Beyond "Detached Concern": the cognitive and ethical function of emotions in medical practice

Jodi Lauren Halpern
Yale University

Follow this and additional works at: http://elischolar.library.yale.edu/ymtdl

Part of the Medicine and Health Sciences Commons

Recommended Citation
http://elischolar.library.yale.edu/ymtdl/3360

This Open Access Dissertation is brought to you for free and open access by the School of Medicine at EliScholar – A Digital Platform for Scholarly Publishing at Yale. It has been accepted for inclusion in Yale Medicine Thesis Digital Library by an authorized administrator of EliScholar – A Digital Platform for Scholarly Publishing at Yale. For more information, please contact elischolar@yale.edu.
BEYOND "DETACHED CONCERN": THE COGNITIVE
AND ETHICAL FUNCTION OF EMOTIONS IN MEDICAL PRACTICE

A Dissertation
Presented to the Faculty of the Graduate School
of
Yale University
in Candidacy for the Degree of
Doctor of Philosophy

by

Jodi Lauren Halpern
November, 1993
ABSTRACT
Beyond "Detached Concern": The Cognitive
and Ethical Function of Emotions in Medical Practice
Jodi Lauren Halpern
Yale University
1993

This dissertation analyzes the ideal of "detached concern" in medical practice. This ideal arises as an attempt to bridge the gap in medicine between managing diseases and recognizing patients "as persons." First, physicians take their emotions to interfere with making objective diagnoses and making every aspect of their practice "scientific." Second, physicians idealize detachment as the stance of the impartial moral agent who is able to care for all types of patients out of a sense of duty. Third, physicians also recognize the need to be empathic; however they conceive of empathy as a purely cognitive capacity that is compatible with detachment.

Chapter one analyzes the features of emotions that contribute to and also threaten rational agency. Chapter two analyzes Descartes' theory of the emotions, which is the outcome of his "scientific" method for understanding reality. Descartes' legacy to physicians is not only the capacity to build powerful mechanistic models of diseases, but the failure to account for human experience via such models.

Chapter three considers the turn to Kantian ethics to restore respect for patients "as persons" to the practice of medicine. Kantian impartiality is shown not
to require detachment. Further, the practice of Kantian ethics in medicine is impoverished when physicians are not affectively engaged.

Whereas chapters two and three show the limitations of the arguments for emotional detachment, chapters four and five give positive arguments for the role of emotions in medical practice. Chapter four examines the cognitive and affective aspects of clinical empathy, and argues that emotions are essential for directing the empathizer to imagine what the patient is experiencing. The final chapter argues that given the importance of emotional engagement and the fact that emotions can obstruct rational and moral agency, physicians need to regulate their emotions without detaching themselves from patients. Physicians can best meet the goals of medicine by cultivating overarching emotional attitudes like curiosity and courage to effectively move themselves towards a more realistic and respectful appreciation of patients.
Acknowledgments

First, I want to thank my dissertation advisor, Karsten Harries, whose curiosity has inspired me, and whose encouragement and guidance have been essential for sustaining the creative spirit. Second, I thank Judith Wilson Ross for carefully reading drafts of several chapters and for encouraging the writer in me. I thank Judith Broder for deepening my understanding of clinical empathy. For kindling my initial interest in philosophy I thank my teachers, George Schrader and Maurice Natanson. For teaching me to believe in my own thinking, I thank my mother Francine Halpern, and my grandmother, Rhea Slominsky. For teaching me how to listen to people's stories, I thank Phillip Krey. Finally, I want to thank my husband Philip Sager, whose support and companionship have sustained me during the years of thinking and writing that went into this work.
CONTENTS

ACKNOWLEDGMENTS ............................................iii

INTRODUCTION .................................................1
   I The Ideal of Objectivity ..................................5
   II The Ideal of Impartiality .................................16
   III Emotional Engagement and Self-Regulation ..........20

CHAPTER ONE: THE INTENTIONAL CHARACTER OF EMOTIONS ..........24
   I The Cognitive Aspects of Emotion ........................24
   II The Volitional Aspects of Emotion ......................36
   III Emotions Determine Salience ..........................40

CHAPTER TWO: DESCARTES AND THE RATIONALITY OF THE EMOTIONS .46
   I The Problem of Irrationality:
      Emotions as Pathogens ................................49
   II Rethinking the Pathogenicity of Emotions ............55
   III The Problem of Arationality:
      Do Emotions Disclose Reality? .........................64
   IV The Essential Role of Emotions in
      Rational Agency .......................................71

CHAPTER THREE: ON KANTIAN IMPARTIALITY ......................89
   I The Conflation of Impartiality and Impersonality
      in Medicine ........................................91
   II The Conflation of Impartiality and Impersonality
      in a Traditional Reading of Kant .....................96
   III Current Arguments Against the Claim that
      Moral Agency Requires Impartiality ..................103
   IV The Core Kantian Ideal of Impartiality as
      Residing in the Autonomy of the Will ...............107
   V A Revisionist Reading of Kant that
      Does Not Equate Impartiality with Detachment ....115

CHAPTER FOUR: THE CONCEPT OF CLINICAL EMPATHY ..............131
   I The Conception of Empathy as Detached Insight .....135
   II The Original Meaning of the Concept of Empathy ......147
   III Three Models of Clinical Empathy as
      Affective Understanding ................................151
   IV A Non-Inferential Model of Clinical Empathy ........168
   V Some Limitations of Clinical Empathy .................180

CHAPTER FIVE: REGULATING EMOTIONS IN CARING FOR PATIENTS...183
   I Emotional Transitions in a Clinical Encounter ....187
   II Emotional Regulation Involves Self-Persuasion ......192
   III Curiosity and the Freedom to Imaginatively
      Reconstruct One's Situation ..........................200
   IV What Affective Standpoint is Needed for Physicians
      to Maintain their Integrity Over Time? ...............214
   V Prospective Conclusions for Future Work on
      the Physician-Patient Relationship ..................226

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.
Introduction

The increasing interest in medical ethics in the United States over the past twenty years arises in the context of widespread public dissatisfaction with physicians. The effective treatments that are the fruits of medical science do not assuage patients' unprecedented concern that physicians will not talk to or listen to them. Patients fear that their suffering will go unrecognized and their dignity will not be respected.

The prevalent ideal that guides physicians in their conduct towards patients is the ideal of "detached concern." "Detached concern" is a complex concept that posits that physicians can detach themselves from their personal emotions, while maintaining a professional concern for patients. Physicians believe that detaching themselves emotionally best meets the special cognitive and moral demands that distinguish medicine from other helping professions.

Twentieth century physicians take the ideals of scientific objectivity and technological reliability as overarching principles for every aspect of medical practice. The current AMA Code of Ethics emphasizes science as the basis of appropriate medical conduct: "A physician should practice a method of healing founded on a scientific basis; and he should not voluntarily associate professionally with anyone who violates this principle." Physicians have interpreted the goal of practicing medicine "scientifically" to require that they turn themselves into reliable
instruments for diagnosing and treating disease, by purifying themselves from the
influence of any personal emotions.³

In contrast, therapists, social workers and nurses, who rarely view themselves
as applied scientists or technicians, have models of professional conduct that allow
an important place for sympathetic emotions. Yet the goal of medicine is the same
as the goal of these other professions -- to alleviate the suffering of human beings.
In the current environment of medicine, physicians translate this obligation to
understand and care for the patient as a sufferer into an obligation to have an
impartial respect for the patients rights and a cognitive awareness of the patient's
feelings.

The ideal of "detached concern" is well-suited to the current climate of medical
care. Most physicians today are trained in hospitals that are like factories trying
to mass produce "health" in the limited sense of repairing body parts.⁴ The
institutional emphasis on technological reliability and economic efficiency are
extended to every aspect of the physician's conduct. If the physician hesitates
during a cardiac arrest she may lose the precious moments necessary to save
someone's life. If she takes time to grieve afterwards, she will not be reliably fresh
and available for the next patient whose body is in need of repair. During surgery
and other invasive procedures physicians avoid looking at the faces of their
patients. And every physician in training is regularly "on call" during which she
gives up not only sleep and food, but her familiar personal environment and
relationships, for endless hours of treating one "case" after another. In this environment, emotional interactions seem extraneous and even threatening to the physician's usefulness.

The increasing anonymity and compartmentalization of medical care has coincided with an increasing emphasis in medical journals on the moral obligation to respect the rights of patients. The AMA Principles of Medical Ethics say that "The principal objective of the medical profession is to render service to humanity with full respect for the dignity of man." The American Hospital Association provides a "Patient's Bill of Rights," beginning with the statement that "The patient has the right to considerate and respectful care." Note the tension in using the language of rights, which arises in contracts between atomistic "entities" where the primary obligation is non-interference, in the realm of considerate and respectful care.

This focus on rights contrasts with a long tradition of emphasizing the physician's beneficent guardianship of the patient's best interests. The ideal of respecting patients as persons has become popular in part out of a rejection of the idea that the physician knows better than the patient what is best for the patient. The recent movement in medical ethics has been inspired by cases in which physicians believed they were making decisions in the patient's best interest, but did something invasive that was tragically at odds with the patient's own wishes. There has also been an increasing concern about the mistreatment of human
subjects. Such concerns, thematized by the increasing involvement of the legal profession in medicine, have contributed to the erosion of the ideal of beneficent guardianship and to the increased emphasis on patient autonomy.

Yet, it is conceivable that respect for the autonomy of patients could be based in an emotionally engaged relationship in which the physician is genuinely moved to respect the patient as a person. It is in the context of a fragmenting physician-patient relationship that respecting patients has been reduced to not interfering with patients' rights. H. Tristam Engelhardt argues that the physician's responsibility to respect patients could not spring from anything but a bare commitment to rational procedure, given the anonymity of current medical practice. What Engelhardt actually shows is not that respecting patients' rights is adequate for recognizing and valuing suffering persons, but rather that such "respect" is the most appropriate moral standard given the detachment that informs institutionalized medicine.8

Just as the pressures of institutional medicine have reduced the richness of the ideal of respecting patients as persons, they have reduced the concept of clinical empathy. Physicians believe that they can understand the subjective experiences of their patients empathically while remaining emotionally detached. In their essay "Training for Detached Concern," Renee Fox and Howard Lief describe the transformation of the physician in medical training.9 Their thesis is that only after medical students go through a period of alienation in which they overcome all their personal responses to patients (in the same way that they overcame their fear and
disgust at dissecting a cadaver) can they develop the special medical skill to listen empathically without becoming emotionally involved.

The ideal of "detached concern" presupposes that, at least theoretically, the basic demands of medical practice for objectivity and impartiality on the one hand, and for a respectful and empathic approach to patients on the other hand, are compatible with the current structure of medical care. The key presupposition is that the physician could recognize and respond appropriately to suffering patients without becoming emotionally engaged. This "could" refers to a conceptual rather than an empirical possibility. Physicians are aware that as human beings they can never actually rid themselves of all emotional responses to patients; the point of such an ideal, like the ideal of objectivity in science, is to provide a standard for which the physician ought to aim. Yet this ideal has important consequences, because physicians take their inevitable personal feelings of grief and affection as marks of failure, and do not question themselves when they are unmoved by the suffering of their patients.

I The Ideal of Objectivity

The physician's first task is to make an accurate diagnosis of the patient's problem. The ideal of accuracy arises in medicine for pragmatic reasons. The efficacy of the physician depends upon "diagnosing," (literally, distinguishing
between) those complaints and symptoms that are relevant to and those that are not relevant to the health of the patient. This goal requires that the physician be wary of judging the importance of things inappropriately, or of missing valuable information.

Physicians have long observed that their emotions can, at times, disrupt the diagnostic process. One need not point to the extreme examples of an anxiety attack or of profound depression making concentration impossible. The physician may be quite able to think, and yet misperceive or misjudge something because of her emotional responses to the patient. For example, she may miss a breast lump out of her own fear of breast cancer, or underestimate the significance of rectal bleeding out of terror at losing a patient to colon cancer. And not only negative emotions but positive emotions can lead one to miss aspects of the patient's situation that less emotionally engaged people would perceive. For example, a physician who admires a patient may be slow to recognize that the patient has a drug problem.

The claim that emotions influence how the world is perceived rests on several interdependent assumptions that I consider in chapter one. First, emotions are not only physiological occurrences or bodily "feelings" on the model of an itch, but intentional attitudes that imbue the world with certain qualities. One is afraid of something fearful, and one is sad about something, even if that includes everything in one's current situation. Second, emotions involve partial depictions of reality.
Like lighting that casts shadows, emotions reveal some aspects of a situation and conceal others. Third, emotions express motives of the agent, so that they may at times be wish-fulfilling.

Because emotions are partial and can be wish-fulfilling, physicians have traditionally been aware of the need to regulate their ordinary emotional responses to people. However, prior to the scientific era in medicine, such regulation was not yet equated with detaching oneself from all emotions. Rather, the physician was to rise "above" selfish and petty emotions out of a compassion that transcends self-interest.

For example, the Hippocratic writings portray the physician as overcoming the lust and greed that interfere with the practice of medicine by developing a special "philia" for all patients. This "friendliness" was not based on an erotic bond with the patient, but on "'physiophilia' or love of universal nature, in its special form of human nature". Yet, however lofty, this philia was an actual emotional experience in which the physician found himself moved. The writer of the Hippocratic work "On Breaths" notes that there are some arts "which to those that possess them are painful, but to those that use them are helpful," and medicine is one of these. The physician "sees terrible sights, touches unpleasant things, and the misfortunes of others bring a harvest of sorrows that are peculiarly his."

Later, under the influence of Christianity, physicians modeled themselves on priests, who were to care for patients out of agape, a transcending, non-erotic love
for humanity. The priestly ideal still lingers in the language of the guidelines of
the American Association of Internal Medicine, which refer to medicine as a
"calling."\(^{12}\) The late nineteenth century and early twentieth century saw the rise
of the gentleman physician, who cared for patients out of benevolent emotions
inspired by a sense of "noblesse oblige." The common theme in these views of the
physician as vessel of natural healing forces, priest, or gentleman, is the idea that
the physician's special compassionate understanding of human nature allows him
to overcome ordinary feelings of resentment, lust, anger, etc. in order to care for
all patients appropriately.

In the Hippocratic writings, the physician struggles to tame his hubris, lust, and
greed because of his overarching interest in *physiologia*, an understanding of nature
based on logos, or "true reason."\(^{13}\) By the nineteenth century, this interest includes
understanding human nature as it really is, a goal that echoes the Hippocratic view
but adds a psychological emphasis. For example, in his tract on the physician-
patient relationship written in 1849, Worthington Hooker says of the physician:

> He sees them [the patients] in their unguarded moments and when suffering
> and trials of every variety ...are acting upon them as tests, searching and
> sure. He sees much that glitters before the world become the merest dross
> in the sick chamber; and he sees too the gold shining bright in the crucible
> of affliction. He sees human passion in every form and condition...thought
> and feeling are often revealed to him [the physician] unconsciously, and the

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.
very fountains from which they rise are almost open and naked to his view, and I may add to his influence also.14

Hooker portrays the physician as a non-judgmental witness to the variety of human passions. The physician's openness and clarity presuppose emotional self-regulation. He neither idealizes patients out of romantic affection, nor denigrates them out of anger or bitter disappointment upon witnessing their transition from gold to dross. Instead, this "disillusioning" exposure to human weakness leads the physician to a certain emotional skepticism. He will not believe too easily his immediate emotional judgments about other persons. After all, he sees apparent bravery and kindness evaporate under the stress of illness, and cannot help but question his initial responses of respect or affection for such persons. Yet the physician does not question the validity of the emotions of respect or affection in general; for as Hooker immediately points out, there will be times when such judgments will be borne out, when the gold shines bright in the crucible of affliction. The physician cannot see the gold without feeling respect or affection; hence a necessary price for appreciating the valuable aspects of human experience is suffering disappointment.

According to Hooker, the capacity of the physician to understand human nature as it really is, is based on his heartfelt appreciation of the patient's position, as well as his wide exposure to human emotion. Hooker emphasizes that the physician could not influence the patient therapeutically if he were not genuinely moved by
the patient's suffering. But given that Hooker is also aware that the physician is an ordinary human being, whose emotions can mislead him into confusing dross for gold, what is to prevent the physician from being mislead? Hooker implies that the physician's exposure to a wide range of human emotion is an educational process that transforms the physician's own emotional reactions. For example, living through various situations with patients educates the physician to overcome ordinary prejudices about what types of persons will respond with courage or fear in the face of illness. However, Hooker never explicitly addresses the issue of how the physician is to guard against being led astray by his emotions.

The very force of our concern about whether the physician can be certain that her emotions have not mislead her from an accurate understanding of the patient's illness was probably inaccessible to Hooker. For this insistence on certainty is distinct to the practice of medicine in the twentieth century. The modern physician has become much more effective than her predecessors by basing her interventions on an understanding of the body mechanism that conforms to the standards of scientific certainty. From a twentieth century perspective, the history of medicine is predominately an account of how wrong conceptions of the functioning of the sick body left physicians impotent in the face of disease and death. The incentive of the modern physician to purify every aspect of her practice from anything that is not testable by scientific methods can only be understood with respect to this history of impotence and failure.
Sir William Osler, "father" of modern medicine, addresses the topic of the physician-patient relationship only seventy years after Hooker, but with the new mantle of the scientific physician. In "Aequanimitas" Osler emphasizes that the physician must strive to control all of his bodily emotions towards patients. His goal is not only to neutralize all outward show of emotion, such as blushing or sweating, but to control the interaction of his mind and his body so that his blood vessels will not constrict, his heart rate not go up when he sees terrible sights. Osler takes the physical state of "imperturbability" to be a necessary condition for the mental state of "equanimity." In addition, "equanimity" requires a reflective understanding of human nature that aims for "clear knowledge" of what human beings are. This clarity requires seeing through the illusory emotional attitudes towards life that ordinary people have. Osler describes a three step process through which the physician is able to achieve "equanimity" by transcending his emotions:

The more closely we [the physicians] study their [the patients] little foibles, of one sort or another in the inner life which we see, the more surely is the conviction borne in upon us of the likeness of their weakness to our own. This similarity would be intolerable if a happy egotism did not often render us forgetful of it. Hence the need of an infinite patience and of an ever-tender charity toward these fellow-creatures. At first glance, Osler's idea that the physician needs to overcome ordinary
emotions to have special access to the "inner life" of patients seems closely related to Hooker's conception. The ideal physician "detaches" himself from sentimental attitudes in order to see persons unmasked by suffering. And the special experience of the physician involves witnessing not only the hidden "weakness" of others, but also his own hidden "weakness." The physician sees human frailties that most people never see, even on introspection.

But consider the difference between Osler and Hooker's pictures of what understanding the patient involves. Osler speaks of "seeing" into the "inner lives" of patients. Hooker speaks of "seeing" human passion, but instead of referring to the "inner" workings of the patient, he speaks of seeing the fountains from which thought and feeling rise. The image of rising from a well-spring refers to origins rather than to mechanisms. Hooker's physician seems to have special insight into the psychological roots of his patients' feelings: their emotional characters, as rooted in their histories.

Hooker does not take the physician to have a theoretical understanding of human nature that is independent of his experiential responses to patients. Rather, he pictures the physician's understanding of the patient as practical: the physician "knows how" the patient feels because of his capacity to imaginatively grasp a variety of affective attitudes. The physician develops a special ability to recognize the expressions of human feeling. He learns to grasp what the patient shows indirectly in gestures and words, even when the patient himself is "unconscious,"
in the sense of unaware, of his own emotion. This emphasis on a practical understanding correlates with Hooker's statement that the fountain of thought and feeling is "almost" open to the physician's view. Hooker's ideal physician seems to have an approximate understanding of, rather than certain knowledge of, his patient's thoughts and feelings.

In contrast, Osler is interested in knowing the mechanisms of his patients' "inner lives." Osler uses the term 'inner life' to refer to a hidden reality that only someone who had achieved a kind of imperturbability could see. He emphasizes that the physician's special knowledge of human nature is not merely a result of his extensive exposure to personal feelings and suffering. Rather, it is because physicians are trained to neutralize all of their own emotions that they can understand human nature objectively. Osler does not trust ordinary feelings of compassion any more than he trusts lust or greed: it is false to see a patient as really heroic, pitiable, or lovable, and hence wrong to be moved by feelings of respect, sympathy or affection. This is because the emotions that disclose such qualities are necessarily transient and tied to the body and social life.

Osler's view is still prevalent. Consider how physicians focus most of their efforts on "objective" disease processes that can be measured and observed by anyone, rather than on "subjective" complaints that do not fit into mechanical models of disease. Whereas the term "health" means flourishing, the patient's "health" is taken to be an empirically knowable condition that is defined as the
well-functioning of the body parts. To say that someone is healthy is to say that she has a healthy cardiovascular system, and a healthy renal system, and so on, but not to imply anything more about her personal condition. A healthy cardiovascular system is a clear descriptive concept based on the measurement of various physiological and biochemical parameters. In circular fashion, equating the patient's condition with measurable physical occurrences justifies the picture of the ideal diagnostician as the observer who will be omniscient about the workings of things.

Osler views the diagnostic process as detective work in which the physician attempts to "see through" both the patient's illusory subjective complaints and the physician's own misleading reactions to the patient\textsuperscript{17}. This reflects the Cartesian assumption\textsuperscript{18} that there is an objectively real disease process behind these subjective experiences that can be modeled mechanically and thus changed through technological interventions. Osler presupposes that whatever really influences the patient's sickness can be known in the same way that the objects of physical science can be known; the model of observing and measuring entities is taken to be adequate for observing and measuring illness in human beings.

One cannot help but think here of the influence on Osler of the daily autopsies that were a routine part of medical practice in his day. The disillusioned tone of his reference to the real "inner" lives of persons echoes the common experience of physicians who not only see their patients die, but then dissect their bodies. This
access to the "inside" of patients, the blood and guts, reveals that human beings are actually physical objects through and through. As living organisms, humans are bound to die and decay. The personal features are lost, and the 'inner' material remains are all that is left of the person.

For Osler, a truthful understanding of human nature requires getting "behind" the veil of all emotional qualities. Emotions in general are false "projections" rather than genuine perceptions of reality. And Osler's concern is not just with the workings of the patient's body, but with the patient's "inner life," which includes necessarily subjective phenomena. He says that the physician can "see into" the patient's "inner life." This presupposes that the physician can set before his "mind's eye" a representation of the patient's psychological life that is entirely independent of the physician's emotions towards the patient. Osler thus extends the ideal of "objectivity," which has shown its utility in the understanding of disease processes, to the overall approach to ill persons.

But Osler's confidence that the psyche could be observed from a detached standpoint is an unwarranted extension of the Cartesian wish to make every aspect of nature transparent. In chapter two I show how Descartes' work on the emotions leads him to see the inadequacy of mechanical models for representing psychological experience. Ironically, Descartes himself is less "Cartesian" than Osler and modern physicians in that he recognizes the limits of detached observation for apprehending all aspects of reality.
II The Ideal of Impartiality

One way of explaining patients' increasing dissatisfaction with physicians as well as the rise of interest in medical ethics is that the Oslerian perspective leaves no room for patients as "persons" in two senses — as affective selves or personalities, and as centers of initiative and value. In chapter three I consider the turn to Kantian ethics as an attempt to provide this missing acknowledgment of patients as "persons." Kantian views are well suited to medicine since Kant faces a challenge similar to the physician's: to explain how human beings can be free moral agents without disturbing a mechanistic causal explanation of nature (including human nature). However, I show how the traditional reading of Kant favored by medical ethicists perpetuates impersonal interactions between physicians and patients.

Kant envisions moral relations independently of affective ties by emphasizing the moral agent's impartiality. There are good reasons for physicians to strive for "impartiality" in two senses that derive from Kant. First, the physician is expected to value all patients as persons, i.e. because they are persons, regardless of her inclination to like or dislike them. She is expected to overcome prejudices towards certain "types" of people that might prevent her from responding with appropriate "concern" to the needs of all patients.¹⁹

Second, the physician needs to be like the Kantian moral agent in striving to
overcome not only narrow prejudices, but all motives that lead her to treat some persons differently from others when their situations have similar moral features. Kant's ideal moral agent is independent of all influences except that of her own reason in committing herself to action. Kant argues that to recognize one's moral obligation, one must consider one's acts from a standpoint that creates an equivalence of value among the needs of all persons. To act impartially is precisely to act from motives endorsed from this disinterested point of view.

This more radical notion of impartiality is well suited to the doctor-patient relationship. The physician needs to be impartial in the sense of transcending narrow self-interest. Consider the example of a physician motivated by lust or greed, who responds attentively to all patients, and hence meets the standard of overcoming prejudice. She is still not impartial in the moral sense. The ideal of an "impartial concern" rules out the physician's manipulative use of the patient as a means to the physician's own gratification, narrowly construed. Such a use of others is partial to one's own interests. The physician is expected to value the patient as a subject in her own right, without ulterior motives. This ideal is expressed by Kant as the obligation to treat other persons as ends and not as means.

However, the idea that one could appropriately respond to the moral features of patient's situations from the detached position is of a piece with the idea that one could understand and effectively act on the medical features of the patient's
situation from the detached position. Both ideas are based on the notion that what is important is a state of affairs that is independent of one's relationship to the patient. The Kantian agent's allegiance to others derives from his allegiance to reason. But respecting patients as persons in medical practice requires an appreciation of the "weal and woe" of such persons, and an "appropriate" response to such emotionally laden events as suffering, dying, recovery, and grief. The problem with fitting the ideal of pure practical reason to medical practice is that it does not account for the moral significance of recognizing and responding to the concrete human situation of patients. Recall Sir William Osler's idea that from the position of equanimity, the physician could have an "infinite patience" and an "ever-tender" charity towards his "fellow-creatures." We have already questioned whether Osler's ideal physician would be equipped to understand how his fellow-creatures felt about anything. It is also questionable whether Osler's physician could value, and hence recognize the relative importance of, the personal experiences of his patients from the position of equanimity? How could he respect them, in any concrete sense? Ordinarily, respect involves perceiving another as courageous in the face of suffering, honest despite the painful facts, etc. Respect also involves particular motives; these show in the physician's attempts to reassure the patient without deceiving her, to help her retain as much dignity as possible while inserting tubes in her body, and not to abandon her in the face of her impending death. These perceptual and motivational attitudes ordinarily are rooted
in the physician's compassionate emotions for the patient.

Given the importance of compassionate emotions for respecting patients as persons, I rethink the linkage of impartial respect and detachment in chapter three. Kant's contention that a moral agent must be independent of all influences except her reason is inherently ambiguous. According to a traditional interpretation of Kant, the independence of the moral agent characterizes her as she acts morally. That is, she must not be moved by anything but her sense of duty in doing what she does. Hence, to be moved by ordinary feelings of compassion is as problematic as being moved by greed. Both lead to actions based on empirical interests, rather than on the independence of pure reason, and hence lack moral worth. In contrast to this traditional reading, recent Kant scholars argue that it is the moral justification for one's actions that must be independent of one's personal feelings; but in acting one can be moved by emotion, so long as what one does receives the endorsement of one's disinterested reason. Building from this revisionist reading, I argue that acting from duty requires acting in accordance with rules that hold regardless of one's transient emotions, but that these rules themselves involve lasting emotional commitments.
Having argued that physicians need not detach themselves emotionally to diagnose diseases "objectively" and to respect patients impartially, I then turn in chapter four to the question of what role emotions play in understanding patients. It is not only a moral imperative, but a practical imperative that physicians understand the "weal and woe" of their patients. Patients do not usually display observable disease processes to the physician, but present a story about how they are feeling. The capacity to understand how the patient is feeling is crucial for making a correct diagnosis. For example, the physician needs to be able to differentiate between lassitude and exhaustion unaccompanied by pessimism and guilt, which might indicate anemia, from the same exhaustion with pessimism and guilt, which might indicate depression. And in order to know how to ask questions to gain more accurate information about patients, physicians need to understand a great deal about how individual patients see the world. As twentieth century physicians have become increasingly occupied with seeing diseases objectively, they have started to describe explicitly what used to be taken for granted -- that physicians need to listen to patients to understand and treat them successfully. A substantial number of pages in clinical texts since the turn of the century describe what being a good listener involves.21 In the past thirty years there has been increasing medical use of the more "specialized" concept of "empathy" where
"feeling into" the patient refers to understanding (via listening and other means) how the patient is feeling.22

In addition to the diagnostic role of empathy, physicians are beginning to come up with scientific evidence that understanding how the patient feels promotes healing. For example, emotions have been shown to be causally efficacious components of illness: grief suppresses the immune system; anxiety exacerbates chronic illness; certain personality traits predispose one to bowel disease. Research has also shown the importance of good physician-patient communication for promoting patient adherence to medications, diet and exercise.23

The view that clinical empathy is compatible with detachment is motivated by the observation that sympathy can obscure physicians' understanding of and effective treatment of patients.24 In chapter four I take seriously the difference between empathy and sympathetic merging, and show why physicians who take on their patients' problems as their own may fail to be empathic. Nevertheless, I argue that "sympathetic" or resonance feelings are essential for directing the physician's imaginative grasp of the patient's situation.

However, there is nothing inherently truth-seeking or respectful about resonance feelings, which are also the basis of mob hysteria, etc. Given that emotions can be concealing and wish-fulfilling, physicians do need a way to regulate their emotions in order to meet their goals of objectivity and impartiality. Yet, what is needed is not a way of detaching themselves from all their emotions,
but a way of using their emotions in the service of the goals of medicine. In chapter five I give an account of emotional self-regulation that does not reinvoke the need for detachment at any level. Instead, physicians can use reflective emotions like curiosity to loosen the hold of emotions like anger and fear that constrict their understanding of patients. They can also use curiosity when they are too disengaged, to better focus their attention on patients' stories. In addition, physician's can cultivate emotions like courage to help themselves endure loss without abandoning their obligation to genuinely care for patients. In doing so, they develop their own moral character rather than divorce their affective selves from their professional roles.

In summary, "detached concern" is medicine's attempt both to bridge and to hide the gap in medicine between managing diseases and recognizing patients "as persons." In chapter one I analyze the features of emotions that contribute to and also threaten rational agency. In chapter two I take seriously Descartes' theory of the emotions, because it is the direct outcome of his project of making nature transparent and hence modifiable. Descartes' legacy to physicians is not only the mechanistic management of disease, but the problem of fitting human beings into this world-picture.

In chapter three I consider the turn to Kantian ethics to restore respect for persons to the practice of medicine. I argue that Kantian impartiality does not require that physicians detach themselves from patients, and that the practice of
Kantian ethics in medicine is impoverished when physicians are not affectively engaged.

Whereas chapters two and three consider the arguments for "detached concern" and show their limitations, chapters four and five give positive arguments for the role of emotions in medical practice. In chapter four, I offer a conception of clinical empathy as a unique form of cognition, in which emotions are used to imagine the "weal and woe" of others. In the final chapter I argue that given the importance of emotional engagement with patients and also the fact that emotions can obstruct rational agency, physicians need to regulate their emotions without detaching themselves from patients. Physicians need to develop their own emotional characters in order to move themselves towards a more realistic and respectful appreciation of patients, thus meeting the goals of medicine.
Chapter One: The Intentional Character of Emotions

In this chapter I consider the features of emotions that have implications for their role in rational agency. I begin with the claim that emotions are intentional, and then consider some attempts to assimilate emotions to intentional acts -- autonomous beliefs, desires and choices. In contrast to these views, I argue that emotions essentially involve being moved by circumstances and have a temporal structure and generality of focus that distinguishes them from these mental acts. By arguing that emotions do not spring from an agent's autonomous reason or will, I emphasize the passivity of emotions, thus aligning with the view that motivates the ideal of "detached concern." But in contrast to this position, I argue here and in the following chapters that the passivity, inertia, and generality of emotions are compatible with an essential role for emotions in providing knowledge of reality.

I The Cognitive Aspects of Emotion

My most basic premise is that emotions are intentional, they are about something of which the agent has at least potential awareness. In contrast, philosophers have argued that emotions are reducible to non-relational feeling states. For example, William James equates emotions with the sensory experience of the physiological changes one undergoes in emotion, so that anger
is awareness of one's accelerated heart beat, etc. In opposition to the Jamesian view, Cannon points out that we cannot identify emotions based merely on the consciousness of any bodily feelings. The same bodily feelings may occur with anger, grief, disappointment, or frustration. Further, if emotions were simply non-relational feelings, like itches, than they would only be contingently connected to intentional objects. But to be afraid is necessarily to be afraid of something frightful, to be angry is necessarily to be angry at something infuriating. Also, if emotions were mere consciousness of bodily feelings, than they would only be contingently related to one's motives and actions. But, the relation between wanting to flee and fear is not contingent. To be in fear includes being disposed towards certain actions and thoughts, rather than others. Emotions are essentially, rather than contingently, related to mental objects and ends.26

How are emotions, which are experiences an agent endures, like the "intentional acts" that have been well characterized by philosophers: beliefs, desires and choices? First, are emotions assimilable to beliefs? Let us consider three properties that differentiate emotions from beliefs in general. First, although some emotions are like beliefs in having propositional objects, other emotions do not have such objects. For example, my being angry implies that I am angry at x for doing (or not doing) y, which entails that I believe that x did (or did not do) y. But my sadness may be much more non-specific. I may be sad about moving far from my home of many years, missing friends and familiar places. Thus a first
difference between emotions like sadness and beliefs is that although sadness is about something, it is not about a matter of fact. Rather, sadness is about an entire situation -- it is a generalized attitude that colors the whole of an experience, like lighting.

Emotions like sadness are less like beliefs than they are like Heideggerian "moods." Heidegger conceives of "mood" as a generalized orientation of one's attention, that unites specific objects and circumstances out of an infinite range of possible objects, into one's own "situation." This leads us to further specify the sense in which emotions differ from other intentional mental processes. The idea of intentionality brings to mind the picture of an arrow with a sharp focus: one believes that the Pope is Catholic. This picture fits certain emotions, like anger: one is angry about a certain rebuff. But other emotions, like sadness or anxiety, may lack a sharp focus, and be about a large field of one's experience, or even one's entire situation in the world.

A mood can be thought of metaphorically as the space one carves out in the world as one's own. A mood is a kind of "mattering map" that relates events and objects by how they weigh on one. Ronald de Sousa captures this notion of emotions as creating a context of interest by drawing an analogy between emotions and a committee chairperson. Emotions create the opportunity for certain beliefs and desires to come to the fore in the way the committee chair steers the agenda by asking certain questions that guide the work of the committee.
If we use de Sousa's image of the committee chair to begin to imagine the cognitive function of the emotions, then it seems that while emotions may not be assimilable to beliefs, they interact causally with beliefs. This leads us to the second issue in our comparison of emotions and beliefs, which is whether emotions interact causally with beliefs in the way beliefs interact with other beliefs?

We expect beliefs to influence other beliefs according to their explicit conceptual relations. There are two aspects of this expectation. First, we expect that the connections between the beliefs that one holds will be logical. That is, if I believe that people swim in the ocean only on warm days, and I believe that people are swimming in the ocean now, I believe that it is warm outside. If I change my belief about it being warm, I then question one of my two previous beliefs. Second, we expect believers to hold onto or reject beliefs insofar as their knowledge remains static or changes. That is, we expect that if one changes one's beliefs, it will be because one has changed other, supporting beliefs.

But emotions do not function vis a vis beliefs in the way other beliefs do. Emotions often persist despite relevant changes in one's beliefs, showing what Descartes calls "inertia." For example, imagine that I feel sad about the misfortunes of other people. I focus on poor Smith, who really got screwed by his corporation. But the belief about Smith is not itself essential to the sadness; if I find out that actually Smith made a fortune out of his run in with his employer, I may either stop focusing on Smith altogether, or think about how Smith's children
have taken advantage of him, thus maintaining my generalized focus on the pathetic. Or consider the example of anger that persists despite changes in belief. My anger may not fade upon my hearing that the harm you did me was completely accidental; I may believe it was an accident, and yet still feel angry at you, for your failure to feel genuinely apologetic. The fact that certain emotional attitudes involve us in seeking out new, relevant beliefs, when the original beliefs no longer merit the emotion, implies that emotions cannot just be beliefs, but rather must be something else that influences beliefs.

Further, emotions not only show inertia with respect to beliefs; they also can subside without any adequate explanatory changes in one’s beliefs about the object of the emotion. For example, consider how one can just stop loving or resenting someone without a prior change in one’s beliefs about that person. Such shifts do not necessarily suggest that the agent is an irrational or disturbed person. They may be explainable in psychological terms that show that the person is acting in an understandable way. Perhaps the person stopped needing to be dependent on a parental figure, and the object of her erotic love was a parental type. Such psychological explanations help us understand how shifts in emotion make sense in the life of the agent. But such explanations are importantly different from the explanations we expect for a change in one’s beliefs. One changes one’s beliefs because one obtains new information that undermines the justification of an earlier point of view. But one can stop loving or resenting someone without such a shift.
in one's information about the loved or resented person. The tragic aspect of someone asking her lover why he has stopped loving her would certainly be undermined if the answer were as simple as citing a new belief.

Thus emotions like sadness, resentment, some types of anger, erotic love, etc. are relatively independent in their temporal course from changes in belief. This fact is significant for our model of emotions. For we cannot simply take emotions to follow from beliefs, as for example, Joseph Fell does. Fell argues compellingly that emotions are not reducible to beliefs because they have an essential passive, embodied aspect. But he goes on to assert that emotions are the bodily response to changes in belief. But this view does not account for the inertia and transience that are typical of many emotions.

Fell's idea is that where there is an emotion, there is a belief that caused it. He has in mind emotion instances like the occurrence of grief on hearing of the death of a loved one, in which the emotion would not have occurred if a particular belief had not been formed. But emotions can occur without beliefs as causes. Emotions are often occasioned only by the occurrence of an emotion in another person: for example if someone addresses one in an angry voice, one will often become instantaneously irritated and angry, prior to judging the other person as rude, pushy or "just like my brother." Fell might argue that in such cases there is a hidden or unconscious change in belief that occurs prior to the responsive anger, which occasions the anger. But, as I discuss in chapter four, experiences of
emotional resonance with other people, such as "contagious" joy, sadness or fear, need not be mediated by belief formation.

So far we have argued that emotions may differ from beliefs in the generality (versus specificity) of their objects, and in the relation of their temporal shifts to shifts in beliefs. A third way in which emotions differ from beliefs is that we assess the appropriateness of emotions in a different way than we assess the appropriateness of beliefs. Patricia Greenspan argues that an emotion may be appropriate when there is insufficient evidence for a justified belief. Her view is that emotions differ from beliefs in that they play a motivational role in rational agency. Emotions allow agents to keep their focus on certain evaluations. They do so by virtue of their essential impact on the agent's level of comfort or discomfort. She defines emotions as affective states of comfort or discomfort that are about evaluative propositions. For example, fear involves discomfort at the thought that danger looms. There are thus two layers involved in any emotion: an evaluative layer that involves a thought like: danger looms; and an affective state that takes as its intentional object the evaluative thought. Thus, in contrast to Fell's view of the affective element of emotion as following the cognitive element, Greenspan's view of the affect as intentionally related to the evaluative judgment entails their temporal unity.

Greenspan argues that because of their essential motivational role, it is proper for emotions to rest on lesser evidence than would be necessary for a belief. She
makes this point through examples, including the following example of wary suspicion in regard for one's own self-interest. She imagines being involved in a business transaction in which the salesman's (X's) body language -- especially the darting of his eyes -- leads her to become uncomfortable at the thought that X is involving her in a bad deal. Yet she has reason to believe, and no good reason not to believe, from other people's testimony, and from X's work, that X is entirely trustworthy. Greenspan's claim is that even if there is nothing about X that explicitly justifies the belief that X is untrustworthy, the suspicion may be appropriate, so long as it is...

'controlled by' some relevant features of the perceptual situation. I might have at least prima facie evidence for belief, if I were able to specify these features at least roughly; but as things stand now, I do not know enough about the 'subliminal' sources of my emotion even to attribute them to its object. I am reacting to something about X's eye movements, say something whose relevance to untrustworthiness could be explained by a developed science of "body language" if there were one.35

So some aspects of the situation are sufficient to merit suspicious feeling even though they are insufficient to justify the corresponding belief that the salesman is untrustworthy. Greenspan thus concludes that "the emotion may be appropriate in a case where its corresponding belief is neither warranted nor held."36

Greenspan's point is not that the emotion is appropriate if it is triggered by any
perceivable aspect of the situation. Rather the emotion must rest on some pertinent perceptual cues that begin to form a pattern that would, if further information followed in an expectable way, be sufficient to substantiate a belief. Greenspan argues that it is essential for emotions to require a lesser, or more partial evidential basis than belief in order for emotions to sustain our focus on what is salient, and to motivate us to do the kind of further exploration that is needed to form pertinent beliefs, and engage in appropriate actions. The key differences for the evidential basis of an appropriate emotion versus a justified belief are that the evidence may steer the agent without the agent recognizing it, and the evidence need not be so complete as to rule out competing views.

Thus Greenspan's view helps explain the common phenomenon of emotional ambiguity in an apparently rational agent. One situation may offer perceptual warrant for 'opposing' affective responses. For example in the above case of wary suspicion it is also conceivable that Greenspan might feel a sense of affectionate gratitude towards X in virtue of his consistently responsible behavior in the business deal, even as she struggles with her sense of wary suspicion. According to Greenspan's view both of these emotions would be appropriate because of their perceptual warrant in X's conduct, even though to hold the relevant beliefs that X was untrustworthy and that X was trustworthy simultaneously would be irrational.

I would go beyond Greenspan in distinguishing the criteria by which we judge the appropriateness of emotions from the criteria by which we judge the
appropriateness of beliefs. It is not just that the propositional content of emotions properly rests on lesser, more partial, or non-verbal sources of evidence. Rather, I take it that the appropriateness of emotions is not something that can be ascertained by looking at the evidence on which the propositional content of the emotion rests. For example, consider a case in which someone becomes overwhelmed with grief because her goldfish dies. Our concern whether this response is appropriate is not about whether this person is correct in judging the goldfish to be dead. Rather we are concerned about whether the target of the emotional judgment is worthy of the response. In describing what is awry we use evaluative terms like "overreaction" that weigh and measure the suitability of the emotional response to its object. When we imagine cases of inappropriate emotional responses, we think of such grief, or fear or anger as overreactions, or of examples of failing to be moved by suffering or threat as underreactions. The evaluative element of our thought here is apparent: the agent is not giving the situation its due. These judgments are like the judgments involved in noting the failure of a person to help another person, or seeing someone as being unduly preoccupied with herself. They are about the value of the agent's response rather then about the truth of the proposition embedded in the emotion.

Greenspan's point that emotions properly rest on more partial evidence than is required for true beliefs is correct but incomplete; she leaves out the characteristic evaluative, non-evidential basis of our assessments of emotional appropriateness.
Consider for example someone who feels intense fear on spotting a schizophrenic person on the street talking to herself. Is this an inappropriate reaction? We could just say with Greenspan that it is not inappropriate, for the same reasons the suspicion of the salesman is not inappropriate. The schizophrenic person's behavior suggests psychosis, and people who are psychotic from street drugs like PCP can be dangerous, so that even though the belief that this person is dangerous is unwarranted, the feeling of fear has perceptual warrant. But such an assessment sidesteps the more central question of whether the agent is overreacting to the schizophrenic person because something is awry in her own attitude towards mental disorder/ nonconformity, etc. To assess this question we would want to know much more about the agent's beliefs and their connection to her reactivity to this street person. Is the agent simply ignorant about schizophrenia, so that with some education she would recognize the patient's disorder and no longer be afraid? If so there seems to be nothing inappropriate about her emotional response. Or would she respond this way regardless of understanding fully that this person presents a relatively low risk to her safety. If so, then it seems that her intense fear would be inappropriate in the sense that the intense grief over the goldfish is inappropriate. I would apply the same questioning to Greenspan's salesman example to tease out whether her suspicious focus on the non-verbal messages given off by the apparently trustworthy salesman represents an unduly vigilant approach to other persons and is therefore inappropriate.
Emotions disclose reality but they do not do so in the way that propositions do. Rather, they allow happenings in the agent's world to weigh on her to different degrees, and thus determine the openness of the agent to various aspects of her situation. In assessing the appropriateness of emotions versus beliefs, we ask something more of the agent, not just something less. We do not ask whether she is able to step out of her partializing perspective, but rather whether her partializing perspective is an adequate approach to the world. For example, we are not asking a man who is angry at (or loves) a woman to justify his anger (or love) in terms of her qualities, as if he could provide evidence that she is worthy of his anger (or love). Rather, we are asking, as in the case of the person who grieves in consolably over her goldfish, whether we feel that this person is losing proper perspective on what matters in the world. In assessing the appropriateness of an emotion, we are thus making an evaluative judgement, rather than a judgment about the truth or falsity of a proposition.

In summary, emotions considered qua judgments, still differ from beliefs in at least three senses: the generality of their objects; the way they interact causally with beliefs; and the criteria by which we assess the appropriateness of emotions. This is sufficient argument to reject strict judgmentalist accounts of emotions that reduce emotions to beliefs, or accounts like Fell's that portray emotions as bodily responses to changes in belief. Greenspan's picture of emotions as affective states that are about evaluative thoughts but that rest on a lesser evidential basis than that
required for belief is compatible with all the above properties of emotions. Further, Greenspan takes emotions to play a special extra-judgmental role in motivating the rational agent precisely because of their inertia and partiality. Let us now consider the extra-judgmental or motivational aspect of emotions in order to fully credit emotions for their role in rational agency.

II The Volitional Aspects of Emotion

In what sense are emotions related to behavior, desires, and strategic actions? Emotions have a sense of purposefulness that is not captured by comparing them with beliefs. If we accept that emotions involve dispositions to act, as for example fear involves the disposition to flee, than we have committed ourselves to accounting for emotions in terms of practical as well as theoretical reason. And in characterizing emotions as directing attention like a committee chairperson, I already imply that emotions straddle our usual conceptions of practical and theoretical reason.

Since I cannot fully consider here the range of conative theories of emotion, I turn to two thinkers who represent the boundaries of these theories, Ryle and Sartre. Ryle38 represents the extreme behaviorist tradition, arguing that emotions are nothing but dispositions to behave. Ryle's claim is that a necessary condition for ascribing emotions to persons is that they behave in discernible patterns. But
while some emotions lead directly to behavior of some kind, others do not. The sadness described above need not show up as behavior. Perhaps the subject tends not to cry or to express sadness in gestures. In that case, if we want to make out a dispositional structure inherent in such an emotion, it will be a disposition to feel badly when the old home town is mentioned, etc. But if the disposition defining the emotion is a disposition to feel a certain way, then the behaviorist goal of defining a non-observable attitude in terms of patterns of behavior is undermined. Thus, the fact that emotions can be analyzed into dispositional attitudes does not support the idea that emotions can be analyzed into behaviors.

Other analytic philosophers attempt to reduce emotions to basic pro-attitudes or functional desires. But de Sousa points out that an essential structural feature of such desires is absent from emotions. When one has a functional desire, one is focused on an immediate goal of some sort. But if one cannot achieve that goal, one can regress through a hierarchy of more general wants, to focus one's desire on another concrete object. For example, if I want to take a walk in order to relax and refresh myself, and I cannot take a walk, I may take a shower for the same general purpose. De Sousa says that such a hierarchy does not arise with emotions. I think this is not quite right. Emotions do transfer from less accessible to more accessible "objects" all the time -- for example, we learn to love in spouses what we once loved in our parents. But in such cases, it does not quite make sense to speak about the new object satisfying the emotion in the way the
new object may satisfy the desire. The transferred love is still haunted by the earlier object. This is because emotions are essentially rather than contingently tied to their targets -- to what they are about. Emotions cannot fit into the construct of "functional desires," which are ahistorical "functions" that are satisfied, emptied out, and open for the next variable.

In the continental tradition, Sartre offers a distinct and influential conative theory of emotion.40 His view is that emotions occur because of the agent's freely chosen desire to preserve her self-esteem. In emotions one does not act efficaciously in the world. One acts "magically," by changing one's embodied state to transform an intolerable situation. For example, one faints in fear and thus "magically" escapes the situation without being effective in the world.41 Sartre's view is that emotions can only effect the agent's situation through such distortion. He posits a radical dichotomy between the world of efficient causality, through which work gets done in the world, and the world of emotion, which only has "magical" or non-causally efficacious impact on the world.

In support of Sartre, there are common examples of "magical" transformations via emotions. For example, one can idealize through admiration or love, overlooking flaws in another, or seeing reciprocal feelings where there are none. Such emotions are wish-fulfilling at the expense of enhancing the agent's realistic appraisal of her circumstances. But this does not show that emotions have no place in the nexus of efficient causes of action in the world. Behind Sartre's
division of reality into efficient causal relations and magical relations is the assumption that the emotion's own occurrence is unconditioned by any causal events implicating the body or the person qua psychologically determined agent. This assumption is essential for equating emotions, as Sartre does all other mental acts, with radical choice. But how can emotions be understood as the chosen actions of an autonomous will, independent of the causal nexus of conventional, embodied beings?

Consider an example given by Robert Solomon, following Sartre: a wife picks a fight with her husband by accusing him of spending too much time with a co-worker she is jealous of; in picking the fight she creates distance between them that prevents them from going out to a party she does not want to go to. So her emotion functions strategically. Is it plausible to understand her jealousy as a radical act of will, undetermined by her embodied/social situation?

First of all, to see emotions as strategic does not require seeing emotions as uncaused. We can make strategic use of illnesses that have obvious physical causes. Secondly, and more importantly, for an emotion like the wife's anger to be operate successfully, there must be a pre-defined context to establish the correspondence between the emotion and the situation she is in. That is, there must be a convention that allows anger to mean what it does and thus to work strategically. Anger serves a purpose in this example because it fits with the evidential basis at hand: a neglected wife responding to her husband's close
relationship with a co-worker. This evidence of a possible betrayal serves as perceptual warrant, in Greenspan's terms, for anger, but not for terror or joy, because of what anger means to embodied conventional beings like us. But it is impossible to see how any situation could offer perceptual warrant for one affect rather than another without relying on bodily/conventional determinants for human behavior.

What we can say after considering Sartre and Solomon, is that emotions are overdetermined. The occurrence of a particular affect now may serve a strategic purpose. But in order to explain how an emotion instance can serve a purpose, we still need to explain how the agent's situation has typical features that warrant that emotion. For we cannot choose unrestrictedly which aspects of life merit anger, joy, fear, etc. Rather, it is only because we learn through social, embodied experience what type of events merit anger that we can use anger in an appropriate way. So for emotions to work strategically, they need to be part of the nexus of causation in the real world.

III Emotions Determine Salience

De Sousa offers a developmental account of emotions that serves as an excellent response to what is missing in the Sartrean account, and also in the
previously discussed attempts to equate emotional judgments with beliefs. De Sousa starts by focusing on how we learn emotions: his view is that people are biologically programmed to have basic proto-affective responses to certain stimuli; but the fact that there are heritable emotional dispositions does not entail that there are full-fledged emotional primitives (as Descartes, for example, presupposes). De Sousa says that "we do need a repertoire of primitive instinctual responses, but emotions are not mere responses." Rather, it is possible for infants and children to learn how to have full-fledged emotions from other persons because their proto-affective responses are triggered by typical situations. Children learn that certain responses are related to certain scenarios: for example, normally children learn that affection is related to being lovingly held; but in pathological cases, affection may relate to disturbed ways of being attended to, including being hurt. De Sousa's hypothesis is this: We are made familiar with the vocabulary of emotion by association with paradigm scenarios. These are drawn first from our daily life as small children and later reinforced by the stories, art, and culture to which we are exposed. Later still, in literate cultures, they are supplemented and refined by literature. Paradigm scenarios involve two aspects: first, a situation type providing the characteristic objects of the specific emotion-type [what the emotion is about], and second, a set of characteristic or 'normal' responses to the situation, where normality is first a biological matter and then very quickly becomes a cultural one. It is in
large part in virtue of the response component of the scenarios that emotions are commonly held to motivate. But this is, in a way, back-to-front: for the emotion often takes its name from the response disposition and is only afterward assumed to cause it.44

This passage makes two important points. First, de Sousa's claim is not only that we learn what emotions are about, but that we also learn how to respond emotionally, from these proto-typical situations of childhood. De Sousa summarizes his view by saying that "the role of paradigm scenarios in relation to emotions is analogous to the ostensive definition of the common noun."45 But it might seem that the ostensive definition of a noun only gives a name to some experience, without teaching one what it is like to experience the thing named. I take it that de Sousa is presupposing here a contextualist view of ostensive definition (he calls himself a contextualist elsewhere). That is, he seems to view the social naming of an experience as of a piece with the differentiating of the experience within the stream of consciousness: to recognize how it feels to be angry is already to recognize some experiential features of anger that tag this experience as the same one that has occurred on other occasions.

Second, de Sousa says that the emotion takes its name from the response disposition, "and is only afterward assumed to cause it." I think de Sousa is right, because emotions could never be learned as they are, if they were divisible into a prior cognitive event and a responsive bodily feeling. For infants and children
would then need to make inferences from their fear about the bogeyman, to the affect associated with fright. But we can find no rules of thought to appropriately ground such inferences (the reasons behind this claim are put forth in chapter four in my argument against an inferential mechanism for empathic imagining).

Rather, de Sousa's point is that emotions can be learned via paradigm scenarios because they are bodily attitudes that inherently portray situations. De Sousa is close to Greenspan here. As noted above, she argues that an emotion is distinguished from a collection of thoughts and bodily sensations by the relation of the bodily sensations to the thoughts: the affective aspect of the emotion is about the evaluative aspect of the emotion.

In summary, I agree with de Sousa and Greenspan that although emotions have essential cognitive and volitional significance that likens emotions to beliefs, desires and choices, the act model of emotion is incorrect. Emotions do not originate in a spontaneous act of theoretical reason or in an independent act of will. Rather, emotions are embodied, learned responses to situations.

In contrast to Descartes, who views emotion as the intrusion of animal nature on rational agency, de Sousa and Greenspan see emotions as essential for rational agency. De Sousa takes the inertia of some emotions to changes in belief, and the generality of focus of emotions like sadness, to show that emotions are conceptually prior to beliefs, rather than physiological responses to beliefs. De Sousa's point is not that emotions happen to direct one's attention but rather that
emotions are necessary for focusing one's attention on a limited set of objects out of an infinite array of possible objects. He says that emotions are needed to "determine salience."

Greenspan makes a similar point with regard to the special motivational role of affect. She says that emotions "'register' evaluations in positive or negative affect," thus exerting an ongoing motivational influence on the agent. An agent in a state of discomfort at an action requirement will have a compelling extra-judgmental reason to act. The physician who feels moved by the patient's suffering will have her own discomfort as an immediate reason to focus on helping the patient, in addition to her other reasons for helping the patient. As embodied responses to one's situation, emotions direct one's attention and sustain one's motivation, and therefore determine a pattern of interests in, rather than a chaotic response to, one's circumstances in the world. Without such focusing, one would lack the capacity to form relevant concrete beliefs in the first place.

Following de Sousa and Greenspan, I take the inertia, partiality, generality and passivity of emotions to indicate their primacy in directing attention prior to any act of judgment or volition. Emotions "give us frameworks in terms of which we perceive, desire, act and explain."47 Emotions can direct our attention only because their efficacy is somehow independent of, and thus deeper than, and irreducible to the rules of thought. This independence from rule governed behavior shows itself in the inertia and partiality of emotions. This depth shows itself in the passivity
and generality of emotions: one's entire situation in the world can "weigh" on one in emotion prior to any activity of deliberating. I now turn to Descartes to argue against the conception of emotions as disrupting rational agency that is still prevalent in medical practice today.
Chapter Two: The Rationality of the Emotions

In this chapter I challenge the Cartesian picture of emotions as irrational that is central to the ideal of "detached concern." I first show that although Descartes account of emotions as rooted in the interaction of mind and body in the pineal gland is often caricatured, it thematizes issues that are still problematic today, particularly in medicine. I consider sympathetically Descartes view of emotions as essentially passive, bodily, conventional responses that cannot be about "things as they are in themselves" in the way scientific judgments can; but I argue against the assumptions, rooted in Cartesian ontology, that these aspects of emotions entail that they are therefore "projections" of imaginary objects into an otherwise clearly given human world.

In his work Passions of the Soul, Descartes describes emotions in terms of the interaction of the soul with the tiny physical "animal spirits" at the locus of the pineal gland. Descartes uses the term "soul" to refer to the quality of being mental or conscious; the emotions straddle the realm of physical interactions (res extensa) and conscious experience (res cogitans). Descartes says that the "animal spirits" are "nothing but material bodies and their one peculiarity is that they are bodies of extreme minuteness." Descartes conceives of "passion" and "action" in animals as entirely lacking intentionality. An approaching tiger causes a movement in another animal's visual receptors that travels to the pineal gland, moving the
"animal spirits," which then cause a muscle movement causing the animal to flee. And Descartes describes pre-reflective "passion" and "action" in human beings as equally mechanistic:

If someone quickly thrusts his hand against our eyes as if to strike us, even though we know him to be our friend...that he will take great care not to hurt us, we have all the same trouble in preventing ourselves from closing them; and this shows that it is not by the intervention of our soul that they close, seeing that it is against our will, which is its only, or at least its principal activity; but it is because the machine of our body is so formed that the movement of the hand toward our eyes excites another movement in our brain, which conducts the animal spirits into the muscles which cause the eyelids to close.50

Descartes' mechanistic account is based on his assumptions that human reflexes are like animal motion, and that animals are no more conscious than machines.

However, in human beings, in non-reflexive reactions, there is an additional step in which the movements of the tiny animal spirits lead to the conscious experience of emotion, prior to mechanistically causing further movements. And this conscious experience is not simply a flashing moment of intuiting a discrete quality, like a flash of color, but involves an interpretation of states of affairs. This is apparent in Descartes' teleological explanation of typical emotional attitudes. Fear is accounted for by the tendency to avoid danger, anger by the tendency to
protect oneself from harm. The explanation of revulsion is that phenomena related to death produce an agitation of the pineal gland "which causes the soul to employ all its forces in order to avoid an evil so present." And Descartes says that the soul is moved to feel joy when it sees the brain pattern that is associated with pleasurable sensations so long as it knows the body is healthy, much the way a theater-goer assured of her own safety enjoys what goes on in a show. The idea that the soul's pleasure is a fitting response to the condition of the body entails that emotions have a meaning structure that involves the depiction of situations in the world.

This intentional aspect of emotions subjects them to appraisals of their adequacy as ways of recognizing reality. The reasons why Descartes takes emotions to be falsifying projections can, for the purposes of our project, be distinguished into two different problems. First, there is the problem of emotions overcoming reason by clouding and distorting the perceptions and thoughts of the agent, on the model of hallucinations and delusions. I turn to this problem of "irrationality" next. Second, there is the problem of "projection" as an ontological state of affairs arising from the unbridgeable gap between the causes of emotion (res extensa) and the objects of emotion (res cogitans). Because of their amphibious nature, emotions cannot fit into the form of explanation that gives everything its place in Descartes scientific world-picture. I consider the implications of this opacity or "arationality" of emotions in the second part of this
I The Problem of Irrationality: Emotions as Pathogens

For Descartes it is not simply because emotions are caused by bodily events that they cannot be trustworthy sources of knowledge of reality. Perceptions, which are caused by bodily events, can be sources of true knowledge of reality when there is a correspondence between the cause and the formal object of the perception. This is not just because the causes of perception are observable in the public realm; for Descartes describes proprioception, in which the bodily events causing the information are initiated inside someone's body, as a trustworthy source of knowledge as long as there is a correspondence between the cause and the formal object of the proprioception. In the case of normal proprioception, for example, awareness of the location of one's arm with respect to the rest of one's body, the object of proprioception is the same inner body parts that originate the movements that cause the proprioception. That is, there are tiny movements in the muscles of the arm that, according to Descartes, cause a chain of movement through the blood, into the pineal gland, resulting in awareness of the arm. Hence, there is a reliability and correctness inherent in perception and proprioception in a healthy body.

Descartes takes those cases in which the relational object of perception or
proprioception does not correspond to the cause of the information to raise the possibility of doubting even the most apparently indubitable, first-hand "knowledge."

is there anything more intimate or more internal than pain? And yet I have learned from some persons whose arms or legs have been cut off, that they sometimes seemed to feel pain in the part which had been amputated, which made me think that I could not be quite certain that it was a certain member which pained me, even although I felt pain in it.52

This example of pain in a phantom limb, in contrast to the example of reliable proprioception in a healthy body, illustrates Descartes picture of "projection" as an epistemological problem: to attribute the pain to the phantom limb is to mistakenly impute physical reality to a limb that has no spatial being but only a subjective presence. The cause of the pain imputed to the limb is not the real limb, which no longer exists, but the body's dependence on already entrenched information. The example of the phantom limb shows that Descartes sees error as arising from dependency on pre-reflective, experiential knowledge or "common sense." To grasp and locate each new pain, indeed to use one's body reliably, depends upon a history of associations of bodily sensations with navigating one's body in the world.53 And such body-knowledge is vulnerable to error because of the influence of past bodily experience on present bodily experience. The "projection" of the phantom limb is possible because body-knowledge has inertia: the old body
experiences persist even when they are unwarranted by the objective world. The agent is fooled by her dependence on her own subjective history of bodily experiences, mistaking a subjective limb-experience for evidence of the objective being of the limb.

According to Descartes, all emotions fit the paradigm of pathological perception illustrated in the phantom limb example. First, just as there is a gap between the experience of pain in a phantom limb, and knowledge of the existence of the limb, there is a gap between the experience of emotion and knowledge of the existence of the objects of the emotion. The objects of emotions are propositions that the person holds in mind, but their causes are the movements of the animal spirits. In contrast, Descartes takes not only beliefs but even "pure" desires to be potentially rational because they are caused by the same mental events that they are about. That is, in the case of pure desires, the agent's own will is both the source of and the object of the desire. This makes it possible for the agent to use introspection to have veridical first hand knowledge of her desires according to Descartes. But in the case of emotions the cause is not mental, even though the object is, so the agent cannot use introspection to verify the attunement of cause and object.

Second, emotions differ from perceptions in that their occurrences do not correspond in a linear fashion to shifts in the external environment or in the beliefs of the agent. Rather emotions impose a structure on situations that is rooted in the prior history of the person. Descartes argues that the inertia of emotions is
essential for their pragmatic role as reflexive behaviors that allow the agent to escape harm; but this inertia also entails that emotions lack the sensitivity and flexibility to modify themselves as the agent's environment changes. This factor leads to the idea that emotions impede perception by causing the agent to view the world in certain ways; and that emotions disrupt reflective thought by causing the agent to posit relationships in the world that have no present empirical basis, like the phantom limb.

Descartes interprets the phenomenon of emotional inertia to signify the passivity of reason with respect to emotion. For example, in his descriptions of the interaction of the soul and body in the cases of joy and revulsion, Descartes pictures the movements of the animal spirits in the pineal gland as projecting an image that acts on the respectively passive soul. Thus, if the animal spirits portray the body as healthy and give the message for "pleasure," the soul will respond with joy, even if the body is not really healthy, and the source of pleasure not really good for the body. This case is analogous to the man with pathological thirst in the sixth meditation, who feels pleasure at quenching his thirst, even if the ingestion of water is terribly dangerous for his body. The source of such deception is the arationality of the animal spirits, which cannot correct their movements to insure a correspondence with objective reality, and the passivity of the soul with regard to the animal spirits. Thus the observed property of emotional inertia prompts the view of the passions as akin to a disease that distorts consciousness,
leading one to falsely "project" situations into the world that have no objective reality, and thus to act in ways that are irrational; I will refer to this view as the model of the "pathogenicity of the emotions."

According to the Cartesian picture, the very "projection" of res extensa into res cogitans that is essential to emotion leads to an analogy between emotions as a class and faulty perception: emotions are like delusions and hallucinations in their unwarranted amplification of the products of consciousness. But in the case of faulty perception what is illegitimate is the content of the "projection", i.e. the phantom limb. In the case of emotions what is illegitimate is the process by which something that originates in res extensa takes on a propositional object, thus insinuating itself in the agent's reasoning but not originating in thought. This seamless imposition of a surd mental process into the agent's consciousness is at the core of Descartes picture of the "pathogenicity" of emotions and of subsequent versions of this view in medicine and psychiatry. Descartes views thought as passive with respect to emotions because emotions do not obey reason, yet they become a part of the agent's reasoning.

To summarize the problem of irrationality: Descartes' view of emotions as falsifying "projections" on the model of human error includes a fundamental assumption about the passivity of reason with respect to emotions. Descartes pictures emotions as diseases of reason in which "ideas" are imposed upon reason by foreign forces. The pathogenicity of emotions does not follow from the fact
that emotions are caused by physical events, for such passivity characterizes healthy perception. Rather, because the "ideas" imposed upon consciousness via emotions are not of thought they do not obey the rules of thought. But because these unruly "ideas" have a propositional structure, they engage with and influence beliefs that have their proper origin in reason. Therefore emotions categorically involve "projection" not only because their genuine causal origins (the movements of the animal spirits) are elided from their intentional objects, but also because they amplify reality without sufficient material cause.

The idea that a surd pathogen could intrude on reason, thus disrupting one's judgment, occurs throughout the Meditations. For example in the "First Meditation" Descartes describes delusions, hallucinations and illusions as the result of black bile acting on the cerebella. In these cases of human error, the mind is caused by surd events to portray the world as it does. And for Descartes these examples of the influence of a surd cause on mental life pose a serious threat to knowledge and action: the delusions occasioned by the black bile then continue to disrupt the agent's judgment. The mental effect of the black bile becomes a model for all human error; the background presupposition is that our thinking is always vulnerable to the mental effects of pathogens given the vulnerability inherent in our embodiment and our reliance on habit/convention. Descartes uses the example of the influence of black bile in the "First Meditation" to justify the skeptical questioning of all common sense knowledge that will lead to a new basis
for knowledge. And the example of the phantom limb in the "Sixth Meditation" is used to illustrate the dangerous tendency to mistake common sense for a true understanding of reality.

II Rethinking the Pathogenicity of Emotions

Let us now leave our analysis of Descartes to rethink the assumptions that lead him to picture emotions as pathogens. Let us grant that there are certain emotional experiences that can lead to errors of judgment and misperceptions. The fact that emotions can lead one astray does not prove that it is the nature of emotional experience in general to overcome or misdirect reason. The Cartesian error, as embodied in the Oslerian view of "detached concern," is to presuppose that the passivity of emotions with respect to external events is perpetuated within the emotion occurrence so that the state of emotion itself exerts a foreign influence on reason. According to Descartes, the tiger causes the fear, and the fear "comes over" or "overcomes" reason. Because the emotion has as its real cause some physical event that follows the rule of res extensa and not res cogitans, the emotion imparts into the agent's mental life the unruliness (from the standpoint of the mental) of res extensa.54

I think that Descartes arrived at this picture of emotions as "pathogens" by conflating two different types of irrationality that emotions are subject to: first,
because emotions are psychological occurrences partially determined by arational causes (the movements of the pineal gland) they can come over the agent like states of drunkenness imposed by an intoxicant; second, because emotions have inertia, they can involve a kind of boot-strapping in which they create their own wished-for objects. Let us consider these two types of irrationality in turn. First consider, for example, alcohol induced rage. One may be enraged about one's job, one's spouse or the price of tea in China. But the real cause of one's rage is the alcohol. Such an example shows that a physical event can cause an affect which then "takes on" a focus in the world, thus becoming an emotion. So even if one's rage is then about one's job, this rage was not originally caused by one's thoughts about one's job. This suggests that surd physical causes can occasion emotions that would not otherwise be occasioned by the person's appraisal of her situation. Yet this example does not demonstrate that emotions themselves have pathogenicity, since it only shows that an organic cause can occasion an emotion and not that an emotion itself causes the agent to make unwarranted evaluations. Certainly organic causes can induce beliefs and perceptions as well as emotions, so the fact that emotions can be so induced is insufficient for supporting the model of emotions as pathogens. Rather, the picture of emotions as pathogens requires showing that it is the emotion itself that acts on the agent as the alcohol does in the case of intoxication.

A second type of irrationality, the boot-strapping that follows from the inertia
of emotions, appears to provide an example of emotions themselves functioning as intoxicants. Consider for example emotional transference, in which one is drawn or repelled by situations that would not otherwise merit such responses. For examples, consider an erotic love for someone who reminds one of an abandoning parent, or a fearful apprehension of a non-threatening situation that rekindles a past trauma. These examples show that in the absence of some physical aberrancy, emotions can involve "bootstrapping," or the creation of their own foci, and thus their own self-perpetuation. Does the fact that emotions create their own foci in such cases show that emotions operate like intoxicants, or pathogens, imposing surd influences on consciousness?

If we consider the basis of neurotic transference, we see that emotions do not operate like intoxicants in such cases. In transference it is not the case that past emotions act like surd physical influences perturbing the sensorium of the agent in the way that the agent is perturbed by alcohol. Rather, older, more remote aspects of one's personality operate like independent agents that exert a persuasive influence on one's here and now functioning self. For example, a fearful attitude that one learned as a child has such occurrent psychological force that it moves one in the here and now to fear what might not otherwise be frightening. One is not invaded by transference emotions, one is moved by transference emotions. The irrationality involved in transference is like the error involved in being persuaded by another person to accept their opinions about a situation rather than relying on
information that has primary perceptual warrant. But this type of irrationality, in which one is persuaded by the charisma of another rather than by one's own senses or rational argument, occurs just as easily for beliefs and actions as for emotions. Certainly one can be persuaded to believe that a skin cream will take away wrinkles, and one can be persuaded to break a rule by a charismatic other, even when one's here and now belief system and code of action would have otherwise inhibited such belief/action.

The irrationality emotions can be subject to because of their inertia is similar to akrasia, in which a person acts on reasons that are not her all considered reasons for action. In both cases the agent's attitudes appear to lack something in the way of rationality. In chapter five I compare Davidson's account of akrasia with emotional self-persuasion to show how in both cases there is a division in mental life, so that the person fails one important standard for rationality -- internal coherence. But in such cases it is absurd to characterize such attitudes as anomalous from the standpoint of the mental since such divisions instantiate psychological "laws," for example the "rules" of persuasion on which rhetoric is based.

In summary, transference emotions do not involve reason being overcome by something surd, but rather one aspect of a person being seduced by another, hauntingly familiar aspect of her psychological life. Transference emotions involve irrationality in that they create their own foci, but they do so in a way that cannot
be incorporated into the pathogenicity model.

However, before we reject the pathogenicity model entirely, it might be argued that a better example than transference for emotions invading reason comes from overwhelming states of emotionality. Descartes uses the example of the onrushing tiger to illustrate the way an external cause might set off a chain of events that intrude upon reason. If an overwhelming emotion is caused by a particular external determinant, can we necessarily infer that the agent is acted upon by her emotion?

Robert Gordon considers precisely this issue, and argues that even in the case of emotions like focused fear and embarrassment, which result from something having acted on the agent, the agent is not acted upon by her emotion:

It is a fallacy to infer, from the assumption that the term 'embarrassment' characterizes a person's state as a product of something's having acted on him, that the resulting state - embarrassment - also acts on (much less 'comes over' or 'overcomes') the person. It is similarly fallacious to infer that a second state of affairs, namely that of his being embarrassed by S, also acts on or comes over him. One cannot properly draw the conclusion

\[ X \text{ [} X=\text{state of being embarrassed} \text{]} \text{ is a state that acts on (a person)} \]

from either of the following:

\[ X \text{ is a state of being acted on in a certain way} \]

\[ X \text{ is a state produced by being acted on in a certain way}. \]
This is a consistent argument showing that even when an approaching tiger scares an agent, this does not entail that the state of fear then acts upon the agent.

However, the pathogenicity model depends upon the idea that there is a physical cause of the emotion which sets into motion a chain of physical events that disturb reason. So a Cartesian argument against Gordon would be as follows: the 'S' that is invoked here refers to the focus of the emotion -- the meaning of the tiger-as-scary. But this 'S' is not the real cause of the emotion. The real cause is the tiger itself, whose movements set in motion a chain of physical events that has as its most proximal element the tiny movements of the animal spirits in the pineal gland. And this chain of causes has no place for 'S', the agent's grasp of the tiger-as-frightening.

And while the Cartesian idea that the meaning of the emotion is an epiphenomenon seems wrong, it turns out to be problematic to assert the opposite -- that the meaning of the event plays an essential causal role in the emotion occurrence. For to say that 'S' is an efficient cause of the emotion occurrence requires separating out the focus of the emotion as a distinct event from the experience of the emotion. But this contradicts our picture of an emotion as an affect that is about an evaluative thought. That is, the evaluative thought registers by weighing on one in a particular way through an affective experience. Thus, while we do not have difficulty saying that certain intentional events cause other related events -- for example, beliefs and desires cause actions, we cannot apply
this model to the relation between evaluative judgment and affect within an emotion occurrence.  

It is clarifying to consider here a less focused emotion than embarrassment or fear, like feeling sad. The focus of one's sadness, which might include all the heavy, disappointing aspects of one's current life, registers a certain way via one's sad feelings. But if we try to impose the language of efficient causality on the relationship of the evaluation 'S', where 'S' is 'the disappointing world' to the emotion event of sadness occurring, 'X', we fall into error. For there is no independent event of evaluating the world as disappointing that interacts causally with one's sadness in the way an event of believing interacts causally with an event of acting.

However, if we return to Descartes' pineal gland account we see that it is equally problematic to elide 'S' from the causal explanation of the emotion. For to divide an emotion into a distinct event of detached perception -- ie. the tiger approaching moves the animal spirits -- and then an affective response, makes it impossible to characterize the experience as an emotion. Whereas a belief and a desire are sufficient to explain an action, no simple addition of belief, desire, and bodily feeling is sufficient to explain an emotion. If one believes a tiger is approaching, and that tigers are dangerous, and if one desires to avoid harm, then one will, barring other conflicting desires or beliefs, run away from the tiger. In such a case, one may also experience physiological symptoms that indicate a
generalized state of arousal. But one need not be afraid of the tiger. What, in addition to "nonevaluative" perception is needed to account for the occurrence of the fear event is a statement about how the perception weighs on the agent. To characterize the experience as fear, entails that one's focus 'S' on the tiger-as-scary registers in terms of feelings of displeasure (including bodily feelings) in the face of the tiger. Thus, while we cannot argue that 'S' is the cause of emotion 'X,' if we elide 'S' from our causal account we cannot characterize the agent's reaction to an event of tiger-approaching as fear.

This account of the grammatical structure of emotions leads us to question how the pathogenicity model ever took hold in the first place: for if 'X' emotion event, is not an independent event from 'S', than the passivity of emotions cannot be pictured as the projection of an unruly idea, caused by the animal spirits, into reason: for neither 'S', nor some detached movement which "projects" itself as 'S' can be the efficient cause of 'X' on the mechanical model. But once this model is rejected it becomes impossible to understand what sense the very notion of pathogenicity makes: how can we understand any physical cause as explaining an emotion? And this unclarity in the notion of 'cause' cannot be solved by saying that mental events that interact causally with physical events are also physical events, so that the real reason for mental event 'Y's' occurrence will be the same event as the physical event causing 'Y.'

For, as we showed above, the motivating mental event that is essential to the explanation of the occurrence of an
emotion is not an independent event from the occurrence of the emotion. And there is no way to capture the idea of the person being acted on by 'S' in the language of efficient causality alone.

In summary, the passivity of emotions is an essential grammatical feature that differentiates the way we explain emotions from the way we explain other mental events like beliefs and actions. A unique and defining feature of emotions is that they involve the experience of being acted on by the something that they are about. Yet if one tries to translate this experience into a mechanistic causal claim, one falls into error, because the focus of the emotion is not an independent event from the emotion occurrence.

Let us reject the Cartesian counterargument to Gordon, which depends upon divorcing the focus of the emotion 'S' from the cause of the emotion, because ultimately it cannot ground itself in a comprehensible account of the causal structure of emotions. So we can accept Gordon's argument as it stands. Thus the Cartesian assumption that the passivity of the emotions entails the passivity of reason with respect to emotions is false. And we have already rejected the Cartesian picture of emotions as intoxicants that perpetuate the anomalousness of res extensa in the realm of res cogitans. We can therefore put to rest the Cartesian worry that emotions act upon/overcome reason, that is at the core of the ideal of "detached concern."
III The Problem of Arationality: Do Emotions Disclose Reality?

Let us now address the second Cartesian worry: that emotions are essentially arational because they cannot conform to the ideal of objectivity. Descartes gold standard for knowledge is the "clear and distinct" ideas of the cogito and mathematical principles. These two very different types of ideas are both non-empirical intuitions; they are standards of truth because they can be held free from doubt given the thinkers direct access both to his own thinking and to mathematical propositions. The direct apprehension of the truth of mathematical principles for us qua thinking substance cannot be achieved for knowledge of nature.

Knowledge of nature is based upon putting mathematics to work to build mechanical models. We can build models of nature because there is, divinely guaranteed, a correspondence between our intuition of extension (mathematical knowledge), and the form of nature, insofar as all of nature has extension, or spatial being. In order to ensure that such models are trustworthy we need to build these models from irreducible building blocks, invoking only efficient causal connections. Only mechanistic models will allow for certainty, because only predictions that are based on mathematical principles alone will be unassailable from later vantage points in time and will allow our knowledge of reality to progress. However, even so, Descartes says that the divine guarantee is not one of an absolute correspondence between our mechanistic models of reality and
reality. Rather, our models of nature can only allow us to make accurate predictions about the workings of nature.

Thus, Descartes envisions two kinds of knowledge of reality, direct intuitive knowledge of res cogitans and res extensa, and indirect knowledge of nature via mechanical models. The goal of fitting the workings of nature into the form of thought (pure mathematical intuition), and the inevitable gap between intuitive and empirical knowledge gives rise to the ideal of objectivity. The ideal of objectivity is an extension of the ideal of indubitability that replaces the coincidence of knower and known with the ideal of an aperspectival grasp of nature secured by the coincidence of the method of applying mathematical models and the measurability of nature. Note that the conception of an aperspectival apprehension of reality is unnecessary for intuitive knowledge; it makes no sense to think of ideas like the cogito in which the object of thought is fully given to itself, as aperspectival or perspectival. Rather, the ideal of the aperspectival observer presupposes a gap between the knower and the known so that a correct approach for apprehending reality fully is needed in the first place. Thus the idea of an aperspectival grasp of reality already presupposes the problem of alternative perspectives, of seeing things otherwise, which does not arise for the indubitable clear and distinct ideas.58

The transition from the intuitive knowledge of the clear and distinct ideas to building models of nature therefore involves a subtle shift from secure knowledge
of reality to secure knowledge of our models of reality. This shift becomes apparent in Descartes' later writings as he points out the problem that the indefinite complexity of nature poses for the human knower. Karsten Harries uses Descartes example of two clocks to make the point that for Descartes our ability to build models that predict natural occurrences "need not mean that the real causes have been understood; indeed, given the infinite divisibility of matter it is very unlikely that our finite models will ever allow us to duplicate nature's processes."59

Descartes writes...

For just as the same artisan can make two clocks which indicate the hours equally well and are exactly similar externally, but are internally composed of an entirely dissimilar combination of small wheels; so there is no doubt that the greatest artificer of things could have made all those things which we see in many diverse ways. And indeed I most willingly concede this to be true, and will think that I have achieved enough if those things which I have written are only such that they correspond accurately to all phenomena of nature.60

Descartes' awareness that the being of nature somehow escapes our mechanical models is revealed most of all in his discussion of the emotions. We have already pointed out that emotions involve propositions about our experience in the world as part of nature, and thus cannot be composed only of clear and distinct ideas, which are non-empirical intuitions. But neither is Descartes' account
of emotion occurrences in terms of the workings of the pineal gland a mechanical model of emotion, as is often supposed. Descartes' choice of terms here - the animal spirits are bodies of "extreme minuteness" that are "indefinitely" small - recalls his language in discussing the related problem of saying that God is his own efficient cause. In a response to Arnauld's objections to his conception of God, Descartes says that "intermediate between efficient cause in the proper sense, and no cause, there is something else, viz. the positive essence of a thing, to which the concept of efficient cause can be extended in the way in which ... the concept of a rectilinear polygon with an infinite number of sides [can be extended] to that of a circle." Just as the indeterminacy of the model of the infinite polygon invites us to think of the coincidence of the polygon and the circle, God as in existence, and as the creator of God's existence, the indeterminacy of the movements of the animal spirits as bodies moving the soul, invites us to think of the coincidence of body and soul. But while we are free to think this coincidence, we cannot adequately conceptualize it, ie. build a mechanical model of this interaction.

This inadequacy poses a major threat to Descartes' ontology, which moves beyond the threat to his epistemology posed by the example of the two clocks. It might appear that as science progresses one could develop more sensitive ways of ascertaining the real causes of the workings of the clock. But the very concept of cause cannot be understood as connecting res cogitans and res extensa, as their interaction in the pineal gland demands. So the pineal gland account, far from
explaining the workings of emotions in nature, points to the inadequacy of any mechanical explanations for accounting for the existence of emotions.

Descartes thus faces the dilemma of failing to explain emotions or acknowledging the inadequacy either of his method for objective knowledge of reality, or of his ontology. His solution is to preserve his conception of objectivity but to challenge his own ontology by affirming the subjective reality of emotional experience while denying that emotions represent things as they really are. In a letter written in 1643, he modifies the list of the basic building blocks of knowledge he had defined in *The Rules* (1628). In *The Rules* he argued that only notions that could be grasped with certainty by the reflective thinker could be the basis of true knowledge: number, extension, and the cogito are examples of such transparent sources of information. But in 1643 he adds to this list a new simple, which is rooted in embodied experience:

Finally, as regards soul and body together, we have only the notion of their union on which depends our notion of the soul's power to move the body, and the body's power to act on the soul and cause sensations and passions.62

As a direct challenge to a strictly dualist ontology, Descartes says that it is unscientific to attempt to explain the experiential union of soul and body in extensional or cognitive terms.

Yet, Descartes is forced by his attribution of reality to emotions and his conviction that we can only know those aspects of nature that can be modeled
mechanically to acknowledge the finitude of human knowledge of reality. He says that even though this experiential knowledge of the passions is "fully" given to us in a dependable fashion, such experiences:

- do not have an objective reality to which a formal reality must correspond.
- Real as modifications of consciousness, they are not otherwise res. A part of my composite nature, outside of which they have no reality, for they are the result of it, the divine guarantee works only in the sense that as teachings of nature they constitute a pragmatic guide to the needs of the composite being. If the union were not real or substantial this role could not be efficaciously fulfilled.  

Here we see that even when Descartes fully acknowledges the reality of emotional experience for human beings, he rejects the idea that emotions are sources of knowledge of things as they really are. Rather, he argues that the modifications of consciousness involved in emotion are merely the result of the composite nature of human beings. The implicit thought here is that knowing things as they really are requires approximating a divine standpoint, free of all bodily and social determinants. Descartes adherence to the ideal of objectivity here causes a rupture in his ontology between 'subjective' and 'objective' reality.

We can now explain the generalized Cartesian claim that the arationality of emotions is based on their "subjectivity." The kind of "subjectivity" that Descartes takes to exclude rationality is not based simply on agent-relativity. The problem
with emotions is not that they must be experienced by someone; for all perception is subject-relative in this sense: seeing presupposes eyes, hearing presupposes ears, etc. And perception is not only contingently agent-relative in the sense that all thoughts, including thoughts about mathematics, must be had by someone. Rather, the information involved in perception is relative to the experiencing subject. That is, whereas the thinking that contributes to mathematics is unrelated to the situation of the thinker, perceptual information always incorporates this situation: seeing is from an angle, at a certain distance, static or changing with the agent's movements. And touch and taste and smell are even more agent-relative. Yet Descartes takes perception to provide true knowledge of reality, when it conforms to certain criteria. This is because, the validity of the information derived from perception and proprioception is based on a correspondence between the cause of the perception and the object of the perception. And this correspondence can be ensured by evaluating perception according to a standard or measure that is itself aperspectival. That is, true knowledge of the location of body parts, and of the dimensions of physical objects, is available through scientific measurement. We can build mechanical models to predict the interactions of body parts and other physical objects. Thus, while perception can be faulty and proprioception can lead one astray, these sources of information are always testable and correctable by scientific reasoning, and thus need not be taken to be impediments to knowledge. But there could be no such standard to serve as a measure for emotional
information. There is no entity that transcends temporal, embodied human experience that corresponds to the intentional object of emotional experience.

In summary, Descartes' claim that emotions provide a pragmatic guide to the needs of composite human beings does not imply that emotions contribute to rational agency. For Descartes rational propositions are those that an objective knower would assent to. By attributing to emotional experience only a 'subjective' reality Descartes is affirming the reality of emotional experience while denying that emotional propositions have representational rationality. Emotions are real effects of our composite being; but qua evaluations they are epiphenomena that have no genuine objects in the real world. Thus, although emotions have adaptive rationality in that they are useful for our composite beings, they are inessential for rational agency in the full-blooded sense.

IV The Essential Role of Emotions in Rational Agency

In this last section I will argue that although emotional judgments cannot provide objective knowledge of reality, emotions are essential for apprehending reality. First, I show that Descartes goal of securing knowledge of reality against the arbitrariness of the human knower cannot succeed; I then show that his reduction of reality to what can be objectively known elides the representation of human beings. Given the Cartesian equation of rationality with the ideal of
openness to all further evidence, it is therefore irrational to overlook the representational rationality of emotions.64

The goal of securing knowledge of reality from the arbitrariness of human interests is at the heart of the ideal of "detached concern." Physicians worry that if their rational understanding of the patient is guided by their emotions, then they will be arbitrary in what they attend to. For example, the doctor whose insights into the patient's attitudes are guided by empathy may very well miss the forest for the trees because of a preoccupation with some irrelevant part of the patient's story, or a blindness to the significance of some other part of the story. The paradigm of "detached concern" posits that such arbitrariness must be eliminated by adhering to a method of inquiry that does not depend upon the attitudes of the observer in any way.

A Cartesian (and as we soon discuss, a Kantian) views reason as self-sufficient not in the sense of providing its own content but in the sense of directing itself. That is, reason may need to make use of other faculties, most notably sensation, but it alone directs inquiry. To the degree that emotions direct one's attention, they enforce an arbitrary order on reason, which cannot be trusted to provide the proper access to reality. One hears echoed in this view the Platonic notion of desires projecting their ends onto practical reason from below, from the animal part of one's nature. The idea is that emotions seek their own ends, and these ends are inessential to reason, and thus untrustworthy paths to information about reality.
The rationalistic assumption of the Cartesian view is that our very attention to reality could be determined by logical rules alone. But this assumption is called into question by philosophers as different as de Sousa and Heidegger, who show that the so-called arbitrariness of emotion is a feature of human inquiry in general. De Sousa does not dispute the Cartesian observation that emotions introduce non-rational determinants into human inquiry, but rather views this phenomenon in light of a different anthropology than that of Descartes. De Sousa takes emotions to play an essential role within the economy of higher cognitive beings, rather than to be a carry over of the "animal" part of our being. In fact, de Sousa argues that less cognitively complex beings, call them ant-machines, could get along perfectly well without emotions, because they have a finite range of concrete interests/ a finite 'world' in which their beliefs, desires and actions could be utterly determined by external rules.  

In contrast to ant-machines, beings of our cognitive complexity, who have an infinite range of possible interests, must determine what is salient to their situation among an infinite set of inferences they could be making. According to de Sousa it would be impossible for such beings to think or act efficaciously unless there was something guiding and maintaining their attention appropriately. De Sousa takes the role of emotions in rational agency to be the focusing and maintenance of motivated attention. He says that emotions "determine salience," they encapsulate experience into organized patterns of importance for the agent. De
Sousa argues that there is no non-affective mental activity that can substitute for emotions in this regard. His argument is based on a version of the "philosopher's frame problem," as described in an example by Daniel Dennett: a robot knows everything about bombs and airplanes, and has all the appropriate intentions to preserve itself, when it is informed that it is in an airplane that has a bomb on it that is about to blow up. The robot decides to leave the airplane, but in fact the bomb is on the robot's own wagon, a fact the robot had stored away, but "it had not 'thought' to draw the inference." When the robot's designers then instruct it to draw the consequences of what it knows, it is busy deducing that "pulling the wagon out of the room would not change the price of tea in China" when the bomb explodes. When the designers tell the robot only to deduce what is relevant to it, the robot is busy ignoring thousands of irrelevant implications when the bomb again explodes. Without fear the robot would have no non-random reason to organize its approach in such a way as to prioritize and act on the relevant knowledge needed to serve its given goal of preserving itself.

De Sousa emphasizes that the "philosopher's frame problem" is not the problem of induction; the issue is not which inferences are valid, but rather, before making any inferences, what clusters of information will be relevant in the first place. He argues that for cognitively complex beings, knowledge of reality presupposes selectively attending to some things rather than others; but "no logic determines salience: what to notice, what to attend to, what to inquire about." And de Sousa
points out that this problem of what to pay attention to is a problem both for factual knowledge and for choosing strategies of action given one's complex set of existing desires. De Sousa points out, for example, the insufficiency of Bayesian decision theory for directing action. Bayesian theory dictates "maximize expected gain." Thus, according to this theory, "a fair bet is equivalent to no bet at all." But there is a meaningful difference between minimizing one's losses (by not betting) and maximizing one's gains (by betting). De Sousa acknowledges that an additional principle could describe this difference; but his point is that this additional principle could not "be dictated by rationality alone."69

In contrast to the robot and the Bayesian gambler, affective beings have resources to handle dilemmas of this sort. Fear will certainly give one the directedness to flee a life-threatening situation; and attitudes like boldness and timidity will influence one to gamble or not to gamble. This leads de Sousa to put forth the hypothesis that "emotions are species of determinate patterns of salience among objects of attention, lines of inquiry, and inferential strategies."70

De Sousa's view of emotions as encapsulating reality into quanta in the way perceptions encapsulate sensory experience is appealing, but invokes the problematic language of mechanisms. But de Sousa is using perception as a metaphor. He says that emotions imitate the encapsulation of perceptual modes. I turn to Heidegger for a deeper and more radical account of emotions as determining attention, which allows us to understand the metaphor of perception.
non-mechanistically.

Heidegger argues that all knowing presupposes attending, and that attention originates in being in a 'mood'.71 Second, Heidegger challenges the Cartesian reliance on the visual model of the knower as a subject observing an object that is before her. These two developments in Heidegger's thought make it possible for Heidegger to offer a radical critique of the Cartesian project of purifying knowledge of reality from the arbitrariness of human existence.

Heidegger emphasizes that affectivity, which he refers to as being-in-a-mood or having a state-of-mind, is a constitutive feature of human being in the world. In direct opposition to the Cartesian premise that thought is independent of affectivity, he states that even "undisturbed equanimity" is an affective attitude, or mood. Heidegger also notes that apparently "pallid" states in which one seems to lack feelings for anything are in fact conditions in which one feels burdened by, or uncomfortable about the oppressive details of one's day to say existence; such a mood is the basis of the possibility of experiencing the opposing mood of joy when one feels free of the burdensome character of existence.72 Heidegger's point here is well supported by the common observation that depressed persons do not always feel sad or anxious, but often feel "flat," with an awareness of the burdensomeness of getting through the moments of the day; recovery is often noted by the observation that the events of the day are flowing together again, without the awful weightiness of time on one's hands.
Heidegger's general claim is that "when we master a mood, we do so by way of a counter-mood; we are never free of moods." In contrast to the Cartesian picture of the passions as disrupting consciousness, Heidegger views the continuity of consciousness as the ebbing and flowing of one mood into another; the cessation of mood altogether is the cessation of consciousness. "The fact that moods can deteriorate and change over means simply that in every case Dasein always has some mood."

Heidegger's conception of the function of moods in human existence is close to de Sousa's. Both take moods/emotions to be the primary basis of encountering oneself-in-a-world, where "world" has the significance of an organized field of interests, as in the ordinary language use of the term "world" to refer for example to the world of baseball. Heidegger says that "mood is a primordial kind of Being for Dasein, in which Dasein is disclosed to itself prior to all cognition and volition, and beyond their range of disclosure."

Let us consider what Heidegger means by the priority of mood, and contrast this notion with de Sousa's idea that human beings would lack direction for thought and action without emotional attitudes steering their attention. Whereas de Sousa comes to his notion of emotions as encapsulating reality by way of the analogy of perception, and especially the analogy of vision, Heidegger makes it clear that he does not see the disclosive function of moods as comparable to the perceiving of reality by a subject beholding an object. The metaphor of the subject looking at
an object, s—o, is entirely misleading when applied to the role of moods in determining conscious life; in fact, it is to avoid and undermine this metaphor that Heidegger explores the phenomenon of affectivity.

There are several aspects of the s—o metaphor that Heidegger challenges in his discussion of mood. First, there is the idea that a mood is equivalent to a psychical condition of the subject. De Sousa moves beyond the view that emotions are inside someone's head, by integrating the embodied aspect of emotion with the intentional aspect of emotion. However he still takes the significance of the emotion to derive from the experiences of a particular historic, embodied subject. For de Sousa the possibility of emotional communication presupposes that the affective meaning residing in one individual conveys messages to other individuals who then may or may not take on embodied affects of their own. This Cartesian atomism is hard to overcome insofar as physiological occurrences are essential to emotional occurrences, and individual bodies underlie physiological events. But as we have seen, there is an insufficiency in pointing to the occurrence of physiological events to explain the occurrence of the emotion. And most importantly the resonance between persons manifested in "shared" emotions in clinical empathy, the effects of rhetoric, the placebo affect, love poetry, etc. cannot be explained using an atomistic event model of emotion. This suggests that it is one-sided to see emotional experience as originating in the psychical or even psychophysical condition of a subject.
Rather than taking emotions to reside in the subject, Heidegger views the subjectivity of the person as residing in the emotion, or mood. That is, Heidegger first rejects the idea that a person exists in the world in the way that an entity exists inside a room. Human beings are not mere things, and no thing could be equivalent to a human world. Rather, the very possibility of consciousness, of a reality for a subject, undermines the possibility of understanding the world as a mere collection of things. For the mode of disclosure of reality for human beings presupposes the problem that human beings are not only here in the world, present in the way that entities are present; for reality to be disclosed presupposes that human beings are also there, ahead of themselves, directed outward spatially and towards the future. And the very spatial and temporal structure of experience, the gap between near and far others, and the temporal gap towards near and far projects, changes with changes in mood. For example, in fear, the scary thing is both very near, and yet not certain to occur, i.e., it has essential spatial and temporal significance.

By taking mood to be prior to cognition and action, Heidegger is equating the experience of mood with the basic structure of human existence as being "there" in the world.\textsuperscript{76} First and foremost, human beings find themselves in a world of near and far others and things, which draw them in and repel them. Certainly, an aspect of the world one is "thrown" into from the start is one's own vital body. But just as one's own physical pain and pleasure is of immediate significance to
one, so is the smile of another person, the threatening thunder and lightening of a storm. The key notion is that these things are given to us at a similar level of proximity. The world is disclosed expressively, and the different styles of expression are what Heidegger means by mood. Thus Heidegger posits that prior to any act of perception or standing over and against the world, being-in-the-world, ie. occupying certain configurations of relations with others and projects, with a certain style, is what is involved in being-in-a-mood.

Heidegger's conception of mood undermines the Cartesian visual metaphor s---o for emotion, and with it the equation of emotion in general with "projection." The Cartesian s---o model of cognition presupposes that intuitive knowledge of our own ideas is independent of our knowledge of the world. One builds from one's intuitions, mathematical models of the external world qua extended substance. As the example of the clockmaker shows, Descartes holds a correspondence theory. The way we build our models may not be the way God has actually produced the external world, but we can know without doubt the contents of our own mind and thus our mentally constructed models of nature. This presupposes that there are indubitable subjective experiences, ie. clear and distinct perceptions, which are transparent to the subject, and thus provide a basis for penetrating the opacity of the external world.

Heidegger's model undermines the primacy of introspection for encountering reality. Heidegger points out that "only because the 'there' has already been
disclosed in a state-of-mind [mood] can immanent reflection come across [psychical] "experiences" at all." Heidegger's claim that mood is a style of being-in-the-world means precisely that mood cannot be originally a psychical state of a subject standing apart from, and looking at, a world.

The strength of Heidegger's account resides not only in rethinking the notion of subjectivity but even more in rethinking the notion of the world as an affective world. Heidegger says:

Having a mood is not related to the psychical in the first instance, and is not an inner condition which then reaches forth in an enigmatic way and puts its mark on things and persons.

Heidegger's use of the term "enigmatic" is especially important here. The affective significance of things in the world remains mysterious when affects are understood according to the s—o model. That is, the Cartesian s—o model presupposes that there is a human structural configuration in the world -- a practical reality of being "there" that is independent of affectivity. This assumption, which plays a crucial role in the Sartrean account of emotions, given Sartre's adherence to the Cartesian s—o visual paradigm, has been labeled by Joseph Fell "the two-world hypothesis." The idea is that one could negotiate the social world independently of affectivity. For Sartre, emotions are "magical" precisely because they are subjective projections in which one remakes the world in one's own terms, rather than negotiating the world in a strictly rational, efficacious fashion. Consider the
background assumption of a typical Sartrean account of an emotion: in fear of a scary face in the window one may faint, thus "removing" the scary threat. The background assumption is that whereas one has "acted" in one's emotional "world," in the world of efficacious relations with objects, one has done nothing to protect oneself from the potential assailant. There is a radical discontinuity between the two worlds of affectivity and efficacy.81

We have already indicated a major weakness of the "two-world hypothesis" in our criticism of the Sartrean account of emotions. If emotional qualities are subjective projections onto situations that could be viewed in some non-affective way, than it seems there are no reasons for certain qualities to be attributed to certain situations. But then how can we account for the typicality of emotion, for the fact that human beings have typical responses to being touched, falling ill, having children, etc.? How can we account for the style of a thunder-storm as characteristically scary, and the style of an infant as characteristically adorable (a problem Sartre sought to explain)? If one retains the Cartesian s---o visual metaphor for emotional apperception of reality, one cannot help but attribute to the subject the activity of bestowing affective meaning on the world. But Heidegger contrasts this notion of bestowing affective qualities onto reality with the notion of an originally affective world structure, prior to egological affectivity. The Heideggerian notion that affectivity arises first in the actual style of human being-in-the-world, the way human beings move and respond to each other and things,
rather than in subjective ideas, helps us escape the "projection" metaphor that is at the heart of Descartes' view of emotions.

In using Heidegger to deepen our understanding of de Sousa's idea that emotions determine salience we face the following gaps between the two conceptions. De Sousa's analogy of perception for emotion brings us to envision persons as depending upon their emotions to sense the humanly relevant features of situations. Heidegger leaves us with the more radical idea that affectivity is the material out of which relevance is built in the first place. Heidegger emphasizes that the possibility of anything mattering at all depends upon the condition of human being as being-in-a-mood. This point is often misunderstood as an overly idealistic claim about human experience that ignores the material, bodily facts of human life. But such an interpretation misses the point that Heidegger's argument is operating at a very different level from an account like de Sousa's. De Sousa's claim is a teleological claim about the actual species, human being. Heidegger's claim is an analytic claim about what is essential to Being, rather than what is characteristic of human beings. Heidegger is not making the unpalatable claim that all the material configurations of beings in the world depend upon affectivity. Rather, Heidegger's point is that it is only because human beings find themselves in a world rich with expressive meaning that anything can matter at all. The world moves us, and we can therefore express emotion.82

Heidegger offers a radical critique of the Cartesian ideal of eliding arbitrariness
from knowledge of reality. Heidegger's claim is that the very possibility of conceptual knowledge presupposes an affective way of being towards things that involves directing one's attention. But this should not be interpreted simplistically to mean that any singular emotion occurrence could determine salience, as de Sousa seems to imply in his account. Rather, determining salience involves not only directing one's attention, but maintaining possibilities beyond one's given direction, that will be the basis of future directions. An auditory metaphor is more useful than a visual one here: one must not only attend to the theme at hand but must anticipate future themes as they are foreshadowed, and must recognize patterns of recurrence. Far from viewing mood as a mechanistic way of encapsulating information, Heidegger views mood as an indication of the human condition. The problem of having to determine salience prior to being able to define things ostensively and know things as matters of fact reveals the fact of human freedom. Heidegger says that man is the being whose being is in question. He means that human beings must determine what matters in the world by what they attend to, and what they do.

Thus, while the occurrence of emotions may be considered arbitrary in the sense that such occurrences are prior to deliberation, they are not arbitrary in the sense of meaningless. Rather the very arbitrariness or openness of human attention and the necessity of affect for responding to this openness by "clearing" a space for particular projects are necessary conditions for the possibility of human

...
knowledge.

Our detour into Heidegger was meant to deepen and substantiate the idea that emotions are essential for determining salience, and to undermine the Cartesian hope that human knowers could ever overcome the "arbitrariness" of human knowledge. By linking knowledge to interested activity in the world, Heidegger undermines the divorce of representational from adaptive rationality that is the basis of Descartes view of emotions as arational.83

We have rejected the Cartesian goals of eliminating the arbitrariness of the human knower from knowledge of reality, and of divorcing representational from adaptive rationality. But this does not entail that emotions yield knowledge of an objective state of affairs. Descartes correctly saw that unlike non-affective perception and reasoning, emotional "judgments" could not fit into his scientific world-picture. This problem is very significant for physicians, because the essential elements of Descartes world-picture remain the pillars of current medical science -- the assumption that all events can be explained within an efficient causal framework, and the assumption that these interactions could be modeled mathematically. Now, as then, the scariness of a tiger cannot be captured by any aperspectival representation. Even the most universalizable affective objects, like the warmth and security represented by a mother and child, the erotic feelings represented by two lovers, the threat represented by an angry face are concepts that cannot be divorced from their social origins. Bernard Williams calls moral notions
like honesty and courage that are semantically embedded in their social origins "thick concepts." Their meaning cannot be divorced from the specificity of human cultures and human values, even though such representations are so essential to the idea of humanity that we cannot think "human being" without apprehending these emotional objects. The point is that we cannot speak sensibly of an aperspectival view of such objects, precisely because in a world without human values, these emotional objects are meaningless.

If, as Descartes claims, we cannot speak sensibly of an aperspectival view of the objects of emotional experience, is it rational to exclude emotional evaluations as trustworthy sources of knowledge of reality? I argue in chapter four on empathy that our own emotions are essential for revealing other people's emotions to us. Descartes' account allows a place for empathy as a kind of pragmatic experience, but excludes the possibility that empathy is genuinely revealing. Thus Descartes forces us into a peculiar affective solipsism. Instead of accepting this bizarre conclusion, we might say that it is because our empathic emotions reveal the affective lives of others that they help us function adaptively in the human world. The essential role of our own emotions in revealing the existence and nature of emotions in others entails that emotions are essential to rationality in the full-blooded sense of openness to things as they really are.

The problem of explaining emotions reveals a tension between the concepts of rationality and objectivity that Descartes bequeathed us. For Descartes, the only
way to judge our perceptions of reality as rational is to give objective explanations of why we see things as we do. For example, the perspectival phenomena of vision become trustworthy insofar as there can be an aperspectival account of vision that includes an explanation of perspective and its limits -- the science of optics.85

We have opposed this tight linkage of rationality with objectivity by arguing that all knowing presupposes human interests/being-in-a-mood. But this does not mean that we can give an alternative conception of rationality that does not include some notion of transparency -- of openness to all further evidence. Although, as I discussed in chapter one, we appraise emotions to some degree the way we appraise values, we cannot judge the rationality of emotions without also considering whether the agent's perspective in the emotion is sufficiently open to things as they really are. So for example we cannot assess whether X's fear of Y is rational without considering whether, from some alternative perspective, free of X's biases, there is independent evidence that Y is dangerous. Thus even though the representational rationality of an emotion cannot be grounded in the ideal of objectivity, it is still guided by an ideal of openness to further evidence that is haunted by the wish for transparency.

In summary, the picture of emotions as pathogens is still influential in medical practice today because Descartes' basic assumptions are largely unchallenged: the belief that human nature can be understood scientifically and the belief that a
scientific explanation could be framed in the language of efficient causality alone. Thus the wish to make knowledge of reality transparent exerts a totalitarian force that excludes important aspects of reality, like the emotional lives of other persons. But we have seen that this picture of rationality fails to be guided by an ideal of openness to further evidence, and is thus itself irrational. We can therefore reject Descartes' reduction of our knowledge of reality to what can be modeled mechanistically, and with it his conception of emotions as inessential to rationality. I turn now to the other pillar of the ideal of "detached concern," the Kantian conception of the impartial moral agent.
Chapter Three: On Kantian Impartiality

In response to the reductionistic approach to patients in current medical practice, physicians and patients have turned to ethics to restore what is missing. The prevalent ideal in medical ethics discussions is the Kantian concept of an impartial duty to respect patients "as persons," to treat them as "ends-in-themselves." The idea is that the physician is to have a duty-based commitment to act in light of the patient's "subjective" ends as well as a commitment to the objective task of repairing the body.

The ideal of respecting patients' own choices and preferences contrasts with the possibility of using patients for projects that are not their own, i.e., treating them manipulatively. The suitability of this ideal for concrete medical practice is apparent when one considers the opportunity physicians have to treat patients as means to other ends. Physicians are keepers of esoteric, sought after knowledge and skills that give them power in society. And they are trusted to invade bodies and minds, and to make life and death decisions for other people. The ideal of respecting patients as "ends-in-themselves" directs physicians to promote the concrete ends of patients, and not merely to use patients as "material" for physicians' own projects. Given that such projects can range from selfish financial gain, to attaining scientific knowledge, to promoting the health of society, as the physician sees it, the ideal of respecting the ends of the patient curbs not only
greed but also benevolent paternalism.

Kant's emphasis on the impartiality of the moral agent, whose duty to respect patients is not based in personal ties, has made his views particularly relevant to physicians today. However, in equating impartiality with affective detachment, physicians both oversimplify Kant's conception of moral agency and overlook the particular duties physicians have to recognize and respond to suffering persons. In this chapter, I will take a closer look at the Kantian conception in order to distinguish the concept of acting from an impartial sense of duty from the picture of the detached or impersonal moral agent.88

According to a traditional reading of Kant, acting from duty requires detaching oneself from all affective motives. The commitment to respecting persons as persons is taken to follow from principles of reason, rather than from affective responses to particular persons. The traditional Kantian conception supports the ideal of "detached concern" because detachment from affective, historical relationships to other persons is taken to be both necessary and sufficient for respecting persons as "ends in themselves." But I will show that the idea that respect for others is duty-based, hence impartial, does not entail the detachment of the moral agent.

The traditional reading of Kant is currently being challenged by revisionists who aim to show that sympathetic emotions are compatible with, and even valued within, Kantian moral theory. In the second part of this chapter I turn to Barbara
Herman, Onora O'Neil and Marcia Baron to challenge the traditional idea that acting from duty, in a Kantian sense, requires detachment from sympathetic emotions. However, the revisionists do not fully overcome the equation of impartiality with impersonality insofar as they retain the idea that rational reflection alone is "sufficient" for morality. I argue that to the degree that the revisionist account retains this idea of reason's "sufficiency," it loses its appeal as an account of how moral motivation can be an integrated aspect of the agent's overall personality. But to the degree the revisionists take the agent's formal commitment to other persons to be derived from her affective, historical relationships with others, they leave behind the core Kantian idea of pure practical reason: the idea that there is one reason with one sets of ends (universality and consistency) that rules in both the theoretical and practical sphere. I therefore argue that the Kantian ideal of impartiality is separable from the Kantian notion of pure practical reason, and that the former but not the latter is useful for medical practice.

I  The Conflation of Impartiality and Impersonality in Medicine

The Kantian conception of respect for persons is well-suited to the medical profession because it expresses the ideal of an impartial or duty-based response to all persons. Physicians are expected to have a role-related commitment to respect anyone who is their patient, even if they do not approve of his or her behavior or
attitudes. There are two aspects of this expectation: first, it is presupposed that there is some formal property of persons as persons, or as suffering persons (patients) that can merit this respect; second, the physician's respect for patients is taken to be duty-based, and thus required, rather than an optional, though fortunate, occurrence.

The Oxford English Dictionary defines "impartiality" as "freedom from prejudice or bias; fairness." This definition does not refer to detachment. While the ideal of an impartial judge brings to mind the picture of a detached thinker, this follows from certain assumptions of legal theory, rather than from the definition of impartiality. The concept of impartiality is also used in contexts where a detached attitude would be inappropriate. For example, a parent can strive to be impartial in the treatment of all of her children. She may be free of bias, and fair, yet passionately concerned about each of them. The parent can be understood as caring for her children because they are her children, and thus in virtue of a formal, duty-based commitment to them. (However, we would have concerns about the adequacy of the moral perspective of a parent whose care for her children was only out of a sense of duty; I address a case of a spouse having this perspective at the end of this chapter).

Yet in medical practice, the Kantian ideal of impartial respect for persons is equated with detachment because of the additional Cartesian assumption that emotions are sources of prejudice and compulsion. The kernel of truth in this view
is the fact that overwhelming passions, like lust, greed, and fear of death can interfere with the physician's capacity to care for all of her patients impartially. For example, a physician may perform overly invasive life-extending procedures on a dying patient because she identifies with the patient and is very afraid of her own death. She may fail to act respectfully towards a patient who has committed an act she finds morally reprehensible, out of hatred for a parent who did something similar. Or she may spend too much time with a patient she is attracted to, to the detriment of other patients.

However, as I discussed in the introduction, it is the fact that physicians are increasingly estranged from, rather than too involved with, patients, that makes an ethical view that minimizes personal ties so appealing. Physicians narrowly construe the ideal of impartial respect for the "ends" of the patient to mean respect for the legal "rights" of the patient. However, there are three problems with this reduction of valuing the patient's ends to not interfering with the patient's rights. First, whereas the obligation to respect someone's "rights" is definable in terms of permissible and impermissible behavior, respect for ends cannot be so defined. Physicians show respect for the ends of patients not only when they give informed consent, and share decision-making, but also when they strive to be sensitive to the patient in delivering bad news, and when they are careful not to embarrass the patient on medical rounds. But it is a poor use of the term "rights" to speak of the patient's right to be touched and comforted, listened to and encouraged.
Second, the idea of respecting "rights" is rooted in a liberal conception of the person that sees morality as aiming for non-interference with, rather than connection with, other persons. But physicians are expected to engage with patients, and thus cannot have as their overarching principle the ideal of non-interference. Katz and others have shown that the adversarial legal paradigm is ill-suited to capturing the moral dimensions of the physician-patient relationship. The relevance of the Kantian ideal of treating persons as ends-in-themselves for medical practice depends upon seeing the sense in which this ideal captures more than the principle of non-interference.

While Kant did not directly address the issue of "rights" in the *Groundwork of the Metaphysics of Morals*, (hereafter the GMM), much of his discussion of the formula of universal law (FUL) does embody a liberal conception of moral obligation as a negative constraint (non-interference): one is not to will in such a way that what one wills would prevent others from willing the same. However, the formula of treating humanity as an end-in-itself (FEI) pictures moral obligation from a different perspective than the perspective which generates the FUL. Onora O'Neil points out that the FUL is a response to the question of what rational agency itself commits the moral agent to; the FEI is a response to the question of what rational agents are committed to, given that their acts condition the agency of other persons. O'Neil equates the conception of non-interference represented by the FUL with one aspect of the FEI: the ideal of never treating others merely
as a means. To will a maxim that others could not also will would be to act under a conception that omits the rational agency of others, and thus treats others merely as a means to one's own agency. O'Neil argues that the FEI has an additional, positive sense: one is to promote the concrete ends of other persons, so that rational agency can flourish. I agree with O'Neil, given that Kant does explicitly distinguish between a negative and a positive sense of the FEI:

Now humanity could no doubt subsist if everybody contributed nothing to the happiness of others but at the same time refrained from deliberately impairing their happiness. This is, however, merely to agree negatively and not positively with humanity as an end in itself unless every one endeavors also, so far as in him lies, to further the ends of others. For the end of a subject who is an end in himself must, if this conception is to have its full effect in me, be also, as far as possible, my ends," (GMM, 430).

Third, whereas "rights" language is better suited for the bureaucratic interactions of strangers, the Kantian ideal of respect for persons is meant to guide one's personal relationships as well as one's impersonal obligations as a citizen. The Kantian project aims at isolating the moral commitment that is to be at the heart of private as well as public morality. Further, in contrast to utilitarian thought experiments, Kant's examples are not framed from the viewpoint of the detached policy-maker who expresses a generalized commitment to "humanity", but from the standpoint of ordinary individuals faced with personal dilemmas: whether
to steal, to help someone, to commit suicide. Through such examples, Kant focuses on such personal obligations as the obligation one has to oneself to preserve one's own life (the suicide example). In fact, the Kantian conception does not divide the moral sphere into personal versus impersonal commitments; and this fact makes Kantian morality relevant to physicians, whose special obligation to take-up the ends of patients cannot be so divided. In summary, the usefulness of the Kantian picture of moral agency for physicians depends upon a richer translation of Kant's ideal of respect for persons than the present version of not interfering with patient's rights.

II The Conflation of Impartiality and Impersonality in a Traditional Reading of Kant

I turn now to a traditional reading of Kant to consider how the impartial standpoint of morality comes to be equated with the impersonal standpoint of the pure rational agent. In this section I consider the core arguments that support the traditional Kantian view that detachment is necessary for moral agency, to show that there is a gap between the requirement of impartiality and the requirement of autonomy that leaves room for revisionist readings. I then consider the idea that pure reason alone is "sufficient" for moral agency, to show that this idea is what leads to the impersonality of the Kantian moral agent.
In the GMM, Kant contrasts the case in which a man helps people out of sympathetic emotion with the case in which the man lacks all sympathetic emotion, yet helps others out of a sense of duty (GMM 398). Kant says that only in the latter case does the action have moral worth. This contrast is traditionally taken to show that acts from sympathy cannot have moral worth. The traditional reading gains support from Kant's claim that sympathetic feelings cannot be the basis of the duty of beneficence towards others. Kant says that "sympathetic sadness...would ... be an insulting kind of beneficence, since it expresses benevolence with regard to the unworthy, called pity, which has no place in men's relations with one another" (Doctrine of Virtue, section 34)\textsuperscript{92}. And Kant's claims that morality is "not the mouthpiece of laws whispered to her by some implanted sense," and that empirical motives are "highly injurious to the purity of morals" (GMM 426) are traditionally interpreted to mean that sympathy is at odds with morality.

Some interpreters take Kant to see sympathy and duty as essentially incompatible because of his anthropological assumptions. In The Anthropology Kant portrays emotions as selfish, corrupting forces that impede one from reflecting on one's acts and hence from taking other persons into account.\textsuperscript{93} According to Robin Schott\textsuperscript{94}, Kant's stoic conception of emotions leads him to equate all emotion with hedonistic passions that are narrowly self-interested; given this picture of emotions, the sympathetic man is as partial and selfish as the greediest hedonist,
seeking pleasure rather than acting from duty. According to Schott, Kant's preoccupation with pure practical reason follows from his rejection of the temporal, bodily aspects of human nature. The value of Schott's reading is that Kant's preoccupation with finding a secure, unchangeable, apriori foundation for morality is explained in terms of his own aversive response to bodily experience -- thus deconstructing the purity of the Kantian project.

Kant's remarks that sympathy "stands on the same footing as other inclinations" (GMM 398), and that "sympathia moralis [is] really sensuous feelings of a pleasure or pain at another's state of happiness or sadness" (DOV, sec.34) show that his views of sympathy are rooted in his larger view of sensuous input. Schott's reading brings into relief one major tendency in Kant's writing: his idea that sensibility, including perception and affect, is "like a mob of people since it does not think" (Anthropology, sec.8), which must stand before the tribunal of the understanding. At times Kant seems to see the non-judgmentalness he attributes to sensibility as dangerous in practical matters and as requiring not only "processing" by the understanding, but also stern control. In the GMM, Kant conceives of duty as a constraint against unruly forces. Kant portrays the motive of duty as a counterweight opposing all of man's "needs and inclinations, whose total satisfaction he grasps under the name of 'happiness'." Morality requires that Reason "enjoins its commands relentlessly, and therefore, so to speak, with disregard and neglect of these turbulent and seemingly equitable claims" of
emotion. (GMM 405). Yet in *The Anthropology* as well as in the moral writings, Kant also argues that because sensibility is, by definition, non-judgmental, it cannot be responsible for confusing, compelling or deceiving the understanding (sec.'s 9, 10 and 11). He cites Cartesian type examples of sensory illusion to show that it is not sensibility that disrupts the understanding, but the understanding which confuses itself, by mistaking "the subjective for the objective" (sec.11). In the GMM, Kant is not indicting sympathetic emotion, as Schott suggests, but indicting the understanding for failing to see in its own sole responsibility for judgments the sole source of justification for morality.

An additional problem with Schott's reading is that Kant did not take the stoic dichotomy between sympathetic emotions and moral action to be self-evident to the audience he addressed. In Kant's social world, and in the works of his philosophical predecessors including Hume and the moral sense theorists, the commitment to value the ends of others was presumed to come from sympathetic emotions. Kant himself was quite influenced by this view, as is shown by the passage following his discussion of beneficence in the *Doctrine of Virtue*, where Kant discusses the practice of impartial morality in concrete experience: Kant urges that one visit poorhouses and hospitals to develop one's sense of sympathy in order to be better disposed to do one's duty (DOV, sec.35). In the GMM, Kant argues against "common sense" in challenging the notion that the moral obligation to take-up the ends of other persons is rooted in emotion. Like Descartes in *The
Meditations, Kant attempts to purify "common sense" by rationally reconstructing ordinary concepts, like duty and respect.

In the GMM, Kant's equation of impartiality with detachment comes in his arguments that only acts done from duty are morally worthy, and that acting from duty cannot be heteronomous. Kant's argument is not that sympathy necessarily blinds one or disables the person from doing what duty commands, but that sympathy is itself blind or indifferent to considerations of duty. There are two senses in which sympathy may be considered indifferent to duty. The man who acts because of sympathetic emotions might help someone whom he should not help, if he considered his act from the perspective of universal law; for example, he might assist a murderer escaping from prison. Thus the general motive of sympathy is only contingently related to actions that accord with duty. The second sense in which sympathy may be considered indifferent to duty is that one's intentional object in a sympathetic act is usually the suffering of others and not the rightness of one's actions. The traditional reading of Kant seems to require that the moral agent acts because of her intention to do her duty, and the sympathetic man may not have such an intention.

The idea that morally worthy acts must be done from the motive of duty alone depends upon the Kantian claim that the moral standpoint is the standpoint of "autonomy," where "autonomy" is taken to require that one have no empirical motive to do what one does. According to a traditional reading, the agent's only
subjective motive for acting morally is "reverence," which Kant calls a "practical," as opposed to a pathological emotion, because it follows upon the moral motive, or expresses the moral motive in the empirical world, but does not move the agent to act.\textsuperscript{95} But Kant's definition of reverence rules out the essential features of sympathetic emotion: reverence is self-willed, rather than something that moves the agent, either in the sense of causing the agent to act, or in the sense of being experienced as an involuntary occurrence; reverence is independent of empirical sensation, including ordinary feelings of pleasure or pain; reverence cannot have as its object the weal or woe of another person, but only considers the person insofar as they exemplify a law. So despite Kant's use of the term "emotion," reverence is an attitude which lacks the three criteria of altruistic emotions: first, they have motivational force; second, they are felt occurrences, either via bodily disturbances, or shifts in consciousness; third, they take as their formal object the weal and woe of other persons.

Under this conception, Kant's definition of duty as involving the ruling of pure reason over empirical inclinations is taken to require the replacement of empirical inclination with a pure interest in morality. This reading is supported by Kant's footnote at GMM 413:

the human will can \textit{take an interest} in something without therefore \textit{acting from interest}. The first expression signifies \textit{practical} interest in the action; the second \textit{pathological} interest in the object of the action. The first

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.
indicates only dependence of the will on principles of reason by itself; the second its dependence on principles of reason at the service of inclination - that is to say, where reason merely supplies a practical rule for meeting the need of inclination.

According to a traditional reading, the sympathetic man acts heteronomously because he acts from his empirical interest in the well-being of other people. This means that, by definition, acting from sympathy could not be morally worthy. However, it does not mean that sympathy is necessarily partial, and thus that feeling sympathy is incompatible with acting impartially. This distinction depends upon seeing that since Kant calls all acts with empirical motives "interested" acts, he implies a distinction between narrowly self-interested, or partial acts, and "interested acts." Narrowly self-interested acts, like stealing, taking one's life, failing to be beneficent, will fail the universalizability test; they are ruled out by the impartial tribunal of morality. But many "interested" acts could pass a universalizability test, even if they are not autonomous according to a traditional reading. For example, acting because one is moved by the ideal of perfection is heteronomous because one is moved by an empirical concept; but this does not entail that acting from the ideal of perfection is narrowly self-interested and could not be compatible with impartial morality. The principle "strive for perfection" is not like the principle "seek one's own narrow self-interest"; it could pass the universalizability test. That is, I could imagine myself as a legislator in a
Kingdom of Ends for which the perfection principle holds; in fact, Kant defends such a principle as an imperfect duty to oneself. The point is, Kant does not argue either that all empirical motives are partial, or that sympathetic emotions are always partial; he only argues that all empirical motives are heteronomous sources of motivation because they are not "derived" from reason alone. There is nothing in this argument that rules out the possibility that sympathetic emotions could (even if Kant didn't think they did) pass the test of universalizability. Thus it is the traditional conception of autonomy and not some picture of emotions as essentially partial, that supports the traditional equation of the moral standpoint with detachment.

III Current Arguments Against the Claim that Moral Agency Requires Impartiality

Bernard Williams, Lawrence Blum, Michael Stocker and others argue against the traditional Kantian view that morality excludes acts that are motivated by one's direct interest in the well-being of another person. Their common tactic has been to show that the detached Kantian agent seriously lacks the kind of commitment to others that moral theory is meant to describe. They use examples that involve personal relationships to show how unsatisfactory it is to equate morality with caring about rules or doing one's duty rather than caring about persons. One often discussed example is the "drowning wife case" (herein the drowning spouse case)
in which the moral agent is pictured as a bizarre rule-fetishist who deliberates coolly about whether she ought to save her spouse because it instantiates a universalizable principle to do so. As Bernard Williams puts it, this person seems to have one thought too many. We expect a fully developed moral person to be moved by concern for his or her spouse rather than by the idea of following a rule.

Yet it is misleading to say that what is wrong with the rule-fetishist is that she has "one thought too many," as if the very act of thinking or being reflective rather than acting automatically is what is amiss. Although it does seem odd that the agent needs to reflect about saving the spouse, the full context Kantians have in mind in such examples makes it less strange: usually the choice is between saving the spouse and doing some other compelling thing like saving five unknown schoolchildren. It is apparent that the alternatives to rescuing one's spouse can be made so compelling that some kind of deliberation is in order if one is to act morally, where the standpoint of morality requires some commitment to the ends of other persons in general. And not just in thought experiments, but in everyday living, including medical practice, acts done from compassion and love often require reflection. Parents must deliberate about how to help an especially needy child without deviating from their commitment to care for all of their children impartially. Physicians deliberate on a daily basis about how to give their time and energy to their patients in an equitable manner. The problem with the rule-fetishist is not the fact that she reflects, but the nature of her reflection.
What is so unsatisfactory about the rule-fetishist is the standpoint from which she deliberates; given the traditional reading, the moral agent needs to leave all affective ties to her spouse aside in order to engage in impersonal calculations. Her intentional object in considering what to do cannot be the drowning spouse, viewed affectionately, "poor Richard, whom I love", because this would make her act an interested and hence heteronomous act. Rather, she must have in mind the rules themselves, and a rationalistic reconstruction of her choice: should I save one spouse who I have promised to care for, versus five children who have hardly lived yet, etc.

The traditional reading takes any case in which the agent is moved directly by the suffering of another person to be a case of heteronomous action, of no moral worth; but there are at least two senses of being moved directly by another that are being conflated here. One sense in which one is moved immediately by another, which Kant clearly takes to have no moral worth, is for one to be caused to act because of one's psychological response to another, one's feeling state, straightaway, without any reflection. But another sense of being moved directly by another's suffering seems to be compatible with at least ordinary moral reflection: one deliberates about what to do in light of one's direct concern for another's well-being. For example, in the case where one's action may involve tragic consequences for others, one may not just save one's spouse instinctively or automatically, but may instead consider one's obligation to the drowning group of...
children, etc. However, the intentional object of such consideration is never the rules themselves, but the threatened situation of one's spouse, and the threatened situation of the children, etc. One thinks: Richard, my love, needs help -- but so do these children, poor souls -- what do I do immediately? Of course, the fact that the parties involved are drowning sets an immediate practical limit on how long any "normal" agent would then reflect. In contrast, when one deliberates about whether to tell a friend a secret that would relieve her anxiety greatly, but perhaps harm another friend, one might reflect for quite awhile. The point is that it is possible to express a commitment to duty and to be moved directly by another's suffering in one reflective moment. One need not set aside one's feelings to consider what one ought to do. The reason for invoking this possibility here is to point out that what is wrong with the rule-fetishist is not that she has one thought too many, nor the fact that her act is rule-governed. Rather, what is wrong is that her intentional object, in considering a suffering spouse, is stripped of all affective qualities that express the importance of one particular person to another. It is not what the rule-fetishist has (a commitment to rules), but what she lacks that is important.
IV. The Core Kantian Ideal of Impartiality as Residing in the Autonomy of the Will

The equation of impartial morality with detachment in Kant is based on a picture of "autonomy" as involving freedom from all empirical interests. This conception presupposes as its complementary concept the idea of "pure practical reason": the idea that reason alone is "sufficient" for morality. According to a traditional reading of the GMM, reason not only justifies the principles of morality, but can actually move an agent to act as morality commands. Kant's examples in the GMM accentuate the practical independence of the moral motive from all affective motives for action. In the suicide example at GMM 429, Kant shows that the idea of duty can lead one to preserve one's life even when one lacks any affectively based will to live. And in the case of the unsympathetic man (GMM 398) who no longer has the goal of helping others out of compassion, the idea of duty alone is sufficient to supply him with the goal of helping others. The idea that reason alone is "sufficient" in these cases depends upon the idea that pure reason can motivate one to view one's own existence impartially, and to take up the concrete ends of others.

There are several steps in Kant's argument for the "sufficiency" of pure reason for morality, spanning discussions in all three sections of the GMM. At GMM 400 Kant moves from the idea of a will that is unconditionally good, to the idea of a
will determined by "reverence for the law", an attitude that depends upon reason alone. This discussion is based on three propositions. The first is that the only unconditional good is a good will. A good will is an unconditional good in two senses. It is good in all circumstances, whereas good actions and things may be bad in some circumstances. Second, it is good in itself, independently of its relation to other things. Whether or not anyone ever valued a good will, it would still be good. This claim about the radical independence of the good will, which stretches the idea of something being "good", perhaps beyond recognition, is reinvoked in Kant's discussion of the formula of treating humanity as an end-in-itself at GMM 438. Kant calls rational agency an end-in-itself because it is a "self-existent" end. The German term here, "selbstandig" means "standing on it's own feet," independent of all else, of a context, like a "self-standing" lamp that relies on no other furnishings for a ground. Kant's "argument" here comes down to the apparently analytic claim that whatever is the source of moral worth must be radically "selbstandig", and that only a good will is independent in this sense. Kant collapses here the "goodness" of the good will with the independence of that goodness.

Kant's next two propositions at GMM 400 specify the structure of a will that is independent and hence morally worthy. Kant's second proposition is that an action done from duty has its moral worth, not in its end, but in its maxim. He defines the concept of duty as applying to a good will "exposed ... to certain
subjective limitations and obstacles" (GMM 397). Kant argues that since all moral
worth comes from the worth of a good will (he equates moral worth with absolute
moral worth), and since all inclinations and objects of the will lack the
independence of the "selbständig" good will, only the principle of the will is a
candidate for moral worth. The second proposition focuses on the independence
of the will with regard to it's structure: willing can be characterized in terms of its
principles alone and thus can be abstracted from and hence made independent of,
an agent's particular aims (intentional objects).

The third proposition outlines the only remaining possible commitment that
could characterize the good will, given its independence from empirical principles,
which are held because of empirical attachments. Kant claims that "duty is the
necessity to act out of reverence for the law," (GMM 400). Kant's argument for
this claim is that "an action done from duty has to set aside altogether the influence
of inclination, and along with inclination every object of the will; so there is
nothing left able to determine the will except objectively the law and subjectively
pure reverence for this practical law, and therefore the maxim of obeying this law
even to the detriment of all my inclinations," (GMM 400). (It is puzzling why
Kant does not see subjective reverence for this law as itself a threat to the purity
of the will).

Kant's conclusion at GMM 400 is that acting from duty means acting from
reverence for the law. This claim does not yet entail the claim that rational agents
could act against all inclinations, from reverence for the law alone. This claim about the capacity of rational agents requires a conception of practical reasoning.

Kant specifies his conception of practical reasoning as the source from which the concept of duty "springs" at GMM 412. Kant defines the will as the power to act in accordance with the idea of laws. Rational beings work in the phenomenal world, just like things; they are subsumable under causal descriptions, as things are. Yet, in addition, rational beings work not only in a way that can be described by law; they work in a way that is derived from law. But what can this mean? Here Kant is imputing causal efficacy to the inferential capacity of reason; reason can move the agent from having certain ends to implementing the means necessary to achieve those ends, just as it can move thought in an argument from premises to conclusions. By guiding the agent to follow what the laws of nature dictate to be necessary for accomplishing her goals, reason can generate a commitment to new goals. In the case of impure practical reasoning, this end-generating capacity of reason begins with given affective motives. One desires X, and then reasons that X requires Y, and this reasoning itself can motivate one to will Y. Kant is indebted to Aristotle for this conception of practical reasoning, despite his rejection of Aristotelian ethics.

But Kant then invokes a picture of pure practical reason that is not implied by this [Aristotelian] conception of practical reasoning: Kant claims that in the case of a good will, reason "solely by itself" is "sufficient" to determine the will (GMM
This claim goes beyond the claim that the "good will" is of "selbstandig" worth (GMM 400). Rather, Kant is making a claim about the efficacy of reason in the case of the good will. In the case of God and the angels, who have absolutely good wills, their practical reasoning necessarily generates morally worthy acts. The term "sufficiency" here is a translation into the practical sphere of the idea of "entailment" in logic. Kant seems to picture the case of perfect willing as entailing good action in the way correct premises and correct reasoning entail the correct conclusion of an argument. This presupposes not only that the pure rational agent's inference capacity would be perfect, but also that she would always start with the correct initial premises of a practical syllogism because she would always have the correct ends or goals. The case of pure practical reasoning is distinguished from the case of impure practical reasoning by the fact that in the former case reason sets the agent's ends, independently of all empirical motives.

Kant extends his picture of the "sufficiency" of reason in the case of the absolutely good will to the case of human beings, with imperfect wills. Kant uses the term "sufficiency" in two senses in GMM 412, implying that in the case of human beings reason is and is not "sufficient" for morality. He writes: "But if reason solely by itself is not sufficient to determine the will; if the will is exposed also to subjective conditions (certain impulsions) which do not always harmonize with the objective ones" only then does necessitation, and duty arise (GMM 412). The point here is that humans beings are subject to necessitation precisely because
the human will is capable of being moved by inclination, and hence not wholly in accord with reason. Unlike perfectly rational beings, human beings, and other finite rational agents, do not necessarily aim for the correct end or goal. This point reveals Kant's presupposition that the correct end of moral action is an end that is not only endorsed by pure reason, but generated by pure reason: it is reason's own goal of universality. Kant soon specifies this end as the goal of willing only those maxims that are universalizable (the categorical imperative). Human reason is capable of generating this goal, even if human beings can also generate other goals. Thus, the sense in which human reason is "sufficient" is that human reason alone, independent of all affects or empirical objects, is capable of generating the good end which the Godly reason necessarily aims for. Human Reason can follow the same practical syllogism God follows, even though it may not.

At GMM 448, Kant argues that rational agents must think of themselves as exercising pure practical reason:

And I maintain that to every rational being possessed of a will we must also lend the Idea of freedom as the only one under which he can act. For in such a being we conceive a reason which is practical -- that is, which exercises causality in regard to its objects. But we cannot possibly conceive of a reason as being consciously directed from outside in regard to its judgments; for in that case the subject would attribute the determination of his power of judgment, not to his reason, but to an impulsion. Reason must
look upon itself as the author of its own principles independently of alien influences (GMM 448).

This passage reveals that Kant's idea that reason, independently of all input, can generate its own ends in the practical sphere, depends upon his presupposition that practical reason and theoretical reason are one and the same reason. In the *Critique of Pure Reason* Kant argues that theoretical reason is capable of going beyond "sensible" input.\(^9\) For Kant, the "sensible" includes affectivity, feelings of pleasure and pain, as well as sensation.\(^9\) Kant claims that reason can think what it cannot know, that it can have ideas without the corresponding intuitions necessary for knowledge. Reason is not only capable of going beyond the sensible, but driven to do so by its own commitment to unity and its unlimited inference capacity. Theoretical reason is thus spontaneous in two senses: it organizes sensible intuition into concepts "spontaneously" (via the understanding); and it pushes beyond given input, toward its own goal of unification. Kant's picture of pure reason as capable of setting its own ends is an extension into the practical sphere of his picture of theoretical reason's "spontaneity" in the second sense.

But the Kantian idea that reason can move the agent in virtue of its radical spontaneity presupposes that practical reason can work independently of empirical causation. This metaphysical presupposition follows from the divorce of all empirical ends from the ends of reason. Kant seems to have conceived of all
psychological motives as operating deterministically in the way that physical causes operate deterministically. Thus, the independence of reason in the practical sphere requires that reason be capable not only of justifying moral action, but of actually causing the agent to do the right act. But Kant takes the idea that reason alone actually moves the agent to be incomprehensible if one considers the agent as a phenomenal being, since there can be no gap in the causal chains of nature. Rather, he argues that the idea of an intelligible cause only makes sense when the agent is considered as a noumenal being (GMM 452).

Thus, according to a traditional reading, the idea that reason alone is "sufficient" for morality is ultimately based in Kantian metaphysical assumptions. The "autonomy" of the moral agent has both a practical and a metaphysical sense: practically, the agent sets aside all empirical motives to act from the motive of duty alone. This is possible because, metaphysically, the agent is not merely an empirical being subject to natural causation, but also a noumenal being, who can initiate causal chains de novo (Critique of Pure Reason, B566-586). But this metaphysical assumption is notoriously problematic. Given the Kantian picture of causality as a category of the understanding, and of noumenal reality as, by definition, what cannot be grasped by the understanding, it would seem to be illegitimate to speak of "noumenal causality." Kant says that while such causality cannot be understood by human beings in the way the laws of nature can be understood, it can be thought (GMM 458). His discussion of practical reasoning
is meant to allow us to "think" this idea of moral agency. But we cannot have an idea of an agent abstracted of all the phenomenal aspects of the self. As Brian O'Shaughnessey, Donald Davidson and others have argued, the very concept of agency presupposes given desires, beliefs and intentions that are related in the right way to a person's movements in the world. Acts are intentional; they do not just happen to the agent, but follow from the agent's desires and expectations about the world. If one were not aiming for an empirical object, if one had no desire to do anything, than one's movements could not be grasped as actions rather than mere occurrences.

The noumenal self, stripped of all desires and intentions, cannot even be thought of as an agent. Thus the Kantian idea of pure practical reason leads not only to the "impersonality" of the rule-fetishist, but to an ontological condition of "impersonality." The fully detached "agent" cannot be understood as generating her own projects and pursuits. If the "person" is understood as the locus of a unified set of projects, then the categorical exclusion of such projects entails the exclusion of the possibility of being a person.

IV. A Revisionist Reading of Kant that Does Not Equate Impartiality with Detachment

In response to these problematic metaphysical assumptions and the picture of
the Kantian moral agent in the moment of acting from duty as a rule fetishist who lacks all affectivity, a revisionist reading of Kant has sprung up. Three philosophers who represent aspects of this view are Marcia Baron, Barbara Herman and Onora O'Neil. All three argue that morally worthy acts can be overdetermined: one can act with more than one motive, so that one can act from duty and yet be moved by sympathy. The revisionist reading depends upon making a distinction between the moral principles that guide action and what the agent is aiming for in acting. According to a traditional reading, the maxim or "subjective principle of action" includes a complete description of the agent's intentions. In contrast, the revisionists take the agent's maxims to describe only those aspects of the agent's motives that are relevant to the justification of her act. O'Neil pictures maxims to be general principles, like "help a friend when one can without undue struggle," which one decides upon or commits oneself to out of a commitment to morality. One uses the categorical imperative as a standard by which to judge the suitability of one's maxims. O'Neil implies that the agent need not actually go through this procedure to derive every maxim, but that it must be the case that her commitment to certain maxims in some way tracks her commitment to the standard of universalizability. O'Neil takes these maxims to be guiding principles, an almanac or road map, which the agent follows in particular cases, by aiming for things that are compatible with these principles. The agent would have no reason for using the road map if she did not have ordinary motives, like sympathy, that lead her to
act in the first place. I take O'Neil's point in using the metaphor of a road map to include the idea that the maxims genuinely direct, rather than simply endorse, what the agent does. Perhaps the kinds of maps O'Neil is familiar with are more like travel guides, because she implies that moral principles lead the agent to pursue ends which she would not necessarily pursue otherwise.

This interpretation of O'Neil's is slightly different from, but compatible with Barbara Herman's picture of the relationship between maxims and motives in moral action. In "Rules, Motives and Helping Actions"102 Herman argues that while "motive" and "end" are sometimes merely reciprocal concepts, this need not be the case. According to Herman:

the end, or object of an action, is that state of affairs the agent intends his actions to bring about. The motive of an action, what moves the agent to act for a certain object, is the interest he has in the object.103

By the "interest" the agent has in the action, Herman means a principle or commitment that directs the person to do what she does. One can have a variety of concrete ends, like buying an efficient car, going to a simple, self-service restaurant, etc. that express one's interest in being economical. Herman gives the example of altering eating habits out of concern with one's future health (one's end), which is a case of acting from principles of prudence (one's motive). Prudence might also direct one to visit the dentist regularly, etc. The motives of economy and prudence are "second-order" commitments which serve as filters and
sources of first order motives like buying an efficient car and eating healthily.

Herman applies the distinction between one's motive for acting, and one's end in acting to the case of beneficence from the motive of duty. In the case of helping another person, one's end may be to provide help, rather than to instantiate a rule. But one may have such an end because one is directed by moral rules that tell one to provide help: "The rules direct that in certain circumstances, actions of a certain sort are to be done. A moral rule requires us to help, to provide help, not to follow rules." Herman can thus distinguish the case in which one acts from duty but not as a rule-fetishist, from the case of the rule-fetishist: in the former case "the object [end] of the action is to save this person; the motive is to provide morally called for help"; in the latter case "the object of the action is to do what the moral principle (the duty to help) requires; the motive is to act in conformity with duty."

Herman is able to distinguish a case of acting from duty that is not rule-fetishism, by rejecting what she calls an "externalist" view of moral rules. In such a picture, the acting person aims to instantiate rules, because she is attached to such rules as one might be attached to an external authority, out of a commitment to being obedient. She contrasts this picture with:

the fact (and the way) moral rules are learned. We acquire knowledge of how, morally speaking, things work, and we employ our knowledge in determining what we should do. Moral rules are internalized; when learned
in the right way, they are a constitutive part of the agent's conception of himself as a person. They are neither memorized (as Ryle pointed out, they are not the sort of thing that can be forgotten), nor are they present as mere habits of response. In knowing such rules, we know how to go on.106

The distinction between the maxim and/or second-order commitment under which one acts and the intentional object of one's action makes it possible to argue that affective interest in the object of one's action does not entail that one is acting from interest, or heteronomously. This leads to a reinterpretation of Kant's statement at GMM, 426, that "the proper worth of an absolutely good will ... lies precisely in this -- that the principle of action is free from all influence by contingent grounds, the only kind that experience can supply." According to a traditional reading, this means that if human acts are to be morally worthy the agent must be free of all empirical interests. But according to a Hermanian reading, Kant is only saying that the moral worth of the maxim, not the agent's actions under the maxim, is independent of all empirical interest. It cannot be that the maxim has moral worth because it goes along with one's sympathetic inclinations, since sympathetic inclinations can be indifferent to morality. But neither is the moral worth of the maxim ruled out by the fact the agent who belongs to it has sympathetic inclinations.

It would seem that Herman and O'Neil have given sufficient ammunition to argue against the equation of morality with detachment. The distinction between
acting because one is trying to obey rules, and acting in a way that expresses one's commitment to duty, will permit acting from emotion as long as such action expresses a commitment to duty. The caricature of the rule-fetishist pictures the agent as segmented between a purely rational will that seeks only the end of being rational, and an ordinary will that seeks ordinary ends; the agent acts morally only when reason invades and occupies the whole person, putting its own ends of consistency and universality in the place of ordinary affective goals. The appeal of the revisionist conception is that acting from duty is pictured as ordinary affective activity, which is somehow rule "governed." The commitment to rational reflection is taken to be an integrated aspect of the agent's personality. However, the problem is in understanding the sense in which acts may be both ordinary affectively motivated acts, and yet rule governed in a Kantian sense.

The core of the Kantian conception that Herman and Baron attempt to preserve is the idea that the rules of morality do not just describe what the moral agent does; they direct what she does. This requirement has two aspects. First, the moral agent must know that she is acting from maxims that are universalizable. Second, she must be motivated by the idea that what she does expresses her commitment to duty. If she happens to have been brought up to act according to rules which all pass the test of universalizability, but has no commitment to doing her duty, than her act has no moral worth. And her commitment to acting dutifully cannot be merely an after the fact endorsement of her actions; rather this
commitment to acting dutifully must be what actually causes her to do what she does. This idea distinguishes the Kantian conception from an Aristotelian picture of moral agency. Even if one has been brought up well, and one has virtuous attitudes like bravery and kindness that ordinarily lead to good acts, and one endorses such attitudes reflectively, one is not yet acting from duty.

Herman and Baron both take Kant to assert that only acts done from the motive of duty have moral worth, where acting from refers to the causal efficacy of one's sense of duty. But then we must ask how the moral agent's commitment to the rules that ground rational agency can actually direct her action, if the model of the obedient rule-fetishist is to be rejected? Herman and Baron characterize the person acting from duty as having a counterfactual commitment, such that if what she does out of sympathy were not in accordance with duty, she would not do it, and if duty commands what she would not otherwise do, she would do it.\(^{107}\) I will refer to this counterfactual description of how duty could preempt other motives as "pre-emptive overdetermination."

The revisionist concept of "pre-emptive overdetermination" provides for a deflationary reading of the Kantian notion of autonomy, which avoids the problematic concepts of transcendental freedom and noumenal causation. The appeal of the notion of "pre-emptive overdetermination" is that the agent's freedom comes down to an empirically specifiable freedom to do otherwise; the agent acts autonomously if she does what she does because she believes it is morally
required. However, the picture of agency embodied in the concept of "pre-emptive overdetermination" is far from transparent. The idea that the agent's sense of duty can pre-empt her other motives can be interpreted in two ways, one of which retains, in deflationary form, the Kantian idea of pure practical reason (Herman's reading), the other of which does not.

In "On the Value of Acting from the Motive of Duty" Herman pictures the motive of duty as a commitment of the agents which can move her against all of her ordinary affective motives. Herman says that while one's sense of duty can be a second-order commitment that expresses itself through primary motives like sympathy, one's sense of duty is also capable of being a primary first-order motive, which is "sufficient" to lead to moral action all by itself. She reads the case of the unsympathetic man as a case in which the motive of duty alone moves the agent toward beneficence. And Herman takes the agent's sense of duty in this case to be based solely on his impersonal commitment to act in a way that is universalizable. She notes that it is only because the unsympathetic man lacks any personal reason to do what he does that his action is a case of beneficence in the sense that Kant uses that term in the DOV: "it [an action] is beneficent only if the agent conceives of what he is doing as an instance of what any moral agent is required to do when he can help another, and acts to help for that reason. For Kant, only the motive of duty [alone] could prompt someone to act on a maxim with such content --for no other motive responds to a conception of action that
regards the agent himself impersonally or is impartial in its application.\textsuperscript{110} It seems that Herman envisions the apathetic man as generating his beneficent motive solely from detached reflection about what universal morality commands. But this reinvokes the traditional idea that reason alone is "sufficient" for moral action. Herman concludes this article with the statement that "at the heart of Kant's account of moral worth" is the idea that the moral motive "expresses a kind of independence from circumstances and need, such that in acting from the motive of duty, we are, as Kant saw it, free\textsuperscript{111}.

Herman never explains how rational reflection alone could move the apathetic agent to help others. In the absence of any new picture of what rational reflection involves, Herman seems to rely on the traditional Kantian picture of the "sufficiency" of practical reason as grounded in the spontaneity of theoretical reason, but without the traditional assumption of noumenal causality. But Bernard Williams points out the problem in presupposing that ordinary (ie. phenomenal) rational reflection is "sufficient" for generating a commitment to valuing the ends of others. According to Williams\textsuperscript{112}, reflection about truth brings in an impartial standpoint because it is concerned with how the world really is, and how the world really is presupposes convergence of all knowledge about the world. But reflection about how to act does not presuppose any such convergence. Even if reflection takes me from seeing that my own ends should be promoted to seeing that you too will be committed to the idea that your ends should be promoted, it cannot move
me to promote your ends. The additional step of reciprocity is not a bare logical inference.

The force of Williams argument against Kant depends upon seeing that the "drive" toward unity and consistency of theoretical reason is insufficient for the moral commitment to universalizability of practical reason. One can recognize that one's ends are rooted in some general feature of one's make-up and that other tokens of one's type have the same ends; but this cognitive grasp of analogous ends cannot move one to value the other's ends as-if they were one's own.

If one holds on to the core Kantian idea that impartiality requires the agent to be "free" of all affective motives in the sense that she can act in the absence of all affective motives, than one winds up with the following reading of the case of the apathetic man: His aim in helping cannot be to relieve so and so's suffering, because this would mean he was moved directly by the suffering of another person, and thus in Kantian terms moved by pathological emotion rather than practical reverence for the moral law. Rather, his aim must be to help a person because it is what duty commands. But his commitment to duty in this case cannot be derived from his own reasoning, because, as Williams has shown, detached reflection cannot commit one to take-up the ends of other persons. But if his commitment to help others does not arise from within his own reasoning processes or from within his own ordinary motives, then it rests on no reason of his own. It is simply an arbitrary choice. But this leads to a picture of the apathetic man
as arbitrarily following an external authority; he has become the rule-fetishist.

The revisionists must leave behind the Kantian idea of pure reason's "sufficiency" in order to develop an account of how the agent's commitment to reflection is internally related to, and capable of directing, her other commitments and motives. In their more recent articles, both Herman and Baron do seem to be moving away from the core Kantian presupposition that one kind of reflection is "sufficient" for both theoretical and practical reasoning. In "The Practice of Moral Judgment" Herman discusses the importance of "perceptions of moral salience," which may include affective responses, for the practice of morality. She argues that the ideal of treating persons as ends in themselves commits moral agents to discerning who are persons, and what is owed to such persons, and that such discerning may involve affectivity.

Herman strives to balance her conception of morality as a situated practice with affective, historical components, with the demands of Kantian theory for an ahistorical, unsituated reflective consciousness as the source of moral commitment. Her solution is to conceive of the "fact of reason," which is the agent's consciousness of her moral independence, or autonomy, as the source of the agent's commitment to take-up the ends of other persons. But the kind of reflection that gives rise to a commitment to others involves role-reversal, in which one sees oneself as-if in the position of others. This implies that Herman takes the fact of reason to already involve role-reversal. But how could consciousness of one's
radical independence from one's own affective situation motivate one to put oneself in another's position? In the next chapter, I will address this question in light of the related phenomenological concept of empathy as projection from one's own situation imaginatively into someone else's, thus "freeing" oneself from one's own situation. I will argue that radical detachment from one's own situation is incompatible with empathy, and that awareness of one's own affective ties and needs is constitutive of understanding other persons. At this point it is sufficient to note that if Herman reads the "fact of reason" as involving empathic projection, she has already left behind the Kantian picture of the unity of practical and theoretical reason.

Marcia Baron conceives of the power of duty to pre-empt other motives as rooted in the agent's capacity to redirect her ordinary, affective motives from within. In "The Alleged Moral Repugnance of Acting from Duty"114 she argues, as Herman does, that one's sense of duty is "sufficient" ground for moral action; but Baron does not see this sense of duty as ultimately grounded in pure practical reason. It is not the agent's ahistorical, unsituated rationality or "freedom" that Baron appeals to as an explanation of how a commitment to duty alone can generate beneficence. Rather, Baron argues that the commitment to duty is necessarily historical and affectively situated. Rather than picturing the commitment to duty as immediately efficacious in a discrete moment of action, Baron sees the efficacy of duty as mediated by the deliberate re-education of one's
character over time. She writes that acting from duty "must be thought of not in terms of isolated actions, but as conduct viewed over a stretch of time, and governed by a commitment which unifies and directs the self."\textsuperscript{115}

Baron's emphasis on "conduct" rather than discrete acts, represents an important difference between her conception of duty and Herman's conception. Herman explicitly states that only a small subset of an agent's acts will be done \textit{from} duty, because only acts that are morally required can be done \textit{from} duty; helping acts that are not strictly required are not credited to the agent's sense of duty, except in the tenuous sense that if such acts were impermissible, the agent would not do them. In contrast, Baron sees one's sense of duty as "a concern to do the morally recommended as well as the morally required."\textsuperscript{116} Her view of a sense of duty as informing one's conduct over time identifies the agent's sense of duty with the unity of the agent's projects, and the development of the agent's character, rather than just with the performing of right action. The identification of the sense of duty with the agent's capacity to develop her own character implies that the agent's feelings and ideas, and not just her behavior, will be influenced by her sense of duty.

This leads to the question of what kind of reflection constitutes a developmental commitment to acting from duty, and how does such reflection generate a commitment to role-reversal? Baron says she cannot yet give a full account of such reflection, but she gives the following suggestion:
One who acts from duty will reflect on her conduct and not be left cold by thoughts about how she acted -- nor will she feel only the retrospective emotions (e.g. regret, unlike remorse) which enable one to evade moral responsibility for the conduct in question. \(^\text{117}\)

Baron thus envisions a non-detached kind of practical reasoning, which relies on the agent's capacity to feel remorse, among other things.

Extrapolating from Baron's comments, I picture the unsympathetic man not as a rule-fetishist, but as a depressed person of good character who feels remorse over his failure to care much about others. His sense of duty includes a commitment to taking-up the ends of others, even when he lacks spontaneous feelings of sympathy. In order to make the ends of others "as far as possible, [his] own" he reflects on the situation of others to find elements of their situations which are able to move him, perhaps because he feels that these problems are like his own in some way. There is a sense in which this reflection involves consciousness of "freedom." This man will have to struggle to free himself from the pathology of his depression, the pervasiveness of his feelings of hopelessness, to adequately grasp what is salient about another person's situation. The personal insights that enable him to recover from his depression sufficiently to care for others can be pictured as "freeing" him from the pathology of his own situation; but these insights themselves will not follow from detached reasoning, but only from emotionally engaged reflection. This man's sense of duty expresses itself in his
painstaking commitment to recover from his apathy, in order to adequately care for other persons. In the final chapter of this thesis I give an account of what such emotional self-regulation without detachment involves.

The model of "detached concern" in medical practice presupposes that a detached commitment to rational reflection is sufficient for the physician's moral responses to patients. This presupposition is rooted in a Kantian conception of morality that is most often expressed in medicine as the duty-based commitment to respect patients as "ends-in-themselves." The professional commitment of the physician to respect all patients, regardless of her own personal preferences, requires striving for impartiality. Yet this chapter has shown the inadequacy of the Kantian conception of pure rational reflection as the ground of the impartial commitment to respect the ends of other persons. The traditional presupposition that impartiality requires detachment from one's affective, historical situation reduces the moral agent to a rule-fetishist who lacks the moral sensitivity that medicine requires.

The revisionist reading of Kant shows that one can act from duty and also be affectively engaged. Yet the revisionist reading has retained in part the traditional idea that one's commitment to respect others as ends is rooted in one's ahistorical, unsituated rationality alone. But, as Bernard Williams points out, there is nothing about reflection so conceived that could generate a commitment to take up the ends of others. Rather, the kind of reflection that grounds one's sense of duty to others
must involve "role-reversal." Having used Descartes and Kant to show the inadequacy of detachment for understanding and valuing the subjective experiences of patients, we can now explore the affective basis of "role-reversal" by rethinking the concept of clinical empathy.
Chapter Four: The Concept of Clinical Empathy

There are numerous articles by doctors claiming that empathy is essential for diagnosis and for developing a therapeutic alliance with patients that facilitates treatment. Yet despite this agreement over the need for empathy, physicians have varied, even contradictory conceptions of what the term "empathy" means. Michael Basch describes the widespread disagreement by clinicians over whether empathy should be considered...

An end result, a tool, a skill, a kind of communication, a listening stance, a type of introspection, a capacity, a power, a form of perception or observation, a disposition, an activity, or a feeling.

The confusion over what "clinical empathy" involves is in part a result of the ambiguity and vagueness that adheres to the general concept of "empathy." The O.E.D. defines "empathy" exactly as Theodor Lipps defined the term "Einfühlung" in 1903: "The power of projecting one's personality into (and so fully comprehending) the object of contemplation."

Lipps' emphasis is not quite right for physicians in that physicians are not seeking to contemplate their own personalities in their patients, but rather to learn something about their patients' experiences. (The term contemplation reflects Lipps' interest in aesthetics). However Lipps' description of "projecting" oneself...
into an object to more fully comprehend it is central to the concept of empathy used in clinical discussions, even though there is little clarity about what such "projection" involves.

Despite the confusion over the definition of empathy, there are certain observations that any conception of clinical empathy must fit. First, although the original concept of empathy as "feeling into" another person's experience has led some theorists to equate empathy with direct perception or telepathy, physicians today reject this view. There is no organ for empathy, and no discrete form of energy transmission involved in communicating emotion, as there is in the case of visual, auditory, olfactory, and tactile perception. Rather, most clinicians take empathy to be like but not identical with sense perception in that empathic understanding involves pre-reflective receptivity to messages that another person communicates, voluntarily and involuntarily.

Second, while many physicians note that there are momentary "flashes" of empathy, they also note that the accuracy of empathy increases with time, effort, and increasing familiarity with the patient. Such flashes may be similar to the flashes of discovery one has after long periods of working through one's ideas. They do not in and of themselves contradict the claim that empathic understanding requires reflection and effort.

Third, physicians have observed that empathic understanding depends not only on their own capacities, which will be the focus of much of this chapter, but upon
interpersonal dynamics, including the degree to which physicians and patients can understand each other's language, style and values. Given that such understanding depends upon communication, empathy is both an intrapsychic and an interpersonal activity.\textsuperscript{124}

Fourth, clinical empathy involves building a conception of another's situation-as-lived. Hence there is a sense in which empathy is best understood as constructing a model of the patient's situation, as Buie describes it.\textsuperscript{125} However, the term "model" is misleading in that it implies that the physician's imaginative creation is an entity that can stand apart from the physician's experiences. This reification of the content of empathic imagining is perhaps invited by the fact that historically the concept of empathy originally referred to contemplation of a work of art. But I will argue that in the case of clinical empathy, there is no separable mental model of the patient's world that the physician holds in mind without participating emotionally in the patient's experience. Rather than mistakenly searching for a product of empathy, it is better to think of empathy as a capacity to follow the patient's story affectively and imaginatively.\textsuperscript{126}

However, even though empathy does not yield an "inner model" of the patient that can be extracted from the empathizer, one can generate descriptions about the information one has gathered through empathic imagining, and these descriptions can be used to build models that can be contemplated theoretically. For example, the empathic discovery that a patient with anorexia nervosa actually sees herself
as fat, can be conveyed (I have just done so), and used in a description of what anorexia is like. The reader can understand and use the idea that the anoretic patient sees herself as fat without empathizing. Alternatively, one could convey, in a concrete case history rich with the affective imagery of the real patient, how it feels to be anoretic, inviting the reader to empathize.

The observations that empathy involves communication, improves with increasing familiarity with the patient, and involves both pre-reflective receptivity and conceptualization, must all be accounted for by an adequate conception of clinical empathy. These observations suggest that empathy cannot be easily fit into a faculty psychology that divorces cognition from affect. Our account of empathy must explain how the interaction of cognition and affect enables the physician to learn something new about the patient's experience, which cannot be provided by the physician's other ways of understanding the patient.

In the first part of this essay I argue against the predominant medical model of clinical empathy as detached insight. In part two I return to Lipps' definition of "Einfühlung" as a kind of cognition that is essentially affective. In part three I turn to some psychoanalytic models of empathy that take affective engagement to be essential, but still do not explain how the physician can learn something new about the patient via affective engagement. In the final part of this essay I give my own non-inferential account of how the physician's emotions can be informative about the patient's situation.
I The Conception of Empathy as Detached "Insight"

The definition of empathy as "feeling into" suggests a conflict between the standpoint of empathy and the standpoint of detachment. Yet many physicians today presuppose that the detached standpoint is sufficient for empathy. For example, Renee Fox and Howard Lief state that it is only after medical students go through a period of alienation in which they overcome their personal responses to patients in the way they overcome their fear and disgust at dissecting a cadaver, that they can develop the skill to listen empathically to patients without becoming emotionally involved.127

Charles Aring's "Sympathy and Empathy," Journal of the American Medical Association, 1958, and Herrman Blumgart's "Caring for the Patient," New England Journal of Medicine, 1964 are two classic articles by physicians about "detached" empathy.128 Aring and Blumgart argue that empathy involves making correct inferences about the patient's condition rather than feeling anything in response to the patient. Both physicians argue that for the physician to respond affectively to the patient's situation would not only make the physician unhappy, but would also prevent him from making correct diagnoses, and hamper his judgments about therapy. Aring conceives of "detached" empathy as the use of one's own knowledge of emotional experience to make inferences about the patient's condition. Blumgart conceives of "neutral empathy" as the careful observation of
the habits, and attitudes of the patient so that one can predict how the patient will respond to her illness and to treatment.

The core thesis put forth by Aring and adopted by Blumgart is that empathy can be and should be fully distinguished from sympathy. Their point is not just that one can best empathize with the patient if one does not "sympathize" with her in the sense of pitying her. Rather, they claim that the physician can best empathize with the patient if he refrains from "feeling with" the patient in any way. Aring defines the term "sympathy" as "an affinity, association or relation between things so that whatever affects one similarly affects the other. The act or capacity of entering into or sharing the feelings of another..." And Blumgart adopts Aring's definition of "sympathy."

Aring argues that the emotionally moved physician will become hostile in response to a patient who is very dependent. And Blumgart argues that sympathy is destructive because the sympathetic physician will grieve for patients, and regret his limitations, whereas the "neutral empathetic" physician will simply do what needs to be done without such reactions. These examples suggest that Aring and Blumgart picture sympathetic feeling as stirring up the physician's (often unconscious) personal conflicts. But this shows that Aring and Blumgart take the notion of "participating in another's situation" quite literally to mean taking on another's problems as burdens of one's own.

Aring and Blumgart show that when the physician reacts to the patient's
situation by taking on the patient's conflicts, this disturbs her capacity to listen well to the patient. But such reactivity, in which one actually experiences a need to solve another's problems as if they were one's own, is only one form of emotional responsivity. It is the way an older sibling might respond to the problems of a younger sibling -- by wanting to fight by their side against the bully in the schoolyard. In contrast, a parent interested in guiding her child's development, might experience the child's fear empathically, but with a concurrent awareness that the bully is not really such a threat. By not taking on her child's need to fight the bully the parent might then enable the child to deal with the bully from an entirely different emotional standpoint, using humor or curiosity. But while I reject Aring and Blumgart's assumption that responding emotionally coincides with reactivity, I accept their observation that the former can invite the latter. It takes work for the mother not to respond childishly to the bully. In chapter five I show how the physician caught in emotional conflict can, like the parent, learn to regulate her emotional reactivity while maintaining emotional involvement with the patient.

I also agree with Aring and Blumgart that empathy can be utterly devoid of sympathetic advocacy. Consider how one can say something particularly hurtful to someone whose vulnerability one grasps empathically. But this difference between empathy and sympathy does not show that empathy can occur in the absence of "feeling with" another. For it is not unimaginable that in one's cruelty
one vicariously experiences the suffering of another, and yet enjoys this experience. (There are compelling psychoanalytic theories about how one could develop an eroticized response to the experience of suffering, either in others or oneself).\textsuperscript{134}

The idea that empathy is independent of emotional engagement involves additional assumptions beyond the observation that empathy does not require advocacy. Empathy is pictured as an ordinary form of inferential reasoning that simply has a special subject matter. Recall the words of Osler: the physician capacity to neutralize his emotions allows him to "see into" and hence "study" the patient's "inner life."\textsuperscript{135} This claim presupposes that the physician can "project" before his "mind's eye" the patient's "inner life" as if it were "an image, as from a transparent slide, upon a screen."\textsuperscript{136} A related assumption by Osler is that the physician can also "see into" his own "inner life," in order to recognize what he has in common with the patient. Osler thinks that imagining how another feels depends upon producing relevant images from one's knowledge of what typical emotions are like, and then applying this knowledge to the patient inferentially.

Osler, Aring and Blumgart presuppose that "knowing how" the patient feels depends on the working of the same cognitive faculty involved in knowing how the patient's body is functioning. When used to refer to third personal or impersonal knowledge about a state of affairs, such as the workings of bodies, the term "knowing how" is interchangeable with the term "knowing that." That is, knowing how the stomach puts out gastric acid is a matter of knowing that the
histamine cells stimulate the release of certain hormones. According to the paradigm of "detached concern," knowing how the patient feels also is a matter of knowing that the patient is depressed, or reexperiencing a childhood fear. Such knowledge depends on careful observation and making correct inferences.

But while careful observation of another's words and gestures contributes to empathic understanding, it is not another's observable movements, but what they signify that is the "object" of empathy. And one cannot directly inspect the patient's feelings toward the world in the way one can directly inspect an entity. In observing an entity of complex structure, one encounters hidden aspects that cannot all be presented at once; but these aspects can in some form be made present to sense perception, either through dissection, magnification, radioisotope labeling, etc. But, as we discussed in chapter two, one cannot make the person's inner feelings apparent, and measurable, in this sense.\(^{157}\)

To make sense of the concept of empathy, knowing how a depressed patient feels must differ in some way from knowing that, as a matter of fact, the patient meets the statistical criteria for depression given in The Diagnostic and Statistical Manual of Mental Disorders.\(^{138}\) Lipps introduced the term "Einfühlung" to describe a way of understanding experientially what it is like to be in another person's position, as opposed to observing another person's responses as "matters of fact."

Eric Cassel points out that in order to understand illness as it is lived, the physician must get beyond objectifying the patient's condition:
Injuries to the integrity of the person may be expressed by sadness, anger, loneliness, depression, grief, unhappiness, melancholy, rage, withdrawal, or yearning. We acknowledge the person's right to have and express such feelings. But we often forget that the affect is merely the outward expression of the injury, not the injury itself.\textsuperscript{139}

Cassel's point is that the damage a person experiences includes an entire way of being in the world which cannot be grasped only by considering the attitudes the person displays. For example, the depressed person is not only visibly sad (and not always visibly sad), but also experiences loss of confidence, feelings of guilt, difficulty thinking about things. And each individual will be guilty, worried, and confused about different concrete issues, and these concrete issues will be what largely characterizes her personal experience. The personal and concrete is not yet understood when one simply recognizes that the patient is depressed. For example, the physician needs to know how difficult it is for the patient to walk with this cane, whether suggesting another treatment of chemotherapy makes this patient feel hopeless, etc. In empathy, one is directed toward another person's particular intentional objects.

I take it that Aring and Blumgart, and Fox and Lief, would not be opposed to Cassell's portrayal of the goal of empathy as understanding the particular aspects of the patient's emotional experience. Nor do they entirely reject the idea that the physician must make use of her own past emotional experiences to come to such
an understanding. Their thesis is that it is possible to come to a cognitive grasp of the patient's position by making use of one's knowledge of past emotional experience, without actually re-experiencing occurrent emotional resonance with the patient.

But what kind of cognition allows us to use our past emotional experiences to understand another's particular emotional experiences? Can we make literal sense of the idea that the physician has a capacity to "introspect" and then apply her emotional knowledge to the patient inferentially, without actually experiencing shared emotion? The term "introspection," which literally means "to look into one's own mind," has a narrower and wider use. The narrower use refers to first-personal awareness of bodily feelings, sense perception, and states of consciousness. However, the information gained from introspection in the narrow sense could not provide an adequate basis for empathy, given that the focus of empathy is the patient's attitudes towards her situation. Rather, what is needed is a model of a kind of "introspection" and inference-making that would yield awareness of motives and judgments as well as 'interior' feelings.

One influential model of empathy as detached insight comes from the picture of the psychoanalyst as able to grasp the analysand's internal world while remaining emotionally neutral. But it is a mistake to posit that the insight that enables the psychoanalyst to make inferences about the analysand involves a discrete act of "detached" "introspection." First of all, psychoanalysis uses
heterogeneous methods to lead to increased self-knowledge, including free-association, reliving of past experiences, transference, and dialogue. And these methods rely on heterogeneous modes of response from the patient, including strong feeling, recollection, and fantasy. And although there is active controversy among psychoanalysts about the level of emotional engagement that is appropriate for the analyst, there is consensus that the analyst's own free associations will involve experiencing affects.141

Further, the term "introspection" suggests a privileged first-personal awareness of one's condition, in which the reflecting subject and the object of reflection are identical. But there is no such identity in psychoanalytic introspection. As Tugendhat has shown, we can only attribute motives to ourselves in the way we attribute such attitudes to others -- by considering our own actions, retrospectively, as they fit into relevant patterns.142 And, as Edith Stein argues, when we consider ourselves retrospectively, through recollection, the 'I' that is recollecting is non-identical with the 'I' that is remembered.143 She gives as an example the case in which I can recollect my past fear over something I then believed, without feeling afraid anymore given that I now no longer hold that belief. Thus when the term "introspection" is used to indicate awareness of one's own complex emotional attitudes, it cannot be understood as referring to a discrete mental act of an 'I' turning in upon itself reflectively. But since the term "introspection" is used as a metaphor for some unspecified reflective experience, it does not help us give a
specific explanation of how empathy is possible.

The model of empathy based on introspection and inference lacks an adequate conception of how the physician can detach herself emotionally and yet "see into" her own mind to produce the relevant images. But even if we allow that such insight is somehow obtainable, we face the deeper problem of explaining how the physician can use such insight to infer something new about the patient. Diagnosing disease usually involves making predictions about something not present and observable, for example, the causes of and future course of a pulmonary infection, based on something present and observable, for example, the chest X-ray and sputum sample of the patient. The idea that empathy involves making inferences is modeled on this picture of inference in science: one "predicts" the qualitative experience of the patient, based upon one's insight about emotional experiences in general and one's careful observation of the patient's words and gestures.

The idea that the empathic physician is like a weather forecaster noting the changing conditions of the patient and predicting the shifts in "weather" is so appealing because it captures part of the truth. Empathic understanding involves trial and error, in that the physician needs to test her grasp of the patient's feelings out and modify her views according to the patient's feedback. But all knowledge must be susceptible to confirmation or rejection by some criteria, in order to be meaningful in the first place. This includes practical knowledge, like knowing how
to folk dance. Folk dancers can learn by trial and error because only certain kinds of gestures count as proper folk-dancing. But knowing how to folk dance is not a matter of making inferences -- a dancer whose knowledge is based on predicting where her feet might go next would certainly be a sight to see! Hence the fact that empathy involves trial and error does not entail that the work the empathizer does involves making predictions.¹⁴⁵

The error involved in equating empathy with prediction can be seen by taking a slight detour: Consider the two very different attitudes one can take up toward an action. One can either take up the attitude of the agent, or the attitude of an observer. In the first case, one intends to do something, call it Y. When one commences to will Y, one is committed to the proposition: "I will do Y now." At this point, one has a special grasp of the fact that Y is about to occur. One already lives in the world in which Y is on the way, in the sense that events are organized in one's field of interests around the happening of Y. In contrast, an observer can only predict that Y may occur, based on inferences from prior events, theory, etc. Now, the difference between the observer and the agent is not that the agent is right more often. It may be that the agent thinks she can lift 1,000 pounds, and the observer finds this extremely unlikely, so that the agent is less likely to be right than the observer about the actual lifting. The point is that the observer can only make predictions about the agent. But the agent does not need to make predictions to anticipate her own acts. She can anticipate her acts directly through forming
intentions that lead to striving and, when physically possible, successful action.

The picture of empathy as detached inference likens the empathizer to an observer, whose only way to anticipate the other's acts is through prediction. This would involve remaining equidistant from various possibilities in a field of unrealized possibilities. If the empathizer were like the observer in this sense, then she could simultaneously entertain the possibilities that the patient's tears express excited joy or hopeless despair. But the empathizer cannot simultaneously grasp despair and joy from a quasi first-personal perspective. The object of empathy -- the features of emotional experience from a first-personal perspective -- sets constraints on the mode of empathic understanding. To understand another person's despair empathically is to grasp their despair, not as a possibility, but as-if it is actually present. (The "as-if" aspect of empathy will be analyzed later in this chapter). Empathic understanding is more like the first-person experiential knowledge of the agent anticipating her own acts than it is like the third-personal predictions of the observer. Therefore, it is mistaken to picture empathy as just another use the physician makes of her capacity to make scientific inferences about the patient.

So far my argument against equating empathy with observation, introspection, and inference has focused on the intrapsychic component of empathy. But the "cognitive insight" model of empathy is especially inadequate at explaining the interpersonal, or communicative aspect of clinical empathy. If the physician relies
strictly on detached observation and conceptual generalizations about affective experience she will disserve the patient in several ways. First, she will miss important features of the patient's individual experience that are not contained in her previous generalizations; and she will expect typical reactions that may not be this person's reactions. Second, like the awkward folk-dancer relying on predictions, she will be unable to follow the patient's story experientially, and not know when or how to say or do the right thing. The clinician would be like someone who goes to the theater but only notices those aspects of the drama that correspond to what is in the written program synopsis. She would have no clue about how to gasp, sigh, and cry appropriately with the audience. Third, the patient is likely to notice that she is alone in her dramatic moments, and that her situation is being seen as a typical instance of such and such. Most likely this will make the patient feel that the physician is not really sufficiently interested in or open to her experiences to merit the trust needed to openly reveal her history, and to build a therapeutic alliance.

It impoverishes the concept of clinical empathy to reduce it, as Osler, Aring and Blumgart do, to weather forecasting. Rather, the empathizer must be sufficiently involved in the patient's "weather" to be able to recognize and appreciate in some quasi first-personal way, how the rain and sun feel. And we have seen that Osler's hope that by detaching himself emotionally the physician could clearly introspect and infer what such experiences are like is untenable.
Sawyier, a philosopher of science, notes that: "When we fill in the concept of empathy, part of what we imply is that the empathizer has himself had something happen to him right then; it is not just that he has thought hard, or tried to figure something out." This leads us to reject the insight model of empathy because it denies the two experiential poles of empathic understanding: in empathy one grasps, more or less, how the other person experiences her situation; and at the same time the empathizer herself experiences the other's attitudes as presences, rather than as mere possibilities.

II The Original Meaning of the Concept of Empathy

To clarify the sense in which knowing how another feels necessarily involves the knower as an experiencing subject, and hence precludes detachment, let us return to the origin of the concept of empathy. First, it is notable that Lipps' concern is primarily with aesthetics, although at that time aesthetics was considered a branch of psychology. He invokes the concept of Einfühlung primarily to describe a mode of comprehending works of art, and secondarily to refer to a way of comprehending the psychological life of other persons. He says, for example, that the appreciation of music is the paradigm of Einfühlung. Lipps defines this form of aesthetic comprehension as follows:

Einfühlung is inherent in something perceived by me or in an element truly
belonging to me and me only, i.e. something subjective, understood by me in the subject or in the spiritually corresponding object, but in the object for me or the object as it "looks" to the perceiving subject.\cite{Lipps}

Lipps emphasis here is on the essentially experiential nature of Einfühlung. By "subjective" he means "experiential," where experience has the sense of "Erlebnis" or lived, first-hand experience. The contrast term for experiencing something in this sense is thinking about something. By "feeling into" Lipps has in mind a mode of perception that is essentially affective and thus different from sense perception as conceived by traditional faculty psychology. Einfühlung is characterized by the essential connection between the affective feelings of the empathizer and the perceptual object of empathy. Thus for Lipps the phrase "projecting one's personality" refers most generally to the experience of a connection between one's own affective condition and the object one is trying to understand.

Lipps gives a more specific description of Einfühlung that is directed towards understanding another person. He says of the gestures and expressions of other persons:

These sensuous manifestations are not the "man," they are not the strange personality with his psychological equipment, his ideas, his feelings, his will, etc. All the same, to us, the man is linked to these manifestations. The imaginative, feeling, willing, individual is immediately apparent to us.
through his sensuous appearances, i.e. his manifestations of life. In a movement, grief, spite, etc. is perhaps apparent to us. This connection is created through Einfühlung.150

Here Lipps further specifies a concrete sense in which Einfühlung requires a kind of "projection." The experiencing subject must reach beyond what is apparent to grasp the psychological life of another person. And this reaching must direct her attention from the start, if she is to actually perceive the movements of another as meaningful expressions of emotion. Thus her attention must be directed by certain interpretive expectations.

But this contradicts the model of clinical empathy as detached "insight." If an interpretive act is necessary even to experience the other's gestures and words as announcements of affective experience, then the goal of ridding oneself of one's own affects in order to receive affective input from the patient without any prejudice or distortion, is no longer comprehensible. Rather, empathy involves framing the input that one receives from another person in terms of one's own emotional experiences. To receive another's experience as meaningful requires that one have pre-reflective expectations about what it is like to be angry, sad, etc. Hence, empathy cannot be a kind of detached cognition free from all perspectives and prejudice.

Lipps' view immediately resonates with Heidegger's view, discussed in chapter two, that "moods" are not discrete reactions to independent input, but unified ways
of being in the world, which provide an organizing framework for structuring input in the first place.\textsuperscript{151} For Heidegger, what makes it possible to understand something is the prior possibility of being in relation to that thing, where being means existing -- with the full range of affective and volitional activities. We can apply this to the empathizer, who recognizes the significance of another person's experiences because they are related to her own concerns. The point is not just that the empathizer must be in some mood or other, for this is true of all being-in-the-world. For example, Heidegger takes the theoretical observer to be in a mood, but one that excludes receiving affective input. In contrast, the empathizer must be in a mood that is interested in the affectivity of another.

However, by extending Heidegger's idea that all knowing involves pre-reflective interest to the concept of empathy, I do not mean to assert that one can only understand emotionally that which serves one's self-interest in the narrow sense. To see in others only what was relevant to one's own particular needs or wishes would be extremely stifling and isolating; one could never learn anything new about others if one referenced all of their experiences in terms of one's narrow self-interest.\textsuperscript{152} Rather, the empathizer cares about another's particular feelings because they are relevant to her own existence as an affective being. It is only because we are interested, qua persons, in a broader range of experiences than those that serve our narrow self-interests that we can be emotionally responsive to others without being reactive advocates, as I argued above.
In summary, by rethinking Lipps' discussion of Einfühlung, we can replace the vague phrase "projection of one's personality" with the following description: Einfühlung is an essentially experiential understanding of another person, which involves an active, yet not necessarily voluntary, creation of an interpretive context. One's capacity to respond as another person is what provides content for this interpretive act; this is the fact emphasized by the term "personality" in the O.E.D. definition of empathy as "projection of one's personality." This act of experiencing-interpreting allows one to understand aspects of reality that could not be grasped by a detached entity who could receive sensory input, but who lacked personality.

III Three Models of Clinical Empathy as Affective Understanding

How, specifically, can the physician's affectivity be the basis of an interpretive understanding of the patient's situation? There are various responses to this question, but all of them share a general theme. The empathizer is somehow able to grasp the subjective dimension of the patient's situation by experiencing something representative of the patient's situation. Whereas "projection" is taken to refer to a pure cognitive capacity in the "insight" model of empathy, Kohut, Basch, and Buie take "projection of one's personality" to refer to an affective and imaginative capacity to feel "as-if" one were in the other person's situation. I use
"as-if" here to indicate that empathy does not require actual involvement in another's relationships and projects, but only imaginative involvement.\textsuperscript{153} For Kohut, Basch, and Buie this specifically means that the physician uses her occurrent emotions as imaginative representations of the patient's feelings. These clinicians offer three different views of how the physician's feelings serve as imaginative representations of the patient's feelings: Kohut presupposes that the physician's capacity to feel what the patient feels allows her to identify with the patient to the degree that she can momentarily experience the world as-if she and the patient were one person; Basch takes the physician's emotional responses to be direct indicators or signs of the patient's own emotions; Buie presupposes that the physician constructs a mental model of the patient's attitudes by re-experiencing her own similar emotions.

The first view, which I will refer to as the "merging" model of empathy, has early roots in the work of Freud\textsuperscript{154}, Deutsch (1926)\textsuperscript{155}, Fliess (1942)\textsuperscript{156}, Fenichel (1953)\textsuperscript{157}, and is fully explicated by Kohut (1959).\textsuperscript{158} I consider here a generalized picture representing the key assumptions they hold in common, without being able to do justice to the complexity of any one version. These psychoanalysts presuppose that empathy requires that the physician experience a "merging" of herself and the patient such that she seemingly feels the patient's emotions along with the patient. Helene Deutsch hypothesized in 1926 that empathy involves an unconscious identification between therapist and patient in which the therapist feels
the patient's experiences to be her own, and only on reflection recognizes that the source of the feeling is the patient. Fenichel talks about the therapist's "narcissistic identification" with the patient which involves the "taking over by the subject of the object's inner state." Fliess talks about the therapist regressing to a state in which he has weakened ego boundaries and can experience the patient's inner state from within. Similarly, Kohut equates empathic feeling with the bracketing from consciousness of a distinction between self and other. Kohut posits that when one's critical faculties are not operative, one can experience a boundariless continuum between another's feelings and one's own. By sharing emotion with another in this sense, one gains access to how it feels to be in her concrete situation.

But what exactly does such "merging" involve? First, the term "merging" is a physical metaphor for a psychological experience that requires definition. Many writers speak of the psychological process as one of "identifying" with another person. But the concept of "identification" is itself in need of clarification. The term "identification" is not used here in the strict sense to indicate a phenomenon in which one actually develops structural features of one's personality as a result of imitating an admired (or feared) other. The most ubiquitous example of strict identification is the child's identification with the parent. Rather, the claim is that in empathy one identifies "not with the other person per se, but with what he is experiencing." The "merging model" presupposes that the physician identifies
not with the patient's character, but with the experience of being in the patient's situation.

The idea that emotions could be shared in this sense depends upon a conception of emotions as intentional attitudes with corresponding typical objects. That is, there must be typical ways of seeing, and typical images, that go along with particular emotional attitudes. Only if this is the case can it make sense to think that in taking on an attitude, one can also take on a way of "perceiving" a situation. However, this conception of emotions, which we argued for in chapter one, does not lend particular support to the merging model; as we will see, the typicality of emotional objects underlies the other models of affectivity that we consider next.

In addition to this background conception of emotions, the merging model of empathy involves at least three specific presuppositions: first, that there is a kind of affective communication in which one responds to another's emotions with similar emotions; second, one's feelings will be similar not only in the sense of being the same type of feeling, but also in being directed toward the same type of situation; third, that it is possible to feel the same way another feels towards her situation because empathy involves bracketing from consciousness ordinary awareness of the distinction between self and other, so that one seems to be in another's identical situation.

First, the "merging model" relies on the idea that feelings can be "contagious."
Almost every account of clinical empathy since Freud describes a component in which the physician finds herself just feeling sad or happy with the patient "automatically" without relying on any effort of thought. Buie, following Freud (1892) uses the term "resonance feelings" to refer to those experiences in which one person's mood seems to be directly transmitted to another. Basch argues that such responses are based on an innate capacity to respond to another's expressed feeling with the corresponding affects. He posits that infants respond not only to parental smiles with smiles and tears with tears, but also with the precursors of joy and sadness. Basch bases his argument on the extensive research of the psychologist Silvan Tomkins, which shows that emotional responses originate as innate capacities, conforming to stereotypical patterns.

The idea that we have an innate capacity for resonant proto-emotions was foreshadowed in chapter one, where I argued that children learn how to respond to typical situations by having typical emotional responses. The missing piece of that argument is the idea that children must have a capacity to resonate emotionally with their elders in order to follow their elders in attaching to paradigm scenarios their appropriate experiential component.

But resonance occurs not only in childhood, but throughout life. The reader has most likely experienced a situation in which he or she has responded to another person's tears with "automatic" tears or to her laughter with laughter. Often, such responses are referred to as "natural sympathy." Rhetoricians depend upon this
phenomenon when they deliberately use exaggerated expressions of anger and enthusiasm to incite a crowd. The reason for calling such responses "automatic," "direct," or "natural" is that they do not seem to rely on recollection, imagination, or any other mental act other than noticing how another person feels. Buie describes sitting with a patient who was sobbing out of intense sadness and feeling "also purely sad; tears often rolled freely down his cheeks". According to Buie, this response differed from other experiences in which he recollected related experiences of his own, or tried to imagine how he would feel if he were in the patient's situation. In this case, he felt that this "was not a sadness of his [the clinician's] own." I take Buie's point to be that the feelings of sadness had no personal referent because they were not attached to a recollected or imagined experience of the clinician himself.

But in affirming the existence of resonance emotion, we need not affirm the second assumption of the merging model, that one's resonant emotions will be directed towards another's intentional objects. In the example above, Buie feels resonant sadness, but his sadness does not have as its initial focus the image of having been abandoned by one's parents as a young child (the patient's situation). One can have resonant feelings, without seeing the world the way another sees the world. The physician can experience "resonant" grief with a person whose spouse has died, without actually grieving for a spouse. Similarly, if a patient feels persecuted by the hospital staff, the physician does not need to feel persecuted by
the hospital staff to "resonate" with the patient's feeling of fear. Consider the classic case of "resonance" as the spread of feeling in a mob. Even if, for example, an angry mob shares a general scapegoat (blacks, Jews, etc.) the individual members of the mob will be angry at different, individual fictional or real blacks or Jews, rather than directed towards one object. One person may be thinking of the boss she hates, another of the lover who rejected her. If they have been manipulated by the same hate literature, their intentional objects may happen to share some of the same general features (such as shiftlessness or pushiness), but this is certainly not necessary. So resonance is insufficient for merging, in the sense of sharing the same attitude towards the same typical object. But then resonant feelings in and of themselves cannot be the basis of grasping how, in particular, another person feels about her situation.

Third, the "merging model" presupposes that resonance must occur in the context of setting aside one's critical awareness of the distinction between oneself and another. This idea has its roots in Lipps' account of empathy. Lipps associated the capacity to forget oneself with the ability to take on the concrete feelings of the "object" of contemplation. For example, Lipps says that one can be so absorbed in watching an acrobat that one actually feels his excitement as one imaginatively goes through his moves with him.\footnote{Deutsch, Fliess, Fenichel, and Kohut are using the concept of empathy in the way Lipps intended when they posit that the physician can actually experience the patient's concrete feelings through
an attitude of emotional absorption characterized by the suppression of the distinction between herself and the patient.

These psychoanalysts are aware that the concept of empathy as "feeling into" another's situation includes awareness of the distinction between self and other. But the issue is not whether the physician is aware that she is not actually the patient, but whether this awareness is internal or external to the experiential moment of feeling with the patient. Kohut's claim is that empathy involves two steps: First, in order to directly experience what the patient's concrete feelings are like, one must set aside awareness of the distinction between oneself and the patient; second, the clinician must reflectively distance herself from the patient to consider whether her experiential grasp of the patient fits with other data, including the patient's responses to her empathic communication.172

We emphasized earlier that empathy cannot be divided into a pre-interpretative experiential moment and a subsequent meaning bestowing reflective act: the experiential pole of empathy is itself an interpretive act. Kohut's contention is consistent with this idea. I take his point to be that in empathy one's own feelings are pre-reflectively informed by the project of understanding another, so that they actually seem to have as their source a "we" subject that unites oneself with another. But it is mistaken to presuppose that one must seemingly merge with another in this experiential sense in order to understand what her feelings are like. Edith Stein offers a persuasive criticism against this idea as it was presented.
by Lipps. She questions Lipps claim that in the experiential moment of empathy "there is no distinction between our own and the foreign 'I,' that they are one. For example, I am one with the acrobat and go through his motions innerly. A distinction only arises when I step out of complete empathy and reflect on my "real 'I." Stein agrees with Lipps that actually experiencing the foreign 'I' as a subject, rather, than just thinking about the 'I' as an object, is essential to empathy. But unlike Lipps, she emphasizes that this experience is an imaginative "announcement" and fulfilling explication of another 'I' rather than an imaginative merging with another 'I.'

Stein's point is that even within the moment of feeling with another, one does not take oneself to be in another's here and now situation. Following Husserl, Stein refers to the "here and now" situation as the "primordial" situation. Stein compares empathy with recollection, given that both are imaginative experiences that "announce" the presence of a real 'I' that is not actually "primordial." Thus, these experiences differ from pure fantasy, in which one creates an unreal subject. She describes recollection of a past experience as a detailed imaginative reliving of the past. Recollection thus differs from mere recall in that in recollection, one re-experiences an imaginative connection to the past event rather than just positing that a certain event occurred which is causally related to one's present experience. But, Stein says, this filling-out experience of the past...

does not make the remembered experience primordial. The present
viewpoint of the remembered state of affairs is completely independent of the remembered viewpoint. I can remember a perception and now be convinced that I was formerly under a delusion. I remember my discomfort in an embarrassing situation and now think it was very funny. In this case the memory is no more incomplete than if I again take the former viewpoint.\textsuperscript{174}

The same gap exists between the 'I' who is the subject of the imaginative act of empathy, and the 'I' who is being empathized with. Stein points out that even if I am so absorbed in empathizing with the acrobat that I entirely forget myself, and pick up a dropped program without even "knowing" I did so, this does not show that I have merged with the acrobat. For if I reflect on the experience of dropping my program, it is apparent that this experience was given to me directly in the here and now (this experience had the quality of primordiality), even though this here and now is now past, and only given non-primordially in memory. However, if I reflect on the acrobat's acts, it is apparent that the other's action was only announced, but never given directly, in a past here and now of mine.

As a result of Stein's analysis, we can now clarify the essential "as-if" character of empathy. Stein shows that the "as-if" cannot be dissolved by the listener's self-forgetfulness. We misspeak when we describe self-forgetful empathic absorption as a total immersion in, or merging with another, because the inner quality of the act of empathy is never one of total merging.\textsuperscript{175} That is, it is not just the fact that
we must imagine what the other feels that makes empathy an as-if, rather than a directly intuitive experience. Even if one brackets the third-personal awareness that one is not really in another's situation and considers only what seems to be happening within the imaginative world of empathy, one does not get to a "merging" experience. It is the internal structure of empathy that requires an awareness that another's experiences are not actually presented within my own sphere of experience. This relates to a point made by Husserl, that an essential aspect of empathy is the awareness of the absence of the other in one's own "primordial" situation.176

The kernel of truth in the merging model of empathy is that the physician's capacity to resonate with the patient's feelings contributes to clinical empathy. We will turn to this issue next. However, we saw that such resonance does not explain how one can understand another's concrete experience. And we have argued against the idea that such resonance, in the context of bracketing one's critical faculties, leads to a merging of self and other in which one samples the other's attitudes in concreto. Thus the "merging model" does not give an adequate explanation of how the physician's emotional responses can represent the patient's concrete situation. It is interesting that Kohut, whose theory of empathic "merging" is worked out most extensively, ultimately reinvokes the old assumption that the physician must use her "merging" experiences as a first step in making inferences about the patient.177
The conclusion that empathy cannot be explained on the basis of a merging identification of physician and patient serves our project of coming up with a useful model for clinical distance. For medically oriented physicians who work under the pressures described in the introduction of this dissertation, the equation of empathy with merging would mean that empathy could only constitute a small fraction of their practice. Most of the time, they need more distance from their patients than the merging model permits.

The importance of maintaining an awareness that one is not actually in the patient's situation arises not only during invasive procedures and in the operating room, but even when the physician's primary task is to listen to and comfort a patient. For example, consider the physician caring for a victim of domestic violence. While it would probably help the patient for the physician to grasp the nature of her feelings of fear and mistrustfulness, it would most likely be disruptive if the physician actually felt herself to be a co-victim, surrounded by a frightening and hostile environment.

In addition, when the physician sets aside awareness that she is not actually in the same boat with the patient, there is a great risk of imposing the physician's preferences on the patient. Here, the act of "projecting one's personality" into the "object" of one's regard becomes the psychological defense mechanism of "projection" which is defined as "the unconscious act or process of ascribing to others one's own ideas or impulses or emotions." Consider, for example, the
transplant surgeon Christiaan Barnard, who identified his desire to pioneer the
technology of heart transplants with his patient's desire to survive. After his first
transplant patient died, he felt so disappointed that he ran to the next patient on his
list, and pressured this ill man into a procedure no one had survived. He said to
the patient: "I feel like a pilot who has just crashed...Now I want you, Dr. Blaiberg,
to help me by taking up another plane as soon as possible to get back my
confidence." Jay Katz shows how Barnard overlooked the meaning of
undergoing a transplant for the patient by conflating his own goals with those of
the patient. Even in the cases where a merging identification does not lead to
paternalism or physical harm to the patient, it conflicts with the basic goal of
empathy, which is to learn something about the patient's feelings. When one
projects one's own feelings onto another one does not learn anything new about
her.

A second theory of how the physician's emotions contribute to her
understanding of the patient's emotional situation comes from Michael Basch, who
was trained by Kohut. I will briefly summarize what I take to be the key
conceptual steps of Basch's theory: First, Basch takes the position that infants have
an inborn capacity to respond to their caregiver's feelings with "resonance"
feelings. Basch's theoretical claim is that the infant's responsive joy is a way of
perceiving her parent's joy. Basch's second claim is that as the person develops
and learns, these originally stereotypical perceptual patterns become enriched and
individuated by the associative context they become embedded in. That is, because they have a psychological significance from the start, these responses can be connected to other experiential phenomena: happy feelings may come to be connected with being in certain places, being touched lovingly, etc. Basch's point, which originates with Freud, is that the perceptual content of one's affective attitudes is essentially connected to one's associations. He argues against the idea that the physician must regress and experience a kind of infantile resonance in order to empathize with the patient. Rather, as an adult, the physician responds with differentiated emotions with particular foci.

Furthermore, Basch claims that the physician can apply her own images to the patient's situation through a process of "generalization" in which one attributes feelings to another person while maintaining awareness of the process of attribution and the distance between self and other. But this only explains how one's own images may be made relevant to the other person, not how one can receive new information from listening to the other. Extrapolating from Basch's model, one could argue that the physician is able to learn something new from the patient because her associations are expanded by the story the patient tells. But note that this still does not explain how another person's imagery can take on experiential import for the listener, and become incorporated in the listener's own associative context. Ultimately, Basch (like Kohut) relies on the traditional model of introspection and inference to explain how the empathizer's affective-cognitive
responses can be moved along from their own prior content to new content that is informative about the patient. First the physician observes the patient; then she feels resonance emotion; then she infers from her own feelings and her general understanding of emotional experience some tentative ideas about the patient's experience. But this does not explain how the physician's own emotions can be directed towards the patient's concrete situation, as they must be if they are to be sources of information rather than mere repetitions. Thus, while Basch's developmental view of emotional communication is quite interesting, it still does not explain how one's feelings can be in dynamic communication with another's experience such that they can represent her attitudes in new and informative ways.

A third view, that of Daniel Buie\textsuperscript{185}, takes heterogeneous modes of cognition, including resonance feelings, recollection, fantasy and conceptual knowledge, to contribute to empathy. These diverse experiences allow the physician to build a progressing model of the patient's world. Buie is not the first to describe empathy as involving the building of a mental model of the patient.\textsuperscript{186} However, the innovative feature of Buie's conception is his emphasis on the physician's affective responses to the patient as essential resources for the construction of this model. The physician can only develop and fill-out a model of the patient's situation-as-lived by experiencing an affective relationship with the patient over time. According to Buie, the landmarks of this map are the physician's imagined, recollected, and resonance feelings, which are organized around a guiding thread.
of conceptual knowledge about the patient.

Here is Buie's example of how recollected feelings contribute to the therapist's understanding of a patient. While listening to a patient who discussed feeling sad about an upcoming interruption of her treatment, the psychotherapist...

noted feeling mildly sad also and spontaneously asked 'three-year old sadness?' The patient grew sadder but relaxed as she engaged in working through her residual feelings about a traumatic separation from her mother at age three...the therapist realized that his empathic sadness was his own, arising from a similar enough experience in his childhood.187

In this example, the physician's capacity to imaginatively relive the sad feelings he connects with his childhood abandonment contributes to his understanding of how the patient feels. Yet Buie points out that the physician's own experience of abandonment had importantly different concrete features from that of the patient, and it is exactly these concrete features that the physician needs to grasp in order to understand how a particular patient is feeling. How did reliving an emotion from his own, only typically related experience, contribute to the physician's grasp of the patient's concrete situation? Buie gives a hint in this passage. Without an additional cognitive act of inference, the psychotherapist felt that his recollected sadness mapped on to a three-year old's experience, even though he was actually about five when his own traumatic separation from his parents occurred. Buie takes this to follow from the fact that the working model of the patient already
included as a salient feature an upsetting separation from the mother at age three. The physician's model of the patient's world directed his own affective responses to the appropriate aspect of the patient's situation.

A similar point is made in Buie's description of the use of imaginative imitation to understand how a woman patient feels about her sexuality, given her recollected experience of being molested by some boys as a little girl. The boys, who she had been playing with "pushed dirt and pebbles into her vagina. Her excitement and exhibitionistic pleasure turned to vengeful rage and shame when, in the process, the boys hurt her genitally and made fun of her." While the physician relied on some recollections from his own childhood experiences to elicit some similar feelings to those of the patient, "he became aware that his empathic understanding was limited." He then imagined himself "anatomically and emotionally as a little girl going through all the details of her excitement, trauma, humiliation, and rage." Buie tells us that this imaginative experience enabled the physician to communicate better with the patient about her sexuality.

Here Buie again illustrates the fact that the concrete features of the patient's situation directs the type and content of the affective experiences the physician uses to understand the patient. In this case, the physician made a deliberate effort to fantasize those aspects of the patient's experience that were not available to him through recollection and resonance. But while the decision to imagine himself as a little girl was voluntary, and the result of reflection, the fantasy that ensued was
not reducible to a deliberate act of thought. Rather, Buie implies that by moving imaginatively through the experiences of humiliation and rage, he was led to unexpected thoughts that contributed to his understanding of the patient.

But how can it be possible for one to experience feelings of abandonment as referring to a three year old whose mother was unavailable, when one actually felt similar feelings as a five year old whose father was rejecting? And how can a man evoke feelings in himself that pertain to a little girl's experience of being sexually molested? That is, how can one's own experiential responses be directed by one's appreciation of the feature's of another person's situation? Buie never addresses this point directly. Instead, he too ultimately invokes the idea that the physician connects her feelings to the patient's situation via non-affective inferences. Thus, despite the richness of his analysis, he still does not explain the mode of representation that allows the physician to see new, affective features of the patient's situation.

IV A Non-Inferential Model of Clinical Empathy

My explanation of how one's own affects can refer to the features of another's situation centers around the idea that the physician's modeling of the patient's situation is essentially an act of imagining how it feels to be in the patient's
situation, rather than an act of making inferences about the patient's situation or an experience of merging with the patient. I turn to some work by Edward Casey and Richard Wollheim on imagining how another feels, to bring out the key features of the imaginative representation of another's experience.

In *Imagining*, Casey describes "imagining how" as follows:

> We are capable not only of imaging (objects and events) and imagining that (states of affairs obtain), but also of imagining how to do, think, or feel certain things, as well as how to move, behave, or speak in certain ways. There is a sense of personal agency, or the imaginer's own involvement in what is being imagined, which is lacking or at least muted in instances of sheer imagining-that. To imagine how is to project not merely a state of affairs simpliciter (ie. one in which the imaginer is not a participant) but a state of affairs into which the imaginer has also projected himself (or a surrogate) as an active being who is experiencing how it is to do, feel, think, move. etc. in a certain manner.

The distinguishing feature of the "imagining-how" in the case of empathy is that the images and relations in the imagined world are organized from the perspective of an agent rather than for an external observer. For example, Casey says that in imagining how to lace up a boot one would imagine the kinesthetic sensations involved in the action of boot-lacing. The reader might challenge this point by invoking those fantasies in which one imagines how two persons do
something together; it might seem that in such cases one need not organize one's imaginings from the viewpoint of an agent. Consider, for example, that I imagine how person A gives person B a flower. If what I imagine could be described either from the perspective of the flower giver or the flower receiver, then the imaginative portrayal must itself be multivalent -- it must be open to more than one perspective rather than categorically organized from an agent's perspective. But, as Richard Wollheim points out, it is different to imagine how A gives B a flower than it is to imagine how B receives a flower from A. It may be possible to imagine both things, but only sequentially, never in one act. According to Wollheim, when one imagines how A experiences something, one "liberally and systematically intersperses imagining his doing certain things with imagining his feeling and thinking certain things." I take Wollheim's point to be that in imagining how A feels I follow the flow of A's feelings and thoughts as they would flow in life for the experiencing subject. If B's feelings are noticed, it is as they are expressed for A. For example, if I imagine B scowling at A, this scowl is felt as a rejection, which is its meaning for A, rather than as whatever it felt like for B. If more "interior" awareness of B's feelings intervenes, then one is now imagining how it feels for B to receive flowers from A, and A will only be noticed as B would notice A.

From my own experience I can identify several levels of "interiority" that characterize imagining another's experience from the inside out rather than the
outside in. First of all, there is the level of bodily experience, including kinesthetic sensations--one imagines how it feels to walk on crutches, to try to lift one's body out of bed when one is exhausted; one imagines pains, feelings of nausea, dizziness, etc. Second, there is the level of environment-as-lived -- not the geometric description of a room, but the room as it surrounds one who lives in it, how it looks, smells, sounds. Thirdly, there are the feelings and attitudes the agent has towards visible and invisible others and events in her world. Some of these attitudes will have a concrete focus -- for example, anger at a parent. Others will provide an overall atmosphere -- for example, being in a depressed mood. Looking back at the definition of empathy, we can now replace the metaphor of "projecting" oneself imaginatively into another's situation with the concrete activity of focusing one's imaginative productions on the various "interior" features of another person's experience.

To illustrate these levels of "interiority" here are some first hand examples from my experiences with clinical empathy. The first level, imagining bodily sensations, is one which occurred when patients were experiencing overwhelming pain and discomfort. For example, while caring for women in labor and assisting in deliveries, I watched women go through waves of contractions and release. I imagined how such contractions might feel -- the pressure in one's abdomen and pelvis, the pain that accompanies stretching of muscle wall, as in bowel discomfort -- and the feeling of comfort and exhaustion as each contraction waned. The
second level of "interiority" -- the sights, sounds, smells of the patient's immediate environment -- is one with which I am most familiar. During a two year period I gave regular, primary care to a group of patients. In retrospect I find that I have vivid pictures of the bedrooms, offices, etc. of these patients, even though I never visited these places. These pictures developed while I listened to particular stories. For example, in the case of Ms. D, who suffered from insomnia, I pictured her bedroom on the top floor of her house. I pictured the room as stifling, with no breeze, and high humidity, with one small window with dusty curtains, and utter silence and darkness. I pictured how exhausting and lonely it felt to drag oneself all the way downstairs, trying not to wake up her daughter and grandchildren, who would only reprimand her. I imagined mixing a drink and sitting in front of the television, trying to relax. Corresponding to a time when Ms. D. was recovering from her underlying problem (a major depression) and starting to be active and engaged with her family, I imagined her cooking macaroni and tomato sauce in her kitchen, with her grandchildren there. I did not actually "image" her grandchildren, but felt a nice feeling of having company and being part of a lively family. I did "image" the kitchen -- it looked much like my own kitchen, and was sunny and breezy. I even imagined the smell of the tomato sauce.

As for the third level of "interiority" -- the patient's attitudes towards others and specific issues, or her generalized moods -- the two aspects require separate examples. In the case of an attitude towards something, I will again use the patient
Ms. D. Ms. D had a divorced daughter, with children, who wanted very much to remarry. The daughter had recently come to live with Ms. D for economic reasons, and Ms. D had started to recover from her depression as her loneliness was alleviated. When she told me the "good" news that her daughter had become engaged, and was planning to move to another city, I imagined how it would feel to receive this news. I imaginatively experienced not only the expected feelings of fear at being alone, but also feelings of guilt for not feeling happy for "my" daughter. As an example of taking on the generalized mood of a patient, I recall my experience with a patient who was experiencing the rapid onset of senility. Mr. D described recent experiences in which he got lost going to work, and could not locate everyday objects at home. While listening to him I imagined how it would feel to discover such gaps in one's own abilities, and experienced feelings of embarrassment and shame. In a later conversation with Mr. D, he revealed that he felt ashamed, and that a desire to cover-up his mistakes was motivating his social withdrawal.

These examples, like Buie's examples, bring home the basic fact that the features or images that give content to the imagined subject's world are features of the patient's situation. That is, they are features of the patient's situation as the physician grasps and portrays this situation. This helps explain why empathy takes time, and is essentially dependent on communication. In the first encounters with a patient, a physician will depend largely on her generalized concepts of what
children, or older men, or people with cancer, are like. Hopefully, these concepts will be rich ones, derived not only from medical texts, but from personal relationships, previous interactions with patients, reading literature, etc. But it is only after the particular patient supplies the physician with image-laden stories that the physician can begin to fill-in the specific features of the imagined subject's world.198

In describing the levels at which one can imagine how it feels to be the agent in the patient's drama, I stumbled over the difficulty of identifying our imaginary protagonist -- for example, did I imagine that I now had a daughter, or that I was embarrassed, or was the subject of my imagining-how some representation of the patient, and not myself? It is not insignificant that I was able to say "she" or "I" or to leave this unspecified in the above examples. Casey199 points out that in imagining how an experience feels, the imagined subject can easily be left unspecified. Wollheim and Casey both argue that just because the features of an imagined situation are features of a real person's experience, it does not follow that one must imagine the experience as happening to that real individual.200 Both authors also imply that just because the affective elements of the imaginative experience have as their source the person who is doing the imagining, this does not entail that the imaginer must imagine that she herself is the agent. The general point here is that the knowledge one has of someone's concrete situation need not enter into the imaginative portrayal as descriptive aspects of the imagined subject.
to provide the content of that subject's experience. One need not specify that the imagined agent is a three-year old to imagine the world as viewed from a three-year old's perspective. The "spontaneity" of imaginative experience seems to rest on just this power to landscape the scene with images from disparate sources without having to explain their occurrence, or apparent unity, by the ordinary rules of empirical experience.

This helps us see more clearly the error in the "merging" model of empathy. The "merging" model presupposes that the fact that empathy involves an integration of the physician's own affects and the patient's images entails that the physician must imagine herself in the patient's situation. But the examples above show that one need not specify that it is oneself that is in the imagined situation in order to imagine from an agent-centered perspective.

But this leads us back to a more precise form of the question that we asked previously: How can the physician's own affects refer to the concrete features of the patient's situation in an informative way, via imagining how? That is, how can one's knowledge of the concrete features of another's situation -- for example, the fact that another was three when her mother left -- lead one to feel emotions that are appropriate to these concrete circumstances? And, going in the other direction, how can these feelings lead one to understand and perceive new concrete aspects of her situation that one has no prior first-hand experience of?

The difficulty we face in answering these questions is that our models of
cognition presuppose that the event of imaginatively portraying another's experience is in principle independent of the event of responding affectively to another's experience. The assumption is that the spontaneity to see things differently resides in a cognitive act of imagining, which is somehow causally related to, but independent of, the emotional experience of feeling with another person. The underlying assumption here is that emotions are generalized pre-programmed reactions to typical situations that cannot in and of themselves incorporate new features.

If emotions are ingrained responses, as Descartes thought, then we need to find a mediating pure cognitive act that explains how a non-specific experience can be made into material for building a mental model of another's world. But this forces us (like Basch and Kohut) to reinvoke the same problematic conception of empathy as inference that we rejected earlier. For what is to guarantee that one's imaginative portrayal and one's affective responses will be appropriately linked? We must posit an additional act of inner cognition that attunes the picture thought has painted in the imagination to one's typical emotional responses. But this search for an inner translator leads to an endless regress.202

The feelings that contribute to one's understanding of another are not independent events that result from imagining another's situation. Rather, they are an essential aspect of the event of imagining how another feels. This distinction is elaborated by Wollheim. He says that there are cases in which an event of
imagining how something feels causes an independent event of affective response: For example, imagining how some forbidden sexual act feels, when I reflect on it from outside the fantasy experience, causes me to feel shame. But Wollheim contrasts this with the kind of feeling that characterizes the act of fantasizing itself: while imagining how the act feels, I feel erotic pleasure. Wollheim's point is that there are affective experiences that occur in imagining how this or that feels. These feelings are an integral part of imagining how something feels. Given that the affect is an aspect of the event of imagining how, it makes no sense to ask for a chain of causes between the event of imagining and the event of feeling. Since they are not separate events, they can have no mediating causes.

Returning to our problem, we see that it makes no sense to look for a mediator between the imaginative portrayal of another's experience and the emotional response to that portrayal. Rather, one's feelings are constitutive of the imaginative portrayal of another person's experience. In empathy the physician's concepts and feelings work together as they do in her own first-hand experience. In one's first-hand experience, one's thoughts and feelings are aspects of a unified "totality" that comprises the experience of being in a "world" (Heidegger's concept). In empathy one uses one's imagination to produce a world that also has the character of an experiential "totality."

The picture of the empathizer as moving "from" perceptual images created by active mind to passive affective responses, and from passive affective responses to
the creation of new perceptual images, is misleading. The "from" does not pertain to moving through time in the sense of a chain of efficient causes, but to moving through depths of experience, in which one fills in the affective elements implicit in the patient's story via one's own affective responses. But this shows that empathic "imagining how" is radically different from other cognitive acts that depend upon moving from one idea to another via inference. There is no comprehensible chain of ideas that can be detached from the affective experience of the empathizer, and presented as a logically complete chain of ideas. The movement of ideas in empathic understanding is essentially affective and experiential. In this sense the thinking involved in empathy is more like the thinking involved in dreaming, in that one image is often connected to the next because both images, however different they might appear to a detached observer, express similar feelings. Hence the standpoint of detachment is incompatible with empathic imagining how.

The idea that emotional resonance is possible because humans share typical emotional responses, need not be taken to imply that emotions lack the spontaneity to take on new objects. In chapter two we argued against Descartes that the typicality of emotions does not imply that they are like blind reflexes. Rather, emotions are typical because humans share not only biological traits, but paradigmatic ways of encapsulating the dramatic features of experience. We can respond with the appropriate affect to another's concrete situation because we share
common paradigm scenarios. But, on the other hand, we can imagine new concrete aspects of another's situation because each emotional experience of our own spontaneously incorporates new features of experience. If emotional learning did not occur continuously throughout our lives, then we could only love others who were exactly like our first love objects, only be interested in activities exactly like our first activities, etc. But even in cases of psychological disturbance such rigidity can only be approximated, given the natural mutability of the foci of emotions (hence the instability of the fetishist who can never really relive the fantasized scenario). Normally, the paradigm scenarios defining emotion types are general and flexible enough to allow for the incorporation of new emotional experiences as somehow already familiar. (As we mentioned earlier, one example of this kind of mutability without an awareness of strangeness is the "primary process" associations that occur in dreams).

The normal mutability of emotional objects explains much of the spontaneity of clinical empathy, but not all of it. As Osler and Hooker recognize, physicians are often asked to understand experiences that are normally kept out of awareness. In order to empathize with "abnormal" experiences physicians need a way to cultivate unusual emotional flexibility. Buie shows the flexibility to go beyond his "natural" scope when he imagines what it would feel like to have the body of a young girl. Once started on this new path, the images of being molested as a little girl become the target of Buie's own familiar (presumably childhood) fears of
bodily harm by strong, cruel others. I hypothesize that the variance in physicians' capacities for empathy relates to the degree to which different physicians are free to expand their emotional focus beyond the range of their previous paradigm scenarios. This would support the notion that empathy can be enhanced over time, since processes like consciousness-raising and psychoanalytic exploration (including in time efficient intensive seminars and groups) seem to greatly enhance this kind of flexibility. In chapter five I will focus on the capacity for moving oneself beyond familiar experience, which I see as rooted in the affect of curiosity.

V Some Limitations of Clinical Empathy

The capacity to understand the patient experientially enables the physician to take a better history, to communicate with the patient more accurately, and to form the kind of alliance with the patient that is needed for treatment to be efficacious. In the next chapter I begin with a clinical case to illustrate some of the practical features of clinical empathy. However, there are several limits to clinical empathy that we also need to consider in chapter five.

First, given that the physician's goal is to accurately understand the patient's situation for the purposes of diagnosis and treatment, what she learns about the patient must lead to working hypotheses that can be refined, rejected or approved
through dialogue with the patient; and her imaginings must be guided by concepts and even abstract theories. Hence whereas the activity of "imagining how" cannot be reduced to an act of reflective thought, this activity must be responsive to the physician's reflective thought about the patient.  

Second, the fact that in empathy the physician can never actually experience the patient's here and now situation, but can only grasp it approximately sets important limits on the accuracy, completeness, and efficacy of her imaginative understanding of the patient. As a result, physicians need to be aware of the gap between their empathic creations and their patients' experiences, and to recognize their patients' expertise with regard to their own attitudes. And they need to be aware that empathy is only one mode of understanding a patient, and that even with regard to emotional matters it may not be the best source of diagnostic and prognostic information. For example, psychiatrists assessing patients for the seriousness of their suicidal or homicidal intentions are sometimes mislead by their empathic grasp of the patient's attitudes, and would find a better source of information in looking at the patient epidemiologically.

Finally, empathy does invite strong emotions at times that may influence the physician to behave out of role. But less empathic physicians also experience many other invitations to react emotionally to patients at the expense of their doctoring. Consider, for example, how lust or anger can disturb one's capacity to listen and express ideas. With these issues in mind, I now give, in the final
chapter, an account of emotional regulation that is compatible with emotional involvement and aims for empathy and respect for patients. By criticizing the pictures of clinical empathy as detached insight and sympathetic merging, I have laid the foundation for a more realistic account of emotional regulation necessary for medical practice. We can leave behind the demand that the physician split off her cognitive capacities from her affective capacities, and consider how she can better serve patients by using all of the elements of her personality.
Chapter Five: Regulating Emotions in Caring for Patients

I disliked him immediately. He was one of the last patients I worked up during my internship, a seemingly interminable year that was finally coming to a close. His diagnosis was inoperable lung cancer, the deadly seed of which had already blossomed in his liver and bones. He had refused chemotherapy in the past, consenting only to local radiation therapy. Concerned about his increasing confusion, hostility, and disorientation, his family had brought him that night to the emergency room for evaluation.

As I approached him, I could sense his hostility. He looked older than his 72 years; the ravages of cancer were obvious in his pasty complexion and marked muscle wasting. His eyes sat deep in their sockets, and the sparse white tufts of hair on his head reminded me of dead trees on a mountain ridge.

"I don't like doctors and I don't want to stay," he began. "Who are you anyway? You'd better not be a medical student! The last time, a student tried for an hour just to get some blood from me!" [So begins an interaction in which the patient continually vents anger towards the physician, but submits to an evaluation which shows no physical cause of his agitation and combativeness.]

I decided to look for emotional or psychological reasons for his hostility. With trepidation I approached his bed, sat down beside him, and asked, "Would you mind telling me a bit more about what you've got wrong with you?" He said he knew he had metastatic lung cancer, that it was going to kill him, and that he was willing to accept his fate; it was his family that kept bringing him for evaluation "at the drop of a hat."

"Makes you sort of mad, doesn't it?"

"I just want to die at home," he said. "Save your fancy technology for someone else. I don't want your tubes and catheters; I want to go while I'm still in charge of my life."

When I admitted that tubes and catheters wouldn't help him get well, he relaxed and much of his hostility disappeared.

"When I first found out I had cancer," he said, "I denied it, like anyone would. Then I became angry -- with my family, my friends, my doctor; I blamed them all for what was happening to me...I was mad as hell, too, because they were all telling me what to do. I was the one with the cancer, they were the ones making the decisions. I read up on cancer"
and talked with people who had had chemotherapy. I decided I had lived a good life and the doctors could keep their chemotherapy and its side effects. I had done all I wanted to do in this world and it was time to leave it my way." A certain belligerent set came back to his jaw, and he said, "I can honestly say I'm at peace with myself and my decision. Tell my family that," he said, tapping my chest with his finger. "Tell them there's nothing you can do -- and nothing I would let you do even if you could!"

We spoke for a while longer ..."I'm in pain most of the time now, but I was brought up not to show suffering -- to be stoic. Lord knows if my family learns about the pain, they'll hover over me like I'm a helpless baby. I know I'm giving them a hard time, but fighting is all I have left -- all I have to remind them that I'm still capable of running my life." I asked if he could talk about all this with his family, and he said, "No, it's my problem, and I'm going to stay in control."

I admired his courage, and I wondered if I would be as brave in the same situation. I learned more about the complexities of cancer from him than from any textbook I had read. And when I left that day, I found that my hostility, like his, had disappeared.

I had mixed emotions when he was discharged the next morning. He had his wish to go home. For that I was glad and I hoped that he wouldn't have to come back, but I knew that I would miss a man I had come to like.

In this final chapter I offer an account of how physicians can regulate their emotions in order to meet the demands of medical practice for objectivity and efficacy as well as for empathy and respect for patients. In this clinical vignette, Dr. Linett (herein Dr. L) faces the double task of regulating his anger and disdain, as well as developing empathy and respect. One result of my analyses of empathy and respect is that I no longer find the phrase "clinical distance" adequate for what physicians need to develop. Being too distant (i.e. failing to respond affectively to patients, and thus failing to understand them and to have adequate moral relationships with them) is as problematic as failing to contain difficult emotions.
I choose instead to talk about "regulating" emotions to refer to the process of responding appropriately to patients.

This vignette brings to the fore the following philosophical problems. First, how can physicians' emotions involve being genuinely moved by circumstances and still be in some sense under their control? Dr. L is moved by the patient's threats, sadness, and courage, sequentially. Once the patient is seen as scary, or pathetic, or brave, Dr. L's emotions are already part of these portrayals, in the way fear is already part of the portrayal of the scary tiger. But if Dr. L cannot choose how to feel about the scary, pathetic or brave patient at any moment in time, how can he steer his own emotional course?

Second, as we concluded in chapter two, detached thought alone is insufficient to direct attention, so Dr. L's shifts in attention must themselves be affectively mediated. Does this mean that the shifts in emotion that Dr. L undergoes are determined by the situation Dr. L happens to find himself in and how things in that situation happen to move him? If so, there would be no room for freedom or control of the sort Dr. L apparently demonstrates.

The error in taking Dr. L to be passively moved by external circumstances from one emotion to another is to conflate the idea of the focus of an emotion, which logically determines the emotion one is in, with the actual situation of the agent in the world, which plays an important but not all determining causal role in what emotion the agent will experience. The focus of the emotion does logically define
the emotion one is in, but a focus is not just what is happening in the world -- the person's actual situation as viewable from multiple perspectives. Rather a focus portrays an aspect of being in the world that the person holds in mind in a concernful manner. What is deceptive about examples like seeing a scary tiger, is that the agent's situation appears to determine the focus of her emotion because almost anyone being charged by a tiger will focus on the tiger as scary. But this is an empirical fact, not a logical argument linking situation and foci of emotions in general. We can contrast this with a case of sitting next to someone on a bus, where the situation so little constrains one's focus that different people will have very different responses. One person may be curious about the person's book, another lustful towards the person qua sexual object, another annoyed at his taking up so much of the seat. In general what moves us in any situation we are in is not predictable by describing the "external" situation -- the view that ignores our biography -- without saying something about us, our thoughts and values.

The real difficulty we have in seeing how Dr. L's emotional shifts could be determined by his own agency is that we lack a picture of how Dr. L could "reflect" on and deliberately shift his own emotions without taking "time off" from being genuinely moved by circumstances. I will argue that such time off is not needed, by showing that the agent's shifts in emotion are themselves emotionally mediated, yet directed by the agent's deliberately maintained role.

My account of the kind of "practical reasoning" involved in emotional control
has four parts. In the first part I trace the shifts in emotions that Dr. L undergoes in this clinical encounter. In the second part of this chapter I develop a causal account of how physicians can make themselves shift emotions on the model of interpersonal persuasion. In the third part, I argue that in addition to acting on themselves persuasively, physicians need curiosity, which provides imaginative freedom to invoke alternative conceptions of their situation. In the fourth part, I consider whether or not physicians need to forsake a realistic appreciation of patient's suffering for their other goals of respecting patients and maintaining therapeutic efficacy; this leads to an account of courage as an emotion that allows physicians to face suffering without detachment.

I Emotional Transitions in a Clinical Encounter

Let us begin by tracing Dr. L's shifts in emotion. In the first paragraph Dr. L shows that he is immediately aware of his dislike for the patient, and of how this dislike is fanned into hateful feelings in the context of his exhaustion and resentment at the end of a long year of overwork (his internship). In describing the focus of his dislike he mentions first not the patient's orneriness, but the fact that his lung cancer is "inoperable" and that the patient refused chemotherapy. The nature of the patient's illness and the patient's reactions are viewed together as a threat to Dr. L's own confidence in medicine. Dr. L then thematizes the patient's
provoking behavior, his "hostility" as the "chief complaint," the initial description of the problem that brings the patient to need medical attention. This makes it clear that Dr. L sees it as essential to his role as a physician to address and manage the patient's hostility.

In the next paragraph Dr. L meets the patient for the first time with a reluctance based on his preconception that this will be a difficult man to encounter --"as I approached him, I could sense his hostility." He first attempts to focus on the aesthetic presentation of the patient, rather than on how disturbed the patient's morbid appearance makes him feel. He does not feel chilled at seeing the wasted face of this man but rather experiences a kind of aesthetic pleasure, like a painter or poet, in noticing the "sparse white tufts of hair" like "dead trees on a mountain ridge."

However, the patient resists this distancing, aesthetic view of his situation, and immediately engages Dr. L humanly by telling him that he doesn't like doctors, doesn't want to stay, and further that he senses the youth and inexperience of Dr. L and doubts he is an adequate physician. Dr. L carefully plays down his inexperience (omitted section), suggesting that he feels inadequate before this patient. Through the night he attempts to curtail arguments with the patient by keeping their discussions short, aware of his own increasing hostility towards this angry patient and towards his responsibilities as a physician. He does a brain scan to look for physical causes of the "problem" (the patient's hostility) and then rather
than attempting to engage in further discussion with the patient, given his own condition, he catches the few available hours left for sleep.

In the morning Dr. L feels ready to try to do his job, which includes understanding the emotional causes of the patient's hostility from the patient's point of view. Despite his trepidation he makes himself sit down beside the patient, a literal enactment of his verbal invitation to the patient to tell his doctor about his first-hand experience of his illness. The patient rises to the occasion, and conveys his understanding of his illness and his sense of being abandoned by his family. Dr. L is sufficiently affectively attuned to the patient at this point to see how being dumped in the emergency room "at the drop of a hat" would be angering, and to acknowledge honestly and without the insecurity he felt the night before, that no medical interventions could really help the patient get well at this point. I infer that this kind of confession on Linett's part follows from an empathic appreciation of the patient's longing to have someone else face with him the hopelessness of medical treatment and the inevitability of his death, rather than warding off such experiences with optimistic therapeutic statements. Dr. L's deliberate act of sitting and listening to this disliked patient, as well as the patient's gift of sharing his experience, make it possible for Dr. L to shift into an empathic stance.

The patient becomes much less hostile in response to this shift in Dr. L. He shows for the first time an eagerness to communicate. He identifies his anger and owns it, finding increasing vitality in affirming his need to be in charge of his own
care. His sense of himself is solidified by being able to reprimand his family via reprimanding his non-abandoning, non-judgmental doctor "tell them there's nothing you can do -- and nothing I would let you do even if you could!"

We see in this vignette that empathic listening can be directly therapeutic. The patient is able to speak at a level that goes beyond ordinary conversation, involving the articulation of feelings and concerns that he was probably not even aware of. He identifies as the cause of his fighting with his family the need to protect himself from being a helpless baby. His recognition of his own denial is apparent when he says "fighting is all I have left." This is an ironic statement infused not with defensive anger but with grief and pathos, which however, are still partially disavowed. If Dr. L were doing psychotherapy he would have the opportunity here for a mutative interpretation, a statement addressed to the patient's unconscious that would attempt to unlink the idea that to fight is to protect against feelings of helplessness; such an interpretation might radically alter the patient's assumption that he needs to keep his family at a distance. Instead, under the pressure of time and also with an impending feeling of grief that this patient will soon die, Dr. L does not thematize for the patient the ironic use of his orneriness, but rather accepts the need for it with admiration for the patient's courage.

Dr. L shifts here from an empathic stance to a feeling of admiration which has as its focus not what the world feels like to the patient (it feels threatening) but the patient as viewed from outside as the hero of a narrative. He talks about admiring
the patient's courage and wondering if he could be as brave. In chapter four, I argued that this kind of identification in which one thinks about one's actual self in another's situation is not empathy, although it certainly builds attachment. Dr. L's felt reverence for the patient is also revealed in his statement that he learned more from this patient then from any textbook about cancer. Moved by admiration, Dr. L does not interfere with the patient's decision to leave the hospital and refuse further care, nor does he pressure the patient to talk more openly with his family. Dr. L thus apparently fulfills his moral obligation to respect the patient in the context of a feeling of admiration for the patient as a particular person who has a particular way of coping with his illness. I will consider later in this chapter whether such admiration is a necessary aspect of respecting the patient, and also how such admiration interferes with understanding the patient's situation fully.

Finally, Dr. L notes that he will miss the patient, whose discharge represents to Dr. L his ultimate parting. This brief but intense encounter touched Dr. L at a personal level; he feels a kinship with this patient. We could easily imagine Dr. L feeling emotional pain if the patient were to break down his defenses and show terror or a dissolution of self, or if he were to be mistreated by his family. By the end of their encounter, Dr. L's natural reactions to "bad" news about this patient would probably be more similar to those of a friend or family member than they would be to the deliberately non-judgmental listener who sat next to the patient on his bed and got him to start talking. In part three I consider what kind of
emotional perspective is needed for physicians to face the suffering and loss of patients who, as Dr. L puts it, they have "come to like" over time.

II Emotional Regulation Involves Self-Persuasion

Dr. L regulates his emotions in this encounter by shifting his affective focus in a way that enables him to be a better physician. He begins with a sense of the patient in the emergency room as an enormous burden to the overworked, abused intern he feels himself to be, and ends with a view of the patient as a kind of spiritual teacher to the receptive, maturing physician he feels himself to be. This does not necessarily imply a shift in beliefs, as our discussion of emotional inertia in chapter two made clear. Dr. L may have throughout the entire interaction the unwavering belief that he is a dedicated doctor trying to help an ornery man whose orneriness is a defense against his fear of helplessness. How can we make sense of the observation that Dr. L directs his emotions by directing what he focuses on, given that we have argued earlier that the focus is itself not a detached thought but a construct with built-in affect? How can Dr. L deliberately move towards the affective experience of feeling admiration for the courageous aspect of the patient?

Robert Roberts\textsuperscript{208} gives an account of shifts in how one construes something that allow for emotional control and redirection. He defines "construing" as "bringing some perceived paradigm, or some concept or image or thought to
bear.

His point is that construing is "noticing an aspect" of one's situation by seeing the current scenario in terms of some other thought or image. Roberts gives the following example. He imagines standing on a tall, wobbly ladder in order to rescue his daughter from a fire.

I do not need to cease judging the situation to be dangerous to cease feeling afraid. To cease feeling afraid (or start feeling less afraid), I need to refocus the situation in some appropriate way...Instead of construing the situation as a threat to my well-being -- say in terms of an image of myself plummeting 25 feet to the pavement -- I construe it as a rescue task -- say in terms of an image of walking down that ladder with my daughter safely in my arms.

The point of Roberts' example is to show that even in the heat of a very strong emotion like fear of falling to one's death, one has the capacity to imaginatively invoke alternative images that will change one's occurrent emotions. But what Roberts assumes but does not explain, is first, that one has the freedom to imagine another portrayal of one's situation, and second, that one can make the alternative affectively charged images -- about saving one's daughter -- efficacious in dampening (or exacerbating) the pre-existing emotion. In the third part of this chapter I show how the affective attitude of curiosity allows one the freedom to imagine things otherwise. But the discussion of curiosity presupposes that we can make sense of the causal claim that an imaginative reconstrual of one's situation
could be efficacious in redirecting one's emotions. I therefore begin by analyzing the idea that one can move oneself, i.e. that one's imaginative reconstrual could have causal impact on an occurrent emotion.

Like the person on the wobbly ladder, Dr. L faces a situation that upsets him by invoking an alternative construal of his situation. He approaches the patient with trepidation, and sits down by his bedside, asking him to tell his story. At that moment Dr. L was still feeling a kind of frustrated anger towards this devaluing patient (emotion A), and did not yet feel sufficient empathy or respect for the patient. Yet, he was motivated to be a good listener, out of a duty to be a caring and thorough physician. In part three of this chapter we will consider what enabled Dr. L to imagine his situation otherwise and to invoke his goal of being a good physician. But our present question is how an imaginative reconstrual of one's situation, whether of being a good physician, or of rescuing one's daughter, can cause one to shift away from anger or fear?

Given the holistic model of affect, image and motive that we argued for in chapter one, let us assume that the tacit thought about being a good physician was embedded in a particular emotion, such as pride, guilt or some other self-relational emotion. We do not know enough details from the vignette to say which of these emotions was in fact strongest for Dr. L, but let us say that it was pride that imbued his construal, so that he was happily motivated towards his idealizing thought of himself as an excellent physician who listens well even to difficult
patients. Let us call this prideful construal of his interaction with the patient emotion B. With this reconstrual of his situation, and an actual shift in his situation brought on by sitting down next to the patient, the patient begins to talk to Dr. L. From the vantage point of pridefully construing himself as a composed listener who can get into the patient's world, he is able to be moved by what it feels like to be impotent and abandoned by a worn-out family; he experiences the patient's sadness empathically (emotion C).

The Dr. L example brings out a point about the causality of emotions that is not obvious in the wobbly ladder example, but that I take to be crucial for showing how reconstruals can be efficacious. In this example it is not just an image or a thought, but a full fledged emotion B (pride) that is brought to bear on pre-existing emotion A (anger), and that shifts the agent out of A, so that he can go on to emotion C (empathic sadness), rather than A' (ongoing anger). The fact that B has causal efficacy can be expressed by a counterfactual statement -- without B acting on emotion A, Dr. L would have responded otherwise, with A'.

What Roberts fails to explain, and what we must explain, is what meaning of "cause" we have in mind in positing that one emotional construal has causal impact on another. If we cannot make sense of Dr. L's pride as causing the remission of his anger, then we cannot say that Dr. L is steering his own emotional transitions. But how does emotion B move the agent out of A?

In particular, what we need is a conception of B as a "psychological cause"
since it is in virtue of the psychological event of B that the agent moves from A to C. Donald Davidson says that to say that one mental event causes another involves the linking of two events under a reason explanation in which event A provides a justification of the occurrence of event B; in addition the two events must also fit under some type-type physical description (inaccessible by translation from the psychological) that explains how the former can be a physical cause of the latter.\textsuperscript{204}

Familiar models of practical reasoning involving beliefs, desires and actions, fit this conception of psychological causation. Beliefs and desires are reasons for the beliefs and desires they cause, and beliefs and desire complexes are reasons for the actions they cause. However, while Dr. L's prideful thoughts about being a good physician may explain why he doesn't want to be angry, they are not reasons against his angry construal of the patient.

We are thus faced with the particular problem of accounting for how a mental occurrence can be a psychological cause without being a reason for what it causes. Donald Davidson considers this problem, which on the face of it presents a contradiction, in his article "Paradoxes of Irrationality."\textsuperscript{205} He is trying to explain akratic action, in which an agent's desires/beliefs can be mental causes of her action without rationalizing her action. He concludes that in such cases there must be a kind of boundary between the cause and the effect of the sort that occurs when one person's desires/beliefs influence another person's action. The idea of
a boundary presupposes that there will be more rationality or internal consistency to separating the aspects of the agents behavior and explaining them as causally interacting sub-systems than there is to unifying all aspects of the agent's behavior. Such a separation allows one to invoke dual perspectives on the reason explanation of the action. As Davidson explains, using the example of someone whose wish to have a well-turned calf causes his belief that he has a well-turned calf: "What his wish to have this belief makes rational is that this proposition should be true: He believes that he has a well-turned calf. This does not rationalize his believing: I have a well-turned calf."²⁰⁶

The idea that Dr. L's construal of the world via emotion B acts as an external, persuasive influence on his pre-existing emotion A takes an explanatory step beyond our observation that B is a reason for the remission of A that is not included in the agent's first personal reasons for/against A. Consider how a person may be moved by another's interested approval from an angry to a listening stance. Dr. L is moved by his own pride in the same way, as a persuasive conception of himself that invites certain attitudes and makes others unlikely. The idea of an intrapsychic equivalent of interpersonal persuasion does not depend upon the idea of an unconscious influence, in that one may be quite aware that one is being affected by a non-reason and still be affected.²⁰⁷ Rather, the key point is that within one person there can be two independent forms of thought determining attention, which influence each other in the way that two separate persons may
influence each other.

If the breakdown of reason-relations defines the boundary of a subdivision of emotions, there is a close link between the possibility of emotional irrationality and emotional self-control. In chapter one we discussed how emotions can be wish-fulfilling -- the agent can construe situations in ways that serve her yearnings. This possibility is inherent in the idea that emotions are not just responses to the situation one is in, but rather involve the agent's construing her situation in one way rather than another. What is not obvious in our discussion of wish-fulfilling emotion is where the locus of the capacity to be irrational lies. Given that there is no "space" between the emotion and the construal that is its focus, there is no sense in asking whether it is rational or irrational for the agent in love, in fear, in anger, to see things as they do. Rather, what such questions presuppose is a perspective that is on the other side of a boundary of the sort Davidson describes. To ask (as I did in chapter two) if a physician's fear of the schizophrenic patient is rational, is to ask whether we can place a wedge between the psychological cause of the fear and what rationalizes the fear. If we can place such a wedge then the possibility that the fear is irrational must be considered.

By positing a boundary between emotion B and emotions A and C so that emotion B serves as a mental cause of the ending of A and contributes to the occurrence of C, we raise the possibility of emotional irrationality. The same capacity to influence ourselves irrationally makes it possible to influence ourselves
in a role-directed fashion.\textsuperscript{208}

However, the model of interpersonal persuasion being invoked here excludes the idea that the link between the causally interacting emotions will be \emph{arational}. That is, there is an important difference between pride causing a shift from anger to empathy, and a toothache causing a shift from empathy to anger. The toothache, which has a psychological component because it includes the experience of pain, does not meet the criteria of "psychological cause" that we have in mind. Bodily feelings like pain, cold, or fatigue acting on one's occurrent emotions cannot be analogized to another person deliberately influencing one's imagination. The way they cause feeling states is more like the way physical events like being rained on, or hit over the head, cause feeling states. We cannot say that a toothache \emph{persuades} one to feel angry, even though toothache's make people angry, and of course, someone may give a particular symbolic meaning to having a toothache.

In summary, we can now explain how the idea that emotions involve being moved by circumstances is compatible with the idea that the agent can reconstrue circumstances and hence direct her emotions. The idea of a partitioning of emotional experience allows us to posit that Dr. L's prideful reconstrual of his situation (emotion B) acts on him from the outside in the way that other people act on him emotionally. This could lead to irrational, wish-fulfilling emotions, but it need not necessarily do so. The causal connection between particular emotion
instances is itself insufficient to determine whether an agent is deceiving herself or directing her emotional responses towards greater understanding of things as they really are.

III Curiosity and the Freedom to Imaginatively Reconstrue One's Situation

Returning to our clinical vignette, consider how Dr. L's shift from anger to empathy is revealing of the patient's situation, whereas his shift into admiration is concealing. From the standpoint of admiration, Dr. L does not himself see, never mind point out to the patient, that the patient's attempt to keep his family from recognizing his pain by fighting with them about his care is futile. That is, the patient is vulnerable and suffering, and his family is bound to recognize this. Yet Dr. L's response is so strongly admiring that he is not cognizant of the pitiable aspect of the patient, of the way in which the patient is not brave but cowardly, in his inability to face the pain in his family's eyes.

Ironically, having persuaded himself into the standpoint of admiration, Dr. L seems to lack something like "freedom" to imagine the patient's situation otherwise. The idea of such a "freedom" includes more than the capacity to act on oneself persuasively that we have just accounted for by positing a boundary within mental life. For one can act on oneself in a way that shifts one to an idealizing emotion that invites no further imaginative exploration. For example, it is probable that
during the risky exit, the man on the wobbly ladder fixed his imagination on the
idealized image of rescuing his daughter to the exclusion of subsequent
reconstruals.

In such cases of holding in mind an idealized image in order to keep out other
images, the lack of freedom does not reside simply in the agent's inflexibility. We
can imagine a situation in which one carefully maintains a prideful feeling about
one's work, or an affectionate feeling towards a friend, despite experiences that
might lead a more "flexible" person to feel differently. Maintaining these
emotions may require acting on oneself persuasively. For example, one may
temper one's anger at an inconsiderate act by a friend by invoking loving thoughts
about how caring she was in some other circumstances. In such cases of
maintaining a chosen emotional attitude over time we speak of the agent as
committed. But in the case of Dr. L's admiring view of the patient, it is not his
commitment to finding something admirable about the patient (an issue we turn to
later) that is problematic. Rather, it is the observation that Dr. L's admiration
involves missing important aspects of the patient's situation. His idealizing
admiration is thus more like pride in work that overlooks the shoddy or weak
aspects of the work, or blind loyalty that leads one not to see that one's friend is
disappointing one. These cases involve boot-strapping, or wish-fulfilling
emotions, in which one "whistles a happy tune" to avoid shifting into other
perspectives that would involve less comfortable emotions. These examples
suggest that the idea of an agent being free to imaginatively reconstrue her situation already includes in it the idea that one can shift one's emotions in ways that reveal rather than conceal things as they really are.

What standpoint allows physicians to move themselves out of reality-concealing emotions like some instances of admiration, fear and anger? Given our arguments against the concept of the detached observer, we need to avoid reinvoking a model of "reflection" that detaches judgment about one's emotions from the realm of emotional experience. Thus I begin with the question of what kind of affective attitude helps one's regulate one's other emotions towards the goal of being realistic? I think the attitude of the physician necessary to enhance her openness to things as they really are is curiosity about her own and her patient's emotional reactions. The term "curious" comes from the latin "curiosus" which means careful, diligent; this term is akin to "cura", which means care, concern, and is closest in English to the word "cure." However, I take it that the physician's curiosity acts also as a way of distancing her from an overly involved concern. To show how, paradoxically, curiosity involves carefulness that can be therapeutic or curative, precisely because it is an emotion that liberates one's affective focus, I want to consider an example of the use of curiosity in clinical practice in some depth.

Because the physician needs to be curious about conflicting emotions, whether within herself or between herself and the patient, or within the patient, the clinical
example I choose is of a couple therapist functioning as a participant-observer. The efficacy of the couple therapist resides in her capacity to combine an empathic grasp of both partners' positions with imaginative reconstruals of their situation that are outside of their joint view of reality. For example, consider a case in which a husband feels he must bear the responsibility for the finances and feels used by his wife, and the wife feels guilty but also resents not having a say in their decisions because the husband is so controlling. The therapist's empathic participation in each of their points of view, invites each of them to get empathically involved with each other. Once the husband and wife can recognize each other's independent emotional construals of their shared situation, they are liberated to manage conflict safely for the first time because they have the tools to make-up after an argument. For example, the husband can then genuinely understand why his wife feels controlled by his making the budget decisions, while still feeling his own concern that she would put her needs first and take advantage of him if she did the budget.

However, this empathic recognition in no way challenges the couple's shared assumptions about their situation. The more difficult, subtle and powerful aspect of the therapist's job is to recognize the collusive way the two members of the couple construe reality so that both of them can avoid certain risks and responsibilities in the relationship. (Often, it is not what the couple argues about, but what they are compelled to agree about that most constrains their relationship).
In this example it may be that both husband and wife construe their gender roles as requiring a conception of him as strong and able and of her as weak and dependent, and both members may believe that if she earned money, this interaction would be threatened. The therapist must be able to raise their consciousness about the way their joint affective construals of reality creates a kind of ideology that conceals as well as reveals what is at stake in their relationship.

In order to challenge their collusive construal of reality the therapist needs to move from an empathic grasp of their emotions to a non-empathic curiosity about the taken-for-granted scenarios that are the foci of their emotions. As we argued in chapter four, empathic portrayals are not hypothetical. In contrast, curiosity is hypothetical -- the therapist becomes curious about whether or not it makes sense for the husband to see his wife as frivolous, when she may be capable of more responsibility than he gives her credit for.

The therapist's curiosity differs from empathy not only in that it takes up the couple's emotional judgments hypothetically, but also in that it considers the way their emotions function instrumentally. The curious therapist sees the couple's emotions from "outside," and posits teleological causal connections between their emotions. For example, the therapist will begin to wonder about how the husband's view of his wife as irresponsible and helpless relates to his own self-esteem. In chapter one, we argued (against Sartre and Solomon) that emotions can function strategically and still represent things as they really are. This is
because emotions are always partial grasps of things as they really are that conceal as well as reveal. In considering the strategic interplay of the couple's emotions, the therapist need not consider the couple's emotional construals to be lies or illusions. Rather, she abstains from taking any particular emotional construal to be an adequate portrayal of things as they really are.

It is only when the therapist engages the couple to become curious about their own collusive emotions and how they function strategically, that the possibility of real emotional change arises. I think there are two forces that effect change. First, by negating the all or nothing quality of their emotional construals, and accepting that there are new, as yet unknown ways of responding to their situation, the couple is freed to experiment with trial reconstruals of their situation. This openness is necessary but not sufficient for change, since the trial reconstruals will not move the couple unless they fit with things as they really are. For example, the trial reconstrual of their situation as involving a wish to rely on each other more and risk overcoming sexual stereotypes would fall flat if in fact the husband did not yearn to be taken care of and the wife yearn for more responsibility and respect.

A second force for change comes from recognizing the strategic nature of particular emotional responses, because such recognition in and of itself shifts the locus of responsibility for, and control over the emotions. Much of the irrationality of emotions -- their power to influence one against one's all considered judgments,
and to conceal aspects of reality -- depends upon their strategic function remaining hidden. For example, Melvin Lansky argues that flirtation and intimidation are efficacious only insofar as they keep the seductive or threatening message as a kind of disowned backbeat to an emotionally benign foreground communication. If one recognizes and calls the person on their disowned seductiveness or threat, they can no longer flirt or intimidate, even if they retain their sexual longing or anger and reexpress it in some other way. In this example of couple therapy, consider how the husband may use feelings of victimization to invoke guilt in the wife in order to control her behavior. By recognizing the strategic function of the husband's feelings of being used, the couple loses their usual modus operandi. The wife who recognizes that she is being made to feel guilty shifts the locus of responsibility for the husband's suffering from her shoulders to their interaction; this alone may dissolve her guilt and with it her capacity to be covertly controlled by her husband.

The husband's recognition of his own covert power in using guilt to control his wife contradicts his feared conception of himself as ineffective in influencing his wife; this recognition alone may empower him to use his influence overtly and ask his wife to share responsibility for the finances.

These examples rely on a point made in chapter one, that just because emotions involve actually being moved by circumstances does not entail that one is by definition at the mercy of one's emotions rather than a shaper of one's emotional responses. However, from within any particular emotional response,
which involves experiencing oneself as moved by circumstances, there is necessarily a blindness to the way one's emotion functions strategically in constituting one's situation. The therapist's capacity to stand outside any particular emotional construal, and to engage the couple in imagining how their emotions may function strategically, is her most powerful tool to effect change.

How can we apply all this to the example of Dr. L? The metaphor of the couple therapist is meant to relate to both Dr. L's work on his own emotional conflicts, and to Dr. L's work with the "couple" consisting of himself and the patient. Dr. L becomes curious about the patient's particular experience of being ill, dumped in the hospital, and faced with one doctor after another. Yet Dr. L's curiosity involves a distinct shift away from a purely empathic standpoint in that he keeps one foot outside the construals that the patient takes to coincide with reality. Dr. L wonders whether it is really the case that the patient will inevitably be dehumanized by the medical situation, and wonders whether the patient is unaware of, or has not yet been offered, the opportunity to connect with a caring physician, and to feel like an active participant in his medical care. Dr. L is curious about what particular meaning seeing himself as the passive victim of medical interventions has for this patient.

Dr. L also turns his curious gaze onto his own threatening emotions. He notices that his initial aversive feeling towards the patient, which includes a concrete sense that the patient is impossible to communicate with, is a defense
against underlying feelings of helplessness about the patient's illness. Apparently this awareness lessens his sense that it is hopeless to try to talk with the patient, and makes it possible for him to invoke a prideful and hopeful reconstrual of himself as a dedicated physician who can help this man. With hope, pride and courage, he sits with the patient and asks him to tell his story.\textsuperscript{212}

In Dr. L's case, the use of curiosity to decenter from his own hopelessness and anger facilitates his capacity to listen to the patient empathically.\textsuperscript{213} Dr. L no longer focuses on his own experience as the object of the patient's hostility, as he decenters his attention from responding to the patients accusatory remarks to considering what it feels like to be this particular patient. This focusing of his attention on the patient's emotions invites, but cannot guarantee, empathic resonance.\textsuperscript{214}

Although his curiosity about the patient's hostility decenters Dr. L from his own anger and helplessness, and enables him to move himself with pride and courage to listen to the patient, there is a danger in such decentering. Dr. L focuses on the patient's anger and the response it invokes in others as itself a "curiosity" -- a reified problem requiring the attention of a physician who can transcend such reactions. In saying: "I decided to look for emotional or psychological reasons for his hostility," Dr. L is using a kind of role-legitimized curiosity to avoid looking at an intense human encounter that implicates himself as well as the patient. This would be patronizing and incomplete if it did not
include attention to his own vulnerability to overreacting to such a patient. That he is also curious about his own reasons for reacting to such a patient is clear from his writing a story about his interaction with the patient in which he must portray his own character as well as the patients. He describes himself in the story as tired, overworked and insecure about his own role. One way to teach medical students to become curious about their own emotional responses to patients is to invite them to write stories about their interactions with "difficult" patients.

How is the affective standpoint of curiosity different from the standpoint of the detached observer? First, there is no affective "truth" that is the object of the curious physician on the model of aperspectival truth for the detached observer. That is why, for example, couple therapy requires behavioral interventions in which the therapist asks the couple to try out some new behaviors to see what new ways they can relate. The therapist cannot know from some process of "reflection" that the husband's and wife's construals of each other are "wrong." Rather, the therapist can only question their conviction that their construals coincide with reality. The basis of this skepticism is not a concern for truth, which is irrelevant here, but an interest in seeing what is concealed behind the sincerity of their emotions. This is not to say that the therapist thinks they are lying, but rather, that she questions their apparent lack of ambivalence, and their conviction that their being fully coincides with their assigned roles in the relationship.215 Similarly, Dr. L gives his patient and himself the opportunity to "try on" a new
interaction, after a night's sleep, given his awareness that his original, angry reaction to the patient's hostility conceals other reactions the two might have towards each other.

The second way the curiosity of the physician differs from a detached standpoint is that curiosity is not disinterested; it is profoundly interested but not driven by the need to secure oneself. The affect of curiosity is characterized by an unburdened or playful absorption in one's thoughts that allows one to forget oneself; one is not worried or concerned about securing oneself. Plato's and Aristotle's image for curiosity is the stargazer, whose mind is on the heavens rather than his own steps.\textsuperscript{216}

However, the image of the stargazer, which is the origin of the theoretical standpoint of the scientist as well as the philosopher, is not quite proper for the curiosity of the physician, because the doctor is not contemplating something distant, but something close to home, which he is suddenly able to see as strange, uncanny. Just as the family therapist must experience an empathic resonance with each member of the couple to be effective, the physician must have as the focus of his curiosity the particular emotional construals that he and the patient have. And as we argued in the empathy chapter, there is no road to understanding the particular affective construals of others except through empathic imaginative involvement.

However, there is an important way in which curiosity about emotions
dampens emotions. The physician's curiosity about the patient's and his own reactions decreases the compelling quality of their construals; it decreases the degree to which he will be actually moved by their plight, since part of what is moving is the idea that what is being depicted is exactly how things really are. But the physician wonders if this is the case. The standpoint of wondering, of entertaining hypothetical construals of their situation, involves a departure from an empathic to a non-empathic stance. Curiosity require suspending judgment in order to allow oneself to be uncertain about how things are; this standpoint departs sharply from the engaged believing quality of empathy, in which one is for example, saddened by that which saddens the patient.217

Dr. L's awareness that the patient's occurrent emotions do not capture every aspect of reality does not lead to the kind of skepticism about the patient's subjective experience inherent in the disease model of medicine. Dr. L is not like the accusatory detective who knows the suspect will kill again. This is because Dr. L refrains from reducing the patient's emotions to a causal explanation which seeks to predict the patient's future. He respects the patient's capacity to author his own future. By acknowledging his uncertainty about the patient's future, he gives the patient a kind of unconditional or non-judgmental regard.218

However, the fact that curiosity lessens one's ties to any particular construal of one's situation is what is behind the ubiquitous association of curiosity with danger. We say "curiosity killed the cat," and Thales' star-gazing ends with him
falling into a well, injured and subject to scorn. The point is not that curiosity is in and of itself destructive, but rather that it involves the risk of setting out to sea without guarantee of a safe journey. We can imagine how dangerous a physician would be if she were moved only by endless curiosity, untempered by empathy. In part four of this chapter I consider what overarching attitudes serve as the physician's road-map over time and assure that her curiosity will be tethered and put to good use.

This image of curiosity as involving a non-collusive, freely moving interest in the patient is resonant with Freud's image of the ideal analyst as having "evenly suspended attention."

Freud tells the analyst to teach his patient how to associate freely by invoking the image of being on a train looking out the window. I think this image is meant not only for the analysand, but for the analyst, since the source of data in analysis is the patient's free associations, and the therapist's attunement with the movement of these associations requires freedom to move with the patient from one focus to the next. When riding a train, one's gaze is not on one place but moves with the landscape. It is this kind of freedom, and not affective detachment, that I have in mind by invoking the affect of curiosity.

In summary, we have seen that Dr. L's capacity for emotional self-regulation is complex. The capacity to act on himself persuasively is a necessary condition for directing his affects. This is illustrated by the man on the wobbly ladder who dampens his fear by holding in mind a courageous construal of his task. But if
self-regulation depended only on the degree to which one could move oneself persuasively, then there would be no distinction between Dr. L's use of pride to shift from anger into empathy, and his use of admiration to cover over his feelings of disappointment. The difference between these two instances is that only in the former case does Dr. L show an ongoing capacity to imaginatively invoke alternative construals of his situation that broaden his understanding of reality. This capacity resides in an overarching affective stance of curiosity.

So far, our account of emotional self-regulation has two tiers. Dr. L's curiosity enables him to decenter from his own anger and hopelessness about the patient. This decentering creates the opportunity for him to alight upon a prideful reconstrual of his situation. His pride then moves him further from his anger, by focusing him on his goal of being a good physician who listens to the patient. All of this made it possible for him to sit and listen to the patient, which made possible but did not guarantee, his subsequent empathic response to the patient.

I have used the contrast between curiosity and idealizing admiration to identify a kind of "freedom" to invoke alternative construals of one's situation that get beyond wish-fulfilling depictions of reality. Yet I would still maintain that Dr. L is moved to wonder about the patient early in his encounter, and yet moved to idealize the patient later. This suggests that the freedom to imaginatively shift gears is not radical. The view of emotions we have defended, as involving being moved by circumstances and determining salience prior to any act of will, entails
that Dr. L's capacity for self-regulation will itself be subject to the conditions of his situation, his bodily condition and habits. For example, Dr. L regulates himself by getting some sleep prior to continuing his discussion with the patient. Dr. L's emotional dispositions will also depend a great deal on his acculturation into the "world" of medicine over time, so that directing his own emotional course will require indirect action via transforming that "world". For example, Dr. L influences the conditions of his socialization into his professional role by sharing stories with other physicians about emotional interactions with patients, rather than just reporting statistics about procedures.222

IV What Affective Standpoint is Needed for Physicians to Maintain Their Integrity Over Time?

Our account of how Dr. L acts on himself persuasively has left us with two questions about what is required for such self-regulation to be in the service of Dr. L's goals as a physician. First, does respecting the patient require idealizing admiration? Second, does the goal of understanding the patient accurately cohere with the goals of respecting the patient and being an efficacious healer?

First of all, it might be argued that Dr. L needs to admire his patients in order to respect them as persons. In this vignette, it may appear that pitying the patient would be devaluing, and would therefore do harm to the patient. Further, if one
favors the Kantian value of respecting patients as end-setters over beneficently
protecting the well-being of patients, than admiration might appear to be more
appropriate than pity. Dr. L's admiration for the patient leads him to see the
patient's choice of refusing care and refusing discussion with his family as rightful.
In contrast, a pitying view of the patient as fighting a futile battle to deny his pain
and his need for familial support might have lead to questioning the
appropriateness of the patient's decision. Does this mean that in order to honor the
patient's own preferences with regard to his health care the physician needs to
admire some aspect of the patient and to avoid seeing the patient as pathetic? That
is, does respecting the patient as an end-setter depend upon feeling admiration for
the patient, so that what one is obligated to do is to drum up admiration in every
case?

The idea that the proper Kantian physician has an obligation to reconstrue until
she can admire all patients is in fact a perversion of our attempt to root Kantian
morality in affectivity. The core of the Kantian conception is the idea that the
moral value of actions is independent from the particular affects that motivate the
actions. The revisionist version we argued for in chapter three suggests that
affective ties are needed to draw one's attention to actions that are capable of being
morally valuable; if one then acts according to maxims that are universalizable, one
is acting morally. What may not have been obvious in our discussion of Kantian
ethics is that the claim that affects are necessary to perceive what is morally salient
is not a defense of a kind of moral relativism in which the feeling of respect, wherever it can be generated, bestows moral value on the situation. If this were true, then the moral value of a situation could be amplified by self-deceptive admiration. What is morally valuable is not having a certain experience of sentiment, but rather being genuinely moved by worthy aspects of another person's situation. Only the agent who is sensitive to what situations actually are morally salient, and then acts in such instances in a way that is generalizable, can be a moral agent.

Further, being sensitive to what is morally relevant in human life depends upon seeing things as they really are. The claim that an adequate moral perspective requires perceptiveness and judgment about what is more or less significant in the human world, entails that the moral agent ought to strive to see things as they really are. In our example, Dr. L's final admiring view of the patient prevents several morally relevant connections from occurring. Dr. L misses the chance of caring for the patient without idealizing him, recognizing his cowardice as well as his bravery and still maintaining his connection to the patient. And the patient misses the opportunity of having his feelings acknowledged by a realistic, caring physician. Such recognition from another might have enabled the patient to recognize that his unspeakable fear about exposing himself to his family could be dealt with as a manageable problem. This might have led to a third moral opportunity, the opportunity for the patient to value honest self-disclosure at the
risk of disappointing his family.

Finally, there is no reason to assume that either experiencing the affect of respect, nor following the rule of respecting the capacity of others to determine their fate should require idealizing their actions. One can feel respect for someone in light of his or her foibles as well as his or her strengths; Dr. L could respect the patient for trying his best given his particular fears. Consider, for example, how women involved in consciousness-raising in the seventies and eighties found that it was only when their male companions removed them from the pedestal of feminine goodness and acknowledged them realistically in light of their strengths and weaknesses that they felt respected as peers. Finally, the Kantian commitment to respect the patient as an end-setter involves supporting the patient's capacity for self-determination, even when the patient's choices are not admired by the physician. (In the last section of this chapter I argue against the idea that physicians ought to direct patients to make choices that physicians consider admirable). In summary, respecting the patient as a person does not require self-deceptive admiration for the patient, nor does it require coercing the patient to behave admirably.

But this problem is part of a more general problem of whether the functioning of physicians over time is best served by emotions that are revealing or concealing of things as they really are, given that much of the reality of medical practice is painful and depressing. For example is it better for physicians to feel realistic
worry about their patient's outcomes so that their hopefulness is guarded, or would they have a more hope-generating effect on patients by feeling unqualifiedly hopeful and optimistic even in situations where there is a low probability of recovery? Is it better for physicians to realistically feel grief at the loss of a patient and to feel realistically guarded at attaching to the next patient, or would a self-deluding feeling of buoyancy leading to an intense interest in attaching to each and every patient be better for meeting the goals of medical practice?

Rene Fox describes the way physicians use magical thinking, which covers over things as they really are, to defend themselves in the face of therapeutic limitation, uncertainty, and existential concerns about the meaningfulness of their work. Fox quotes the anthropologist Malinowski, who said that physicians "ritualize their optimism." They maintain confidence about the effectiveness of their procedures even in the face of strong evidence to the contrary. This helps them to endure and persist and to encourage patients to do so, and thus helps stabilize some patients who otherwise would never have made it, yet also leads to "the nonrational inability to desist, at great physical and psychic, as well as economic, cost to both health professionals and those for whom they care." Fox wrote this in 1980, but in 1993, with the overriding impact of economic concerns on medical practice leading to rationing, I think physicians are just as likely to use magical thinking to ritualize their pessimism, and predict the futility of care in order to desist from care without experiencing their uncertainty.
While we cannot address here these empirical questions about how magical thinking effects medical care, we can return to our clinical example to see how idealizing the patient effects the physician's efficacy over time. Certainly Dr. L's admiration for the patient made the patient feel special, and made Dr. L feel special about his interaction with this otherwise very demoralizing patient. But consider what would happen if the patient returned to Dr. L's service in some overtly regressed condition, with a decline in his health related to non-adherence to medications, and with uncontrollable rage at his family and at Dr. L for his condition. Most likely Dr. L's idealization would be crushed and he would experience disappointment. Often it is just this repetitive feeling of disappointment in patients who do not fare well, that motivates an attitude of detachment in physicians who cannot bear feeling their efforts to be futile. In this vignette, we might speculate that the nurse and others who have repetitively cared for this patient treat him with avoidance and resentment out of such disappointment, which is the flip side of idealization.

This suggests that where there is a continuing physician-patient relationship over time, the efficacy of the physician is best served by a realistic rather than a deceptive appraisal of the patient. However, this needs to be fit into an emotional perspective that allows physicians to endure the losses and disappointments inherent in medical practice over time.

After an intense, but only two day interaction, Dr. L says that he felt attached
to the patient, and that he would miss him. Imagine how much more intense this attachment is, and how much more pressure there is to idealize, when one grows close to a chronically ill patient over many years? We have already mentioned above one risk of such idealization, which is the subsequent devaluation that occurs when the patient disappoints the physician. But an equally big risk is that in order to maintain an idealized relationship with a patient unto the time of her death, the physician simply does not notice the fears, anger, or worry of the patient undergoing heroic therapies. This not noticing guarantees that opportunities to be empathic, efficacious and genuinely respectful of the real individual will be lost.

The affective attitude of facing the pain of another, or of one's own loss unflinchingly, is courage. Aristotle writes "it is for facing what is painful... that men are called 'courageous." It may at first seem that what "facing" pain means for Aristotle can have little relevance to my model of the emotionally courageous physician, since Aristotle's courageous man is not psychologically minded but oriented towards action on the battlefield. And I have in mind a courage that is apparently much more passive, which involves not physical action but an emotional alignment with the suffering of the patient. But there is an essential link between Aristotle's conception of courage in action and my conception of the courage to endure emotions honestly. In both types of courage, one knows that what one faces in the pain and loss of another is the possibility of one's own pain and loss. Aristotle's courageous man has an unflinching sense of his own mortality that
gives the character of courage to his risking his own body to save other bodies. Courage thus contains in it a kind of grief, a sobering realization of one's own finitude. When Dr. L sits down by the bedside of his patient, and experiences trepidation, he is poignantly aware of his own vulnerability and yet committed to taking on the pain of the patient; this double awareness, in which one knows that one is afraid but does not thematize one's fear out of concern for another, is a courageous stance.

In current discussions of medical ethics the issue of whether a good physician needs to be courageous is raised in discussions of risk in caring for patients whose illnesses may physically endanger physicians, most notably, hepatitis B, tuberculosis and HIV infection. Most often courage is seen as a supererogatory virtue that is not necessary for providing good day to day medical care. This ignores the importance of courage in the day to day enduring of the most tragic aspects of human life. Idealization and detachment are strategies for facing these tragedies when courage fails or wears thin, as it inevitably does.

The question of how, practically, to cultivate courage in physicians is a topic for another work, but a few points follow directly from the idea that courage requires a capacity for grief. First of all, one needs time off from the battlefield in order to grieve and repair oneself, and most physicians do not get the time off to grieve and renew themselves in a way that will allow them to face rather than shrink from, each and every patient's suffering. One wonders, for example,
whether another night's rest might have enabled Dr. L to face his patient's discharge with a realistic appreciation of both the patient's fear and of his own fear for the patient's future.

Second, one needs rituals that acknowledge death and renewal in the context of a community. But physicians are extraordinarily emotionally isolated and do not have rituals to grieve. The closest physicians come to breaking their isolation and sharing their painful emotions with each other is through the use of "black humor," which I think is analogous to the use of humor at a wake. Imagine if physicians had communal gatherings to acknowledge the uncanniness of such routine experiences as turning off the respirator of a patient in a coma. Most importantly, physicians often lack any system of faith that might place in perspective their hubristic belief that they are responsible for whether patients live or die, so that deaths, especially unexplained deaths, are personal failures.

In summary, the physician's goals of understanding the patient realistically, treating the patient efficaciously, and respecting the patient as a person require emotional self-regulation. In this chapter, I presented a three tiered model of emotional self-regulation: first, one uses concurrent emotional attitudes to move oneself persuasively out of other emotions; second, curiosity frees one to imagine alternative construals of one's own and the patient's situation, so that one can play a creative role in generating the attitudes that will move one persuasively; third, courage allows one to face the suffering of one's fellow human beings without
having to split-off awareness of one's own vulnerability, thus allowing for maximal awareness of one's emotional attitudes. The "tiers" here are not necessarily consciously experienced as at different levels of reflection. And there is no reason to call curiosity or courage "second-order" emotions, since they have as their focus an aspect of the agent's situation in the world. Rather, these three attitudes work together in apparently simple emotional experiences. When Dr. L sits down beside his patient, despite his trepidation, and asks him to tell in his own words what is wrong with him, he displays all three aspects of emotional self-regulation.

V Prospective Conclusions for Future Work on the Physician-Patient Relationship

How does this model of emotional self-regulation change our reading of Osler's description of the physician striving for "equanimity?" Osler says that the physician's recognition that he has something essential in common with the patient is intolerable, and moves him to forgetfulness. Given our account of emotions as determining salience, it becomes clear that Osler's physician is suffering along with the patient; for to experience something as intolerable, and yet unavoidable, and to desire to flee it, is to suffer. We have come full circle to find that it is not by detaching himself but by resonating in his own being with the patient's suffering that Osler's physician "sees" the "weakness" he has in common with the patient.227

Yet Osler says that the physician's recognition of the pathos of his own being
is resisted by a "happy egotism" that drives him to forgetfulness. This
forgetfulness distinguishes his standpoint from that of Dr. L. Osler's physician
does not become curious about his own or the patient's suffering, or invoke pride
to dampen his pain, or face the patient's pain with courage. Rather, he forgets the
patient's suffering by forgetting its significance for his own being. He strives to
segregate his emotional responses from his conscious thought processes, so that he
is unaware of any feelings, including compassion for the patient.

This kind of splitting of the self, in which the mind does not consider what the
feelings recognize, is commonly experienced by physicians who justify their
experience under the ideal of "detached concern." But to treat patients in a rote
fashion, without consulting one's compassion, makes all kinds of cruelty possible.
This dangerous side of detachment explains why Osler says that the forgetful
physician needs self-control and charity. He is like the Kantian rule-fetishist who
acts out of an external commitment to behave charitably towards the patient, but
is not moved to genuinely respect the patient. If, instead, he could use curiosity
and courage to remain affectively engaged, he would be less likely to treat the
patient as a thing rather than as a person.228

In concluding this dissertation it is notable that once the "happy egotism"
expressed in the ideal of "detached concern" is seen through, physicians face
inevitable conflict. The goals of understanding and respecting patients require
affective connections that are endangered by the structure of current medical practice. This leads to some final questions about the implications of this dissertation for the future of the doctor-patient relationship in a time of massive change in the institutional practice of medicine.

Although analyzing the socioeconomic structure of medical practice is beyond the scope of this work, this dissertation has socioeconomic implications given that it argues for enriching rather than impoverishing the affective ties between physicians and patients. Currently the physician-patient relationship is severely fragmented. Fewer Americans have one primary internist coordinating their care. Physicians are increasingly expected to take on a gate-keeper function in which they have a primary obligation to distributing limited health care goods in an economically feasible way, which competes with their obligation to care for the best interests of patients. Given these two factors, as well as the increasing bureaucratic pressure to spend less time with patients noted in the introduction, physicians are often not sufficiently available and unencumbered to engage with patients in making serious health care decisions.

The idea that physicians are guided by a "detached concern" for their patients provides security to those who do not want to consider the impact of the fragmentation of medical care on patients. The results of this dissertation disturb that security, and provide guidance for reconsidering these issues. First, the issue of whether patients benefit from a primary, continuing relationship with an internist
is clarified by our arguments. We have shown that understanding the patient's illness requires empathy, and that empathy requires building a working model of the patient's world over time and multiple interactions. Thus continuity of care is important not just for comforting patients but for making accurate diagnoses and treating the patient efficaciously.

Second, we have provided an argument against using physicians simultaneously as care providers and gate-keepers by arguing against the picture of Kantian impartiality as involving an impersonal interest in others segregated from one's affective ties. Physicians under direct economic and professional pressures to withhold medical care as institutional gate-keepers may not be able to put the patient's interests at the center of their attention. This would depend somewhat on where the locus of responsibility for limiting care resides. If, as in increasingly numerous health maintenance corporations, the physician's own monthly bonus depends upon not referring the patient to expensive consultants, then it seems that the need to steel oneself from being moved by the suffering of the patient in order to avoid guilt at denying him care would inhibit respect, empathy and curiosity. If, however, the physician is not a directly interested party who benefits from the patient's sacrifices, but rather is rationing care according to guidelines that physicians and patients have previously agreed are ethical, then perhaps the physicians role as an advocate can be preserved. \cite{229} In such cases, the physician would be more like a parent sharing scarce resources as fairly as possible among
many children. However, this parental metaphor, which appeared benign when describing the physician's fiduciary obligations to patients in an expanding health care system, takes on a different tone when used to convey the physician's power to control a scarce resource. A question for future work is what impact on the affective and ethical nature of the physician-patient relationship is posed by rationing medical care?

Third, the question of what role physicians ought to play in patients' health care decisions looks very different once one appreciates the time and mental freedom necessary for empathy, curiosity and respect. In a recent article, Ezekiel and Linda Emanuel proposed that physicians should offer expert moral guidance based on their professional values, regardless of the patient's own values, rather than striving to help patients clarify their own particular health care values. For example, the physician should use arguments to persuade patients to care about the health of others and to therefore volunteer to be research subjects. The Emanuels believe that physicians accrue moral wisdom about dealing with illness as they carry out their other functions, and thus are appropriate moral guides. What is appealing about their view is the idea that there are professional values, such as promoting the well-being of each patient, implicit in medical practice. I argue elsewhere that it is these shared professional values that provide the essential boundaries of the physician's activities. When physicians operate outside of certain guidelines on their behavior -- including the principle of being first and foremost
an advocate of the individual patient -- they lose their moral mandate.

But the Emanuels' idea is that the physician has an obligation to influence the patient's behavior to conform to a set of health values that exist independently of the particular patient's experiences. To some degree, physicians already do this in order to preserve the health of other people besides the patient. For example, physicians are compelled to recommend that people protect their own health and the health of others by practicing "safe" sex. Yet, this extension of the physician's role when patients are likely to endanger others does not entail that it is best for physicians to give moral advice to patients making serious decisions about their own futures, rather than to empathically and respectfully assist the patient in articulating his or her own values.

An important assumption of the Emanuels' article is that physicians in current medical practice can adequately intuit what is of moral significance for patients. But physicians cannot adequately perceive what is morally salient for patients from a detached standpoint. Rather, empathy and respect are necessary to appreciate what is significant for suffering patients. In addition, sensitivity to the moral dimension of patients' situations depends upon a capacity to think reflectively about difficult moral dilemmas over time. I picture moral reflectiveness as a kind of meta-curiosity that loosens the hold of one's own and one's patients' "natural" attitudes. Such curiosity is essential if one is to free oneself sufficiently from primary allegiances to both persons and principles to reconsider their value in light
of new experiences with patients. Finally, by involving oneself empathically in these new experiences with patients the physician can develop moral wisdom that is rooted in the actual practice of medicine, rather than in abstractions. Physicians who do not have the time or support to cultivate their own capacities for empathy, respect, or curiosity are unlikely to accrue moral wisdom. Even if one favors the kind of paternalism supported by the Emanuel's (which I do not), one must question whether physicians under present conditions are suitable moral guides for patients.

In concluding this dissertation it is apparent that the physician's situation in our present society is a tragic one, given that to understand and respect patients requires affective connections that are at best infrequently supported by our current institutions. But the ideal of "detached concern" reifies and excuses the socioeconomic and political limitations of our health care system by idealizing affective disengagement as the truth-seeking, impartial, courageous approach to patients. I have shown that it is none of these, and that physicians can only fulfill their obligations to patients by empathizing with them, and genuinely feeling respect for them. When these attitudes are accompanied by curiosity and courage, both of which require institutional support, physicians can regulate their affective engagement with patients sufficiently to care for patients over a lifetime without losing their compassion.
1. The reflexivity of the term "detach" -- which means detach oneself -- may be misleading if it is taken to require a conscious act of struggling against occurrent emotion. Rather, the goal of detachment is embedded in the very structure of medical interactions, so that physicians need not deliberately try to be unemotional. Rather, they learn to pre-reflectively repress, deny or otherwise divorce themselves from their own emotions.


3. However, the goal of basing one's diagnoses and prognoses on scientific principles does not in and of itself justify the notion that physicians should detach themselves from patients. There is no scientific evidence that emotional engagement disrupts medical care. To the contrary, research on the placebo effect suggests that some emotional engagement between physician and patient may be therapeutic. See Spiro, Howard, Doctors, Patients and Placebos. New Haven: Yale University Press, 1986.

4. In hospitals and HMO's physicians are forced by the institutional emphasis on economic efficiency to take an assembly line approach to patients. Hospitals train physicians to minimize their time with patients by hiring too few residents to work inhumane hours. Physicians are encouraged to be superficially pleasant but not to waste "unproductive" energy by responding emotionally to patients. By the time doctors are out in private practice, they have internalized the value of minimizing time with patients in order to maximize efficiency. Physicians are thus not only instruments but commodities. While they are in other ways a privileged group, they fit Marx's description of alienated labor if one considers how their sympathetic emotions, to paraphrase Marx, are most their own yet most taken away from them in their professionalization. See Marx, Karl, The Economic and Philosophic Manuscripts of 1844. New York: International Publishers, 1964.

5. Adopted by the American Medical Association in 1957 and printed in Beauchamp and Childress, op.cit.

7. American legal cases since the turn of the century have censured physicians for performing surgery against patients' wishes (Schloendorff v. The Society of New York Hospital, 1914). However, the idea that physicians actually need to seek the informed consent of patients before starting invasive treatment is a very new idea, which is explicitly stated for the first time in the 1957 American legal case of Salgo v. Leland Sanford Jr., University Board of Trustees (1957). See Jay Katz, The Silent World of Doctor and Patient, New York: The Free Press, 1984.


11. Entralgo, p.46, op. cit.


17. See Baron, R.J., "Bridging Clinical Distance, an Empathic Rediscovery of the Known," Journal of Medicine and Philosophy, vol.6, 1982, pp. 5-23.

18. There is a striking similarity between Osler's talk of learning to control one's disruptive emotions emanating from one's "medullary center" and Descartes talk about controlling the disruptive emotions emanating from the pineal gland in The Passions of the Soul. See Descartes, The Philosophical Works, E. Haldane and G.R.T. Ross, trans., New York: Cambridge University Press, 1984, pp 336-356.
However, it is an additional assumption to link the obligation to overcome prejudice with the goal of emotional detachment. It is not clear that prejudices, for example, towards cancer or AIDS patients, are essentially related to the emotionality of the physician's response to the patient. On the one hand, one may be both tolerant and emotionally involved with patients. On the other hand, one may have rigid beliefs and habitual avoidance patterns that restrict one's care for certain types of patients regardless of one's occurrent emotion. Hence, it is not clear how detaching themselves would enable physicians to impartially value those persons they are otherwise prejudiced against.

This ideal is not to be confused with the ideal that health care ought to be distributed to all in an equitable manner. Kantian impartiality does not address issues of distributive justice, even though strong arguments can be made for fairness of this sort using Kantian ideas. Rather, the common thought behind the Kantian ideal of valuing others as persons and the medical ideal of respecting all patients as persons, is the idea of responding to others out of a sense of duty regardless of one's personal inclinations.


The concepts of empathy and of clinical listening are not identical despite their overlap in clinical discussions. First, as Jackson shows, the idea of listening in medicine is older, and much broader than the idea of "empathic listening." Second, as Jackson also points out, the term "listening" in "empathic listening" can be used metaphorically to refer to other forms of receiving information besides hearing. See Jackson, op.cit.

In addition, recent work on the placebo effect offers scientific evidence that a good physician-patient relationship plays an independent causal role in healing. This research sets out to measure through non-mechanical models, what for Descartes was in principle not measurable. See Doctors, Patients and Placebos, Howard Spiro, New Haven: Yale University Press, 1986.


26. One problem with contrasting the intentionality of emotions with the apparent non-intentionality of such bodily feelings as itches and twinges, is that even bodily feelings may have an essential intentional component. That is, the picture of bodily 'feelings' as naked qualia, presupposes the possibility of a field of experiences in which body sensations are presented first in terms of their general quality, and then localized to one's own body in order to be hooked in with one's capacity for agency. But this model poorly suits both the foundational body sense, proprioception, and sensation. Consider first proprioception. According to this picture of feelings as qualia, I would first have a general arm sensation, and then contingently refer this sensation to my left arm, here, bent at an angle. But I never have such a sensation. And more importantly, the kinesthetic information I receive about my left arm here leaning on the table, is not something I experience in isolation, but something that is integrated into my continuous activity of moving and settling my body in space. So even proprioception cannot provide an example of a pure "feeling" in the sense of pure qualia disconnected from experience.

Heidegger takes this point much further, arguing against the possibility of pure "sensations" like seeing red, prior to perception of objects of interest to us, like the red car over there. Heidegger claims that we do not first see colors and hear sounds; rather, from the start we see things that interest us. Further, for Heidegger our original attitudes towards the world are essentially affective, a claim which will be considered in chapter two on the rationality of emotion. The point here is that the contrasting notion of non-intentional "feels" used to highlight the intentionality of emotions may not itself have a phenomenological basis. See Heidegger, Martin, *Being and Time*, New York: Harper and Row, 1962.


30. op. cit., *Passions of the Soul*.

31. This is not to overlook the fact that there are cases in which shifting emotion is secondary to attaining new information. For example, if I am angry at someone for committing an injustice, and I learn that she did no such thing, my anger may go away. So emotions can shift because of changes in belief.

32. I am not using the term "belief" here to include "faith." "Faith" can have inertia, and may be more like the affect hope than like purely cognitive beliefs.


35. ibid., pp.86-87.

36. ibid., p.106.

37. Karsten Harries, personal communication.


39. de Sousa, op. cit., chapter eight.


41. Freud's emphasis on "the pleasure principle" as the primary aim of instinctual life is resonant with Sartre's view of emotions as primarily oriented towards gratifying the self by transforming the self rather than the external world. See for example, Freud, Sigmund, "The External World," in *An Outline of Psychoanalysis*, translated and edited by James Strachey, New York: Norton, 1949. However, instincts and emotions are not equivalent for Freud; logically, Freud rejected the idea of an unconscious affect. And of course Freud is quite interested in the way instincts can be transformed into emotions that serve the reality seeking goals of the ego, and the ends of the superego (consider for example, unconscious guilt).


43. de Sousa, op. cit., pp. 181-182.

44. ibid.,182.

45. ibid.,p.184.

46. Greenspan, op.cit., p.4.

47. de Sousa, op.cit., p.24.


49. ibid., p.336.
50. ibid., p.338.

51. ibid., pp.359, 370, 373.


53. In a separate paper I argue that reliable bodily agency presupposes familiarity with one's bodily history rather than introspective access to one's own body as a spatial entity.

54. Now if the concept that an ongoing physical influence causes an otherwise rational agent to behave strangely seems far-fetched, consider how this idea is very much part of our present view of mental illness. The first mental disturbance to be explained by the early 'science' of psychiatry was neurosyphilis, which involves exactly this picture of a pathogen, a foreign physical agent, continuously determining how one perceives and thinks. The idea is that the germ not only makes the agent so sick that she can not think (ie. puts the agent into a delirium) but also exerts a seamless ongoing influence on her thought processes. Currently psychiatrists are trying to find the pathogen responsible not only for states of delirium that are clearly associated with infections of the brain, but for schizophrenia and manic-depressive illness. In these cases the 'pathogen' would arise within the agent's own body as a genetic mutation or chemical imbalance that occasioned an ongoing emotional lability, disrupting the agent's reasoning.


56. Recall my argument against Joseph Fell's two-event model of emotion in chapter one.

57. This is the strategy used by Donald Davidson to explain how beliefs and desires can genuinely cause actions. See "Mental Events," in *Essays on Actions and Events*, New York: Oxford University Press, 1980.


63. Kenny, ibid.

64. Descartes picture of emotions as arational is integrally related to his divorce of representational from adaptive rationality. The former notion of rationality pertains to representing things as they really are, and hence is sometimes referred to as backward-looking in that it depends upon proper adherence or appropriateness to a given (hence prior) state of affairs; the latter notion pertains to usefulness or success in negotiating in the world, and hence is referred to as forward-looking in that it depends upon the influence that one's mental activity will have on one's future condition. The distinction between these two senses of rationality is brought out in the following two examples by Patricia Greenspan: consider how rage may be an appropriate response to an actual injustice, and yet maladaptive for the agent, who brings on further harm by the response of others to her rage. Contrast this with the case of inappropriate but adaptive blame, in which one blames another for misfortunes brought on by oneself, and is thus able to become therapeutically angry rather than depressed. See Greenspan, Patricia, Emotions and Reasons, op.cit.

Descartes grants emotions a degraded form of rationality that derives from their adaptiveness for human beings. His teleological explanations for the occurrence of emotion types suggest that emotions play a needed instrumental role in human life. However, as Descartes makes clear this role is necessary for human life precisely because human beings are not pure rational agents, but rather intermingle res extensa and res cogitans.


66. ibid., p. 193.

67. ibid., p. 193.

68. ibid., p.191.
69. ibid., p.194.
70. ibid., p.196.
72. ibid., p.173.
73. ibid., p.175.
74. ibid., p.173.
75. ibid., p.175.
76. ibid., section 29.
77. ibid., p.175.
78. ibid., p.176.
82. Heidegger's picture of mood as styles of being-in-the-world conveys the importance of emotions for both expressing and sensing what matters to human beings. This link between the expressivity of emotion and the sensing of emotion is what Heidegger adds to our understanding of how emotions determine salience.
83. I take it that this divorce is especially unsuitable for our empathic knowledge of other human beings. Consider for example, Patricia Greenspan's distinction between adaptive and representational rationality in her discussion of "identificatory emotions," emotions that reveal to us aspects of the other's emotional life. (Greenspan's use of "identificatory emotions" corresponds to what I discuss under the topic of empathy in chapter four). She argues that the appropriateness of such emotions depends upon their general adaptiveness -- they
are essential for conducting ourselves in the social world. But it is apparent that the division of rationality into forward-looking and backward-looking is strained in this case. Identificatory emotions are adaptive because they are essential for representing other persons. (I argue in the clinical empathy chapter that the feelings/attitudes of another person are not available objects for detached cognition). The essential role of emotions in representing the emotional life of other persons shows that emotions are rationally required in the full-blooded sense of rationality. See Greenspan, *Emotions and Reasons*, op.cit.


87. Physicians generally speak of respecting the patient as a "person" as the opposite of treating the patient merely as a mechanism or a stereotype. See Alasdair MacIntyre's discussion of the ideal of treating the patient "as a person" in *Changing Values in Medical Practice*, eds. Eric Cassel and Mark Siegler, New York: University Publications of America, 1979.

88. There are two different senses in which the Kantian moral agent can be considered "impersonal". At this early point I use the term "impersonality" to refer to the agent's psychological condition: the "impersonal" agent is disengaged from the ordinary affective ties that are constitutive of personal relationships. The GMM398 example of the man with no sympathetic inclinations, who helps others out of a detached sense of duty embodies "impersonality" in this first sense. Later on, I argue that "impersonality" characterizes the ontological condition of the Kantian pure practical reasoner.

89. For a critical review of the equation of respect for patients with respect for the "rights" of patients, see John Ladd, "Legalism and Medical Ethics," *Journal of Medicine and Philosophy*, vol.4, March 1979, pp. 70-80.


Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.
91. O'Neil also defends the Kantian claim that, despite their differences in perspective, the FUL and the FEI are essentially equivalent; O'Neil seems to accept the Kantian assumption that the kind of reflection that is constitutive of making universal law is the same kind of reflection that is constitutive of taking-up the ends of other persons. See "Universal Laws and Ends in Themselves," The Monist, Henry Allison, ed., vol. 72, no.3, July 1989, p.341-362.


93. See for example, passage 74, [252] where Kant defines "emotion" as the absence of composure of mind, such that reflection is "impossible"; Anthropology from a Pragmatic Point of View, trans. by Dowdell, Illinois: Southern Illinois University Press, 1978.

94. Robin May Schott, Cognition and Eros: A Critique of the Kantian Paradigm, Dissertation for the Department of Philosophy, Yale University, May, 1983.

95. At the end of GMM 400, Kant provides the following footnote:"Yet although 'reverence' is a feeling, it is not a feeling received through outside influence, but one self-produced by a rational concept, and therefore specifically distinct from feelings of the first kind, all of which can be reduced to inclination or fear. What I recognize immediately as law for me, I recognize with reverence, which means merely consciousness of the subordination of my will to a law without the mediation of external influences on my senses. Immediate determination of the will by the law and consciousness of this determination is called 'reverence,' so that reverence is regarded as the effect of the law on the subject and not the cause of the law. Reverence is properly awareness of a value which demolishes my self-love. Hence, there is something which is regarded neither as an object of inclination, nor as an object of fear, though it has at the same time some analogy with both. The object of reverence is the law alone—that law which we impose on ourselves but yet as necessary in itself. Considered as a law, we are subject to it without any consultation of self-love; considered as self-imposed it is a consequence of our will. In the first respect it is analogous to fear, in the second to inclination. All reverence for a person is properly only reverence for the law (of honesty and so on) of which the person gives us an example. Because we regard the development of our talents as a duty, we see too in a man of talent a sort of example of the law (the law of becoming like him by practice), and this is what constitutes our reverence for him. All moral interest, so-called, consists
solely in reverence for the law."


99. Kant attributes to "sensibility" all of those "ideas with respect to which the mind is passive, and by which the subject is therefore affected." See Anthropology, op.cit., sec.7.


103. ibid., p.371.

104. ibid., p.373.

105. ibid., p.373.
106. ibid., p.372.

107. Herman actually states only the negative counterfactual: one would refrain from acting if the act were impermissible (duty as a secondary motive); but her description of the case in which duty becomes a primary motive is a description of the case in which one actually acts against all ordinary motives, from the motive of duty alone. See Barbara Herman, "On the Value of Acting From the Motive of Duty," The Philosophical Review, XC, No.3, July 1981, p.372.


108. Herman, ibid.

109. ibid., p.374.

110. ibid., p.375.

111. ibid., p.382.

112. "Reflective Deliberation about the truth indeed brings in a standpoint that is impartial and seeks harmony, but this is because it seeks truth, not because it is reflective deliberation, and those features will not be shared by deliberation about what to do simply because it too is reflective. The I that stands back in rational reflection from my desires is still the I that has those desires and will, empirically and concretely act; and it is not, simply by standing back in reflection, converted into a being whose fundamental interest lies in the harmony of all interests. It cannot, just by taking this step, acquire the motivation of justice." Bernard Williams, Ethics and the Limits of Philosophy, Massachusetts: Harvard University Press, 1985, p. 69.


115. ibid., p.220.
116. ibid., p.209.

117. ibid., p.220.


   Lipps' statement that one could "fully comprehend" another's experience has behind it both the idea that empathy allows for a quasi first-personal understanding of another and the idea that this understanding could in principle approach certainty. For a discussion of Lipps' view of empathy as involving a complete merging with another, see Edith Stein, *On the Problem of Empathy*, The Hague: Nijhof, 1964, p.16.

   Schutz (and Dilthey before him) extended the idea that one could strive for certainty in one's knowledge of other persons to the idea that empathy could be the basis of a distinct form of objective knowledge appropriate for the human sciences. See Schutz, Alfred, "Intersubjective Understanding," in *The Phenomenology of the Social World*, Walsh and Lehnert, trans., Chicago: Northwestern University Press, 1967. See Berger, op.cit., pp 62-63 for Dilthey on empathy. In contrast, physicians have often used the term "empathy" to describe a non-scientific understanding of human experience. Most of the authors I will consider here sit between these points of view. They are psychoanalysts who have taken empathy to be an important source of knowledge about patients regardless
of its scientific status. Until the 1960's, much of the writing on empathy in medical journals was by psychiatrists. However, with the rise of concern about the adequacy of physician-patient relationships there is now increasing attention by other specialists, especially internists and pediatricians, to the role of empathy in medical practice.


123. See Berger, Buie, Basch, op.cit.


125. See Buie, op.cit.

126. By defining empathy as a capacity, I do not mean to imply that it is reducible to a way of behaving. I take the above detailed analysis of the intrapsychic dimensions of empathy to be sufficient grounds to reject a behaviorist claim that empathy just is the capacity to say and do things that make another person think that they are being understood. Rather, I have given an explanation of how such behavior is possible in the first place.


129. Aring, ibid., p. 449. Given Aring's definition of "sympathy" as including any similar emotion occasioned by another's emotion, sympathy is not equivalent to pity. Rather, one can feel sympathetic joy, anger, etc. The core idea here is that one can resonate with another's emotional state. What such emotional resonance involves is considered later in this chapter.
130. Aring, ibid., p.450.


132. Melvin Lansky focused my attention on the distinction between taking on another's emotional conflicts, and remaining emotionally engaged without such reactivity. Personal communication.

133. In chapter five much more will be said about using curiosity to redirect one's emotional reactions, and also about how one's empathy and curiosity can work together to sustain an appropriate level of emotional participation in another's experience.

134. See for example, Stoller, R.J., *Observing the Erotic Imagination*. New Haven: Yale University Press, 1985. But even if one disregards psychoanalytic theory, an appreciation of the focus in novels, plays and films on adverse human experiences suggests that such experiences can be in some sense vicariously enjoyed. Convoluted attempts to say that the audience does not resonate with the suffering of the characters, but only with their joy at the cessation of adversity, seem to ignore the proportion of time spent focusing on human hardship. And theories that say that the audience enjoys comparing their own safe situations as spectators with their vicarious experience of the suffering of the characters already presuppose the point I am trying to make, which is that there can be multiple perspectives, including resonant suffering and an awareness that one is not genuinely in the adverse situation.


136. For this definition of "projection," see *Webster's New World Dictionary of the American Language*, Guralnik, ed., New York: The World Publishing Company, 1972. Osler's idea that the physician must strive for "imperturbability," or freedom from all bodily sensations in order to attain a truthful understanding of the patient's situation, is rooted in the Cartesian idea that emotions are essentially like perceptual illusions that deceive the knower about what is really there in the world. For Osler, both "introspection" and "insight" strive for clear and distinct perception of affective images, in the way that observing the body requires clear and distinct perception of physical signs. Hence Osler's emphasis on the need for the physician to purify himself from his own emotions, so that the mental medium constituting the projection, like the light from a "projector," would have no interference.
Aring and Blumgart share Osler's concern that the physician's empathic understanding of the patient be free from the interference of present emotions. They speak metaphorically of the physician's sympathetic feelings as tying him down to a particular way of seeing the patient's situation, and hence restricting the maximal freedom of movement that is needed for objectivity. Blumgart warns the physician to strive for detachment in order to keep from losing "objectivity and perspective." But what exactly can the detached physician learn about the patient from the detached standpoint? According to Rosenberg and Towers, Blumgart's ideal physician does not relate to the patient "as one person to another, but rather collects 'data' from the patient's behavioral expressions for analysis and responses within the physician's own perspective. The physician is satisfied merely to observe the signs of the patient's illness, rather than to comprehend its experiential content." See Aring, op.cit., p.449, and Blumgart, op.cit., p.451. See Rosenberg, J. and Towers, B., "The Practice of Empathy as a Prerequisite for Informed Consent," Theoretical Medicine, Thomasma, ed., vol.7, no.2, June 1986.

137. This is not to suggest that nothing objective is expressed in the word and gestures of the patient. Rather, the patient's gestures can express sadness, fear, etc. in a way that any normal "observer" could recognize, hence with a certain "objectivity." Caroll Izard gives evidence of such recognition in Human Emotions, New York: Plenum Press, 1977.

The presence of recognizable features of emotion is a necessary condition for the possibility of meaningful discourse about emotion. If we could not recognize and identify the typical expressions of emotion, there would be no criteria for differentiating sadness from anger, or even from any other moment in the "stream" of consciousness. Furthermore, in our physical world identification requires a mark that can be perceived via the senses. But to recognize that another's gestures show anger or fear is hardly a sufficient condition for understanding what her anger or fear is like. The question at hand is thus: what in addition to recognizable marks of emotion, contributes to a qualitative grasp of another person's feelings?


139. This quote of Cassel's is in Rosenberg and Towers, op.cit.

140. Webster's Dictionary, op. cit.

141. The kind of emotional neutrality that Freud describes for psychoanalysis is not equivalent to detachment in the Oslerian sense. Freud notes that archaic, emotional resonance plays a role in analytic listening. And his notion of psychoanalytic abstinence still allows the analyst to have an emotional interest in the well-being
of the patient. But he urges the analyst to avoid using the clinical encounter to gratify unconscious wishes of his own or of the patients. Also, the analyst is to refrain from taking sides in the analysand's intrapsychic conflicts, since the analyst's involvement in such conflicts would inhibit his intellectual freedom and non-judgmentalness (see my earlier discussion of non-reactivity, and my discussion of curiosity in chapter five).

However, there has been a strong anti-affectivity strand in psychoanalysis. The orthodox view has been that the physician's emotions are "countertransference" phenomena that must be attended to because otherwise they act as resistances that impede the analytic process. Such feelings are informative only to the degree they can be translated into conceptual models that can be used to make inferences about the patient's condition. The underlying presupposition is that the object of the physician's insight is the patient's psychic condition as it is, ultimately for an aperspectival knower. Hence the goal of clinical empathy is to arrive at information that is purified of the physician's experiential responses to the patient. The physician strives to be aware of countertransference phenomena in order to correct for the distortions such phenomena introduce into her understanding of the patient. See Dorpat, T.L., "On Neutrality," International Journal of Psychoanalytic Psychotherapy, 6: 39-64.


143. Stein, op.cit.

144. The idea that empathy involves making inferences about the patient is related to the definition of empathy as "projection." "Projection" is also defined as "prediction...based on known data or observations; extrapolation." Webster's Dictionary, op. cit.

145. There are several reasons for rejecting the picture of empathy as detached inference-making. First of all, there is no phenomenal correlate of this putative act of inference: in empathic understanding one does not independently listen to the patient's story and then compare this information to a conceptual model of typical emotional attitudes in order to make predictions about the patient's inner life. Secondly, there is no explanatory power in positing such a mental act. What would be the criteria one would use to apply the general conceptual model of emotions to the case of the particular patient? How would one know that these features of one's mental model of emotional attitudes were relevant to this aspect of the patient's situation? Wittgenstein's criticism of the Cartesian view of understanding as making 'inner' comparisons is relevant here. The idea that the physician could see the patient's expressions of emotion as similar to some general
feature of emotional experience presupposes that she already knows how to recognize this feature of emotional experience. To equate such recognition with detached inference is to attribute to the physician an 'inner' mental map pairing off particular expressions of emotion with general features of emotional experience. But how could the physician know which aspect of the patient's emotion should be paired off with which general concept? What would direct her towards a quasi first-personal grasp of the dramatic quality of the patient's experience versus third-personal observation of the patient's gestures and facial expression? The physician would need a third diagram or image to make these comparisons. But even this diagram could be understood in a variety of different ways. So another diagram would be needed, and so on, into an infinite regress. The idea of an 'inner' act of inference does not explain how the empathizer is able to recognize and appreciate the emotions of other people. See Wittgenstein, L., *Philosophical Investigations*. New York: MacMillan, 1958.

It might be argued that to equate empathy with insight based on inference does not entail making the Cartesian error of positing a discrete act of "inner" inference-making that is accessible on reflection. Even so the "inference" model of empathy still confuses what the telos, or goal of empathic understanding is. As I argue in the following pages, empathic understanding does not aim for prediction in the way scientific inference aims for prediction.


147. I hypothesize that the fact that empathy involves multiple modes of cognition, including affective imagery, and reflective thought, contributes to the "reality" of the object of empathy. This thought is based on the common observation that using multiple senses contributes to the reality of an object of perception.

148. After this chapter was already written, Stanley Jackson published "The Listening Healer and the History of Psychological Healing," op. cit. In addition to providing a thorough historical analysis of the concept of clinical listening, this article includes important references on empathy, including an early, interesting psychoanalytic work on empathy that is rarely referenced today: Schroeder, Theodore, "The Psychoanalytic Method of Observation," *International Journal of Psycho-Analysis*, London, 1925, vol.vi, pp. 155-170. Schroeder's describes empathy as a form of "inductive introspection" in which one puts "one's own consciousness at the disposal of the unconscious determinants of another's personality" (p.162). He also argues that "psycho-analytic theory does not in the least depend upon logical inference based upon the surface introspection of a relatively static psyche." This point interestingly anticipates my own views, but
Schroeder does not go on to provide philosophical arguments showing the inadequacy of inferential reasoning for empathy (he instead uses the psychoanalytic concepts of projection and introjection to explain how empathy is possible). Yet Schroeder also emphasizes the importance of the psychoanalyst's own "conscious withdrawal of interest" for empathy to be possible (pp.159,162).

149. Lipps, in Hunsdahl, op. cit., p.182.

150. ibid., p.184.

151. See Heidegger, M. Being and Time, op.cit. Heidegger offers a model for the kind of experiential "projection" of meaning that is essential for empathy. He argues that we do not first perceive things as meaningless sense data, and then infer that they are objects that interest us, like tools, forests and persons. Rather, the fact that as perceivers we attend to certain phenomena rather than others, in meaningful, practically useful patterns, reveals our capacity to "understand" the world pre-reflectively. Heidegger says, referring to ordinary things we encounter, that:

The ready-to-hand is always understood in terms of a totality of involvements. This totality need not be grasped explicitly by a thematic interpretation. Even if it has undergone such an interpretation, it recedes into an understanding which does not stand out from the background. And this is the very mode in which it is the essential foundation for everyday circumspective interpretation. In every case this interpretation is grounded in something we have in advance -- in a fore-having [G150].

Something very much like this can also be said of our understanding of persons. There are two aspects of Heidegger's picture of how meaning is "projected" in the human world that I want to clarify. First, the term "totality" does not indicate a collection of unrelated things. Heidegger is referring to a web of mutually implicated things, like the totality of a painting, a workshop, or the world-as-lived [G 102-105] by an individual. Heidegger's claim that the ready-to-hand is always understood in terms of a totality of involvements is the claim that understanding always takes place in a context of interests and projects. We will return to this basic presupposition later on.

A second major point of this passage is that the cares and interests that provide a context for understanding something must be partially pre-reflective, unthemmatized. For Heidegger the experience of being interested in something, which he calls a "fore-having," is necessary for recognizing a gap in one's grasp of the thing and hence for questioning and finally conceptualizing it [G 150-152]. The idea that our overall comportment towards things and other persons underlies the possibility of understanding them, lies behind Heidegger's claim that "moods" disclose reality.
152. Sartre's views all emotions as aiming to sustain one's self-esteem. Thus he would have to have something like this narrow conception of the interests of the empathizer. See Sartre, *The Emotions: Outline of a Theory*, op.cit.

153. The concept of imagining has at least as complex and confused a history as the concept of empathy. Casey traces the term "imagine" to the Latin root, "imaginari," which means "to copy." This root coheres with the Cartesian and later, empiricist theses that mental images are only more or less accurate copies of sensations. But Casey also sees in our current concept of imagining the earlier Greek idea of "phantasia" "which includes any kind of mental seeing or impression in the soul" (Zeno). See Casey, Edward, *Imagining*, Bloomington: Indiana University Press, 1976.


158. Kohut, op.cit.

159. Deutsch, op.cit.

160. Fenichel, op.cit.

161. Fliess, op.cit.

162. Kohut, op.cit.
It might be argued that Stein's criticisms of Lipps are not relevant to the merging model of empathy because the physician is not like an audience to the patient's suffering, but rather is an actual participant in the patient's trials. Some physicians take themselves to share the patient's struggles against disease to such a degree that they think they actually feel the same way the patient feels. For example, some doctors speak of fighting off cancer together with their patients as if it was pretty much the same battle to give chemotherapy as to receive it. But it seems highly unlikely that the physician and the patient could genuinely feel the same way about the patient's pain, suffering, death. It is just not the same thing to recognize that a patient one cares about is dying as it is to recognize one's own imminent death. And it is even more unlikely that the physician could take each and every one of her patient's experiences on in this manner. But even if it were possible for the physician and the patient to feel the same way about the patient's illness, this would not explain how the physician could grasp the patient's feelings via her own feelings.

Stein addresses this point by agreeing with Lipps that it is possible to have an experience of oneness with another person. But she argues that such an experience cannot be the basis of empathy, since it actually presupposes empathy. She points
to the case in which two persons actually share a situation, and respond to the
same news with the same feeling, and recognize each other's feelings, hence
coming to share a response to the same object. For example, two soldiers in the
front lines, hear that the war is won. Both feel joy and relief, and recognize each
other's joy and relief. Stein says in this case the barriers separating one 'I' from
another have not entirely broken down.

"I feel my joy while I empathically grasp the others' and see it as the same.
And seeing this, it seems that the non-primordial character of the foreign joy
has vanished. Indeed, this phantom joy coincides in every respect with my
real live joy, and their's is just as live to them as mine is to me. Now I
intuitively have before me what they feel. It comes to life in my feeling,
and from the "I" and the "you" arises the "we" as a subject of a higher
grade."57

Stein shows here that even when two persons are actually in the same boat, their
experience of merging already presupposes a prior act of empathy, and thus cannot
explain how empathy is possible.

See Stein, op.cit., p.16.

Nijhoff, 1977. Sartre is influenced by Husserl's idea that one can intuitively
experience the absence of the other. However, Sartre takes the more extreme
position that awareness of another creates not only a sense of absence but a
dynamic negation within one's own sense of being. See "The Look," in Being and

177. Kohut, op.cit.


1984, p.132.

180. ibid.

181. Basch, op.cit. Note that Basch takes resonance emotions at birth to be examples
of "fixed action patterns," which are reflex arcs that involve cortical functioning.


183. Basch's theory is Aristotelian in that it emphasizes how natural dispositions can be
educated into complex attitudes that comprise practical knowledge.

185. Buie, op.cit.


188. ibid.

189. Casey, op.cit.

190. ibid., pp.44-45.

191. Casey describes the imagined agent as either the imaginer himself or a proxy. His reason for leaving this open is that he wants to give an account of all imagining-how, including fantasies that specifically pertain to oneself -- imagining how I will feel when I finish this chapter, as well as fantasies that are not necessarily self-referential -- imagining how it feels to dive into a cool swimming pool.

192. Casey, ibid., p.45 footnote 11: "By the terms 'action,' 'activity,' 'agent,' etc. I do not mean to imply that the content of imagining-how is always a form of action in which the subject ... takes the initiative. This is often so in imagining-how, but there are also cases of imagining-how to suffer, to be imposed upon by others, etc. Thus the central notion of 'personal agency' includes a broad spectrum of ways in which the imaginer becomes implicated via self-projection or by proxy in his own imaginative presentation, and some of these ways include adopting a passive stance within the state of affairs contained in the presentation."

193. Casey, ibid., p.45.

194. ibid., p.45.


196. ibid.

197. This does not contradict the point made earlier by Stein that such imagining can only announce the presence of another's ownness, rather than present it fully as here and now. Rather, I take it that Wollheim's point is not that the imagined

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.
experiences will be as complete as actual experiences, but rather that they will be consistently unified according to the perspective of an experiencing subject rather than the perspective of an observer.

198. Howard Spiro argues that attending physicians can "teach" empathy, by including narrative accounts of patients' lives in clinical rounds. See "What is Empathy and Can It Be Taught?" *Annals of Internal Medicine*, vol. 116, no. 10, 1992.

199. Casey, op.cit.

200. Wollheim, op.cit. and Casey, op.cit.

201. There would seem to be an additional meaning to the physician's idea that via empathy she is in the patient's situation. I take this additional meaning to be what Wollheim calls a "master thought": "the way in which we conceive or in which we represent to ourselves our mental processes, or the conception under which a mental process occurs." According to the merging model, the physician takes her imaginative portrayal of the patient to project her into the patient's situation. It is this representation of empathic imagining as a merging experience that is mistaken. Rather, in imagining how it would feel to be in the patient's situation the physician mixes her own affective repertoire with the patients in order to fill-out her portrayal of the patient's experience. Recall Edith Stein's (and Husserl's) point that the guiding intention of empathy is to grasp the situation of another person. This goal, with its explicit awareness of another's separateness, is the genuine "master thought" of empathy. See Wollheim, op.cit., p.192, and Stein, op. cit.

202. To make this traditional conception of the interaction of thought and feeling more explicit, I elaborate on Wollheim's (op.cit.) description of the features of "imagining how" as the "inner" creation and performance of a play for an audience. The mistaken picture of empathic understanding is as follows: one's cognitive faculty writes and performs a play for an audience of two, including a translator who watches and translates the images into affective signals, and an affective audience who responds mechanically to the signals. Given this picture, the possibility for appreciating new aspects of the drama cannot be explained. If the audience somehow gives emotional responses that translate into images that are appropriate for the next part of the drama, this would have to be a matter of sheer luck or based on some magical, extrasensory capacity of the translator.

This model of the interaction of mind and feeling in "imagining how" another feels should be familiar to the reader by now. It brings back into play something very similar to the view encapsulated in the "insight" model of empathy: understanding another's feelings involves "projecting" mental images that convey the person's experience, then using these images to make inferences about another's
feelings. Except in this case, it is not strictly introspection and detached inference that are invoked, but the ability to "observe" an imaginative portrayal of another person's experience, respond to it emotionally, and "apply" one's responses to further imaginative portrayals.

203. See Freud, "The Interpretation of Dreams," S.E., op.cit., vol. IV and V. I read this after writing an initial draft of this chapter. In addition to the structural similarity between primary process thinking and imagining how another feels, there are other concrete similarities. For example, Freud points out on p.534 that in dreams there is an elision of the "perhaps" that characterizes speculative thought in waking life: If one is worried that someone is, perhaps, angry at one, one does not dream that she might be angry, but that she is angry.

The elision of the "perhaps" in dreaming corresponds to the elision of the hypothetical in empathy; empathy involves experiencing the announcement of another's feelings, rather than positing that, hypothetically, the other might feel a certain way.

204. At the beginning of this analysis of clinical empathy, I argued that concepts must guide empathic understanding, rather than just contribute to the retrospective description of some independent pre-conceptual act of empathy. I pointed out that physicians often derive much of the content of their initial grasp of the patient's world from their conceptual understanding of what certain illnesses and certain types of persons are like. I now add the point that these concepts influence the selection of the content of the physician's imaginative experience. For example, if one knows that, as a matter of fact, schizophrenics suffer from feelings of isolation, then one can use this information to pick-up salient features of a new schizophrenic patient's history. For example, one will attend to any indication the patient gives of feeling removed from others at school.

But the richness of empathic understanding depends not upon pre-conceived notions but upon spontaneous communication between physician and patient. The two must work together to co-narrate their imaginative production. Sometimes this will involve a dance that is like a waltz in which the physician follows the patient; other times it is more like dancing to rock and roll in which innovation and even completely changing directions is essential for adequate empathy. For example, the physician might initially imagine that an elderly woman patient, whose husband has died in the past year, after forty years of marriage, is feeling lonely and sad. She might then ask the patient if she thinks about her husband much of the time. If the patient than says that her marriage was not all that happy and she has more time now for her service club, the physician must be able to change gears. The physician may find herself imagining new opportunities, and independence, and beginning to feel the excitement and fear that going it alone involves for this
patient.

Just as there is no reified "model" of the patient's inner life, it is incorrect to picture the "feedback" from the patient as leading to the editing of some already created entity, which is cut and pasted to be made more accurate. Rather, the patient's input must influence the physician's capacity to imagine how the patient feels. The patient's input leads to a greater attunement on the part of the physician to the particular affective imagery of the patient. One concrete manifestation of this process of attunement is that, over the course of time, the empathic physician will begin to use phrasing and gestures that closer to the style of the patient. This seems to be the behavioral correlate of the tuning in process that allows the physician to utilize images that are closer and closer to the patient's own. The physician's capacity to follow the patient's story as one follows a drama, which was mentioned earlier as an important factor for being a good listener, can now be explained. In order to follow the patient, the physician needs to be able not only to imagine how the patient feels in any particular situation, but to modify and direct her imaginings in response to what the patient communicates. We have seen that sometimes one strives to imagine new features of the patient's situation, and at other times the images and affects come involuntarily, and even surprise the empathizer. An example of this follows in the final chapter.


206. Although the point that physicians are socialized into their roles argues for the directedness of the physician's emotions, it does not explain how it is that, like Dr. L, an individual physician can grapple with conflicting emotions towards patients, only some of which are appropriate to his role. The issue we face here is understanding how one deliberately sustains one's role-related responses over time? We have already argued that a detached choice to "act like a doctor" could not account for the flow of emotions in response to patients that is necessary for the doctor to fulfill her role. And physicians do not consciously remind themselves over and over -- I am a doctor, not a lover or parent, so its best for me to be moved by the patient's courage but not his animal attractiveness.

One does not become a doctor the way one becomes a real estate broker, but rather the way one becomes a family member in one's family of origin. One begins as a vulnerable medical student and then intern who is totally dependent on senior physicians to guide one, who is deprived of sleep and family, threatened with incompetency that could cost other people's lives and harm one's own. One practically lives in the hospital for a period of years during third and fourth year of medical school and internship and residency (surgical residencies now take on

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.
average eight years, medicine residencies and fellowships six years, psychiatry, four to six years, etc.). This period of medical "training" does train one's very responses to reality, rather than just offering one a fund of knowledge. This is perhaps most obvious when one considers the way physicians see naked bodies and body excretions differently than other people do.

It might be argued that my comparison of the role of doctor with the role of parent or lover is still inaccurate, because even though doctors are raised to be doctors, the role is still artificial, whereas being a parent or lover is doing what comes naturally. I challenge this artificial/natural distinction. I am influenced here by the feminist analyses of gender that show that not only roles like mother/father and wife/husband but even such apparently natural phenomena as walking and talking like a woman or a man are constituted socially through recognition and reinforcement of certain attitudes rather then others. And the social construction of gender never requires a conscious or deliberate taking on of masculine or feminine roles. However, it is possible to become aware that one's most "natural" responses are in fact directed by socially constituted roles. It is possible to consider other ways of responding and to find that one's potential range is larger than one's role dictates. This is what "consciousness raising" is all about.


201. Roberts, op.cit., p.199.

202. The actual behavior of sitting down next to the patient may have had direct emotional impact on Dr. L. A familiar experience to most people is the "whistle a happy tune" phenomenon in which we can shift our own moods by changing our behavior to something that is more conducive to another mood. Although I have argued earlier against Sartre's view of emotion as generally involving an activity of moving oneself through a kind of magical enactment, I agree with his observation that we can set the stage for certain moods in ourselves, just as we set the stage for moods in others. See Sartre, The Emotions, op.cit.

203. Of course, emotions can simply remit without the agent in any way deliberately redirecting herself, but this does not help explain how an agent can direct her own emotions.


206. Davidson, op.cit., p.298.

207. However, I do think that what is at stake in Davidson's thesis is the core concept of Freud's idea of a dynamic unconscious. Davidson's argument relates to Freud's idea of essential divisions within mental life, but is general enough not to depend upon some archeological metaphor of the mind in which there are literal divisions in consciousness as radical as id, ego and superego. I agree with Davidson, and would extend his work to argue that the necessity of alter-agency in explaining akrasia sheds light on how Freud's innovative conception of the dynamic unconscious is burdened by the archeological metaphor. What Freud was trying to account for was how there could be, in addition to rational instrumental thought and agency, an alternative way of structuring reality and hence an alternative way of causing action, codetermining one's conscious acts. What is not new with Freud, and not essential to the point of dual agencies, is that the alter-agent be determined by a buried past that is out of awareness. (This is not to deny the enormous importance of the idea of forgotten trauma to Freud's particular conception of repression).

The point is that the metaphor of the unconscious as a subterranean force that is blind to the here and now, which occasionally erupts like a volcano into the present, is wrong and misses what is radical in Freud's notion of a dynamic unconscious. In fact, I would turn the archeological metaphor for the unconscious on its side here, and point out what is meant by the dynamic unconscious is not some underlying (hence blind) force that imposes the past on the present without attending to the here and now. This picture of a blind unconscious mind is inconsistent with the core empirical discovery of psychoanalysis, which is that we can read into the manifest content of our utterances/gestures, a latent, primary process message. (For this emphasis on the unconscious as living text I am indebted to James Grotstein, personal communication). The unconscious speaks to us in translation, using here and now events to garb itself.

But if the unconscious mind did not attend to the present aspects of our situation, then how could any semantic links between here and now conversation and the unconscious be established? Our unconscious mind and our conscious mind must co-determine our here and now attention. (This point is also presupposed by the idea of symptoms as compromises, another central tenet of all psychoanalytic theories).
In order to see how Dr. L's invoking emotion B to act on emotion A could be role-directed, consider a familiar example of self-persuasion. Think of a situation in which you felt very angry (emotion A) at the inconsiderateness of someone you love, and tempered that anger by conjuring up thoughts about times when that person treated you lovingly. Such thoughts would likely be accompanied by affectionate feelings (emotion B). Chances are that you spent some time in an ambivalent state in which you genuinely felt both anger and affectionate gratitude towards the same person, construed in two different ways. Depending on a host of factors including your own emotional character/embedded associational ties, but also including the time and freedom of thought available to you, and the nature of the rejecting and loving behaviors that you are focusing on, the impact of B on A could go in several different directions. You might wind up in a complex emotional state in which a feeling of affection coexists with subdued anger (emotion C1). The focus of your emotions would remain your friend, but in his or her different aspects. Or you might wind up in a more self-reflexive emotional response which takes your own ambivalence as its focus. Depending on how harsh or accepting of yourself you are you could wind up with any of a host of emotions including the following: a feeling of self-pitying resignation which has as its focus the unpredictability of your friend (emotion C2), or a feeling of being proud of yourself because you can tolerate the bad with the good in others (emotion C3), or a feeling of shame for trying to flee your angry feelings (emotion C4).

Note that in this example there is no reason to posit that the agent has to go from emotion B to any particular emotion C. It is not that conjuring up loving feelings is sufficient for determining any particular response from C 1-n, but rather that because emotions determine salience, the occurrence of B will direct the agent towards certain further responses rather than others. All of the responses C 1-n are influenced by the affectionate construal of one's friend; they are thus different ways of sustaining the role of being a friend.

Websters New World Dictionary, op.cit.

My discussion of family therapy is influenced by my clinical experiences and the ideas of Melvin Lansky and Benahz Jalili, who teach family therapy at UCLA Neuropsychiatric Institute.

Of course there are important differences in how Dr. L understands a "couple" that includes himself and the patient compared to the couple therapist's understanding of an independent couple. But as I argued in chapter four, the absence of a distinct mental act of introspection suggests that introspective self-awareness involves something like empathizing with oneself. For this reason I believe the couple therapist's functions can be usefully applied here, as long as the point that this is a metaphor is not forgotten.
212. Psychiatric training involves one to one supervision of trainees with senior psychiatrists, where the task is to teach the neophyte to become curious about her difficult emotional responses to patients. One result of such curiosity is an increased ability to empathize with patients others find off-putting.

213. The analyst Eve Schwaber writes about an analysis in which for several months every interpretation she made was immediately rejected by her patient. She felt incompetent and irritated, but these experiences were muted by her strong curiosity about how the patient was experiencing her interpretations. This decentering curiosity about the patient's subjective world allowed her to listen empathically to the patient. She came to understand empathically that the patient was longing for closeness and understanding, but wished to be understood non-verbally and immediately, and thus found the analyst's tentative interpretations painful and rejecting. See Evelyne Albrecht Schwaber, "On the Mode of Therapeutic Action: A Clinical Montage," How does Treatment Help? On the Modes of Therapeutic Action of Psychoanalytic Psychotherapy, Workshop Series of the American Psychoanalytic Association, Monograph no. 4, Rothstein, ed., Connecticut: International Universities Press, 1988.

214. The claim that Dr. L's curiosity and prideful reconstrual of his situation plays a causal role in shifting him into an empathic stance raises certain surface conflicts with my account of empathy. In particular, I argued that for empathy to be possible, it must be possible for shifts in one's affects to arise prior to any deliberate construals of another's situation. I argued that the fact that such shifts could be directed by the other person's affects, via a kind of pre-verbal resonance, helped explain the magic of empathy, the fact that the affective experience one happened to have as a listener could direct one to a new understanding of the other person's situation. It would contradict the results of my empathy discussion to posit that Dr. L deliberately imagines the patient's world in a detailed first personal way that conveys the helplessness of the patient, and then feels sadness with the patient. This would be sham empathy, which involves no genuine communication.

However, what I rejected in the empathy chapter was the idea of being able to infer another's mental contents, not the idea of being able to imaginatively amplify one's own mental contents. Dr. L shifts his construal of his interaction with the patient prior to resonating with the patient, so that his empathic understanding of the patient has multiple determinants, including his own deliberate imagination work. Recall that in the empathy chapter we pointed out that affective resonance alone was insufficient for empathy (argument against the "merging" hypothesis). Rather, the physician relies upon her accumulated knowledge of the first personal details of the patient's life, built into a working model of his or her world as lived, to steer her construals of the patient's situation. Hence the great increase in
accuracy one has in empathizing with familiar versus unfamiliar people. Although this working model can direct one's empathic imaginings unconsciously, one often has to consciously invoke parts of it to gear into an empathic response to a patient. So in this vignette Dr. L deliberately shifts his focus from his personal irritation with the patient, which precludes empathic listening, to an image of himself as a good physician who can listen to his patient. This shift is facilitated by curiosity about the patient's experience, which decenters him from his own irritation and facilitates listening. The activity of listening to the patient or of "trying" to empathize, involves building an initial working model of the patient's world including the idea of having cancer and being alienated from others. But all of this imagination work does not yet include the particular affective grasp of the patient's fear of showing his pain to his family that Dr. L then arrives at through empathic resonance with the patient.

215. My own view of how therapy is effective is that within the safety of a strong therapeutic relationship, the patient is invited to become curious about her own sincerity, in the Sartrean sense. The idea that one can be for oneself only what one is, is challenged. The patient realizes that her emotions are already appropriated by others for strategic reasons, and that she herself has inadvertently colluded in these uses of herself. I am indebted here to Judith Broder, personal communication. See Sartre on sincerity as the apparent antithesis of "bad faith," in Being and Nothingness, Washington Square Press, 1966, pp.100-101.


217. The idea that curiosity detaches one from other emotions is consistent with our view that attention is always directed by mood, because curiosity is itself an affect. One is moved to wonder.

218. It would be interesting to empirically compare the value of the therapist's curiosity versus the value of her warmth in building up a therapeutic alliance with patients, using measures for therapeutic alliance that correlate with outcome measures for psychotherapy.
219. Although we have rejected Freud's statements that the analyst should have affective neutrality, his conception of the analyst's cognitive stance is close to our conception of curiosity. Freud, Sigmund, "Recommendations to Physicians Practicing Psychoanalysis," Standard Edition, 12: 109-120.

220. Freud, ibid.

221. However, I again differ with at least a traditional reading of Freud in that I take it that there is no single archetypal experience of evenly hovering attention, guided by the external landscape, free of particular personal meaning. The train rider's experience of moving away from some town will be colored by a particular attitude -- such as nostalgia, grief or relief -- about the miles separating her from the town.

The image of the moving train suggests how curiosity can attenuate one's emotions without requiring affective disengagement. Consider, for example, how one can reflectively experience a past emotion of one's own with combined involvement and curious disengagement. I may still feel moved by the image of a teen-age boyfriend, but I am also aware of the childishness of his ways, and of the difference between what felt like love then and what feels like love now. This experience is characteristic of the sense of ourselves as changing over time. One relives empathically some of the old longing or fear, but also feels that the construal of the loved or feared one no longer has the same value or meaning.

222. See Spiro, "What is Empathy and Can It Be Taught?" op.cit. Spiro emphasizes the importance of giving narrative accounts of illness on clinical rounds as the patient's situation is presented. I agree with this integrated approach, which sees empathy as essential to, rather than ornamental to, medical care.

223. Fox, "The Human Condition of Health Professionals," in Essays in Medical Sociology, Transactions Press, 1988, p.582. I agree with this integrated approach, which sees empathy as essential to understanding patients, rather than as ornamental.


225. To non-physicians, doctors' jokes appear to be cruel portraits of the ridiculousness of patients -- for example, the obtunded patient with his tongue hanging out of his mouth is said to be showing the 'Q' sign. See Samuel Shem, The House of God, New York: Dell, 1978, a novel about internship.

What is apparent from the inside is that it is usually the most pathetic experiences of patients that inspires humor. The subtext of these jokes is
physicians' own feelings of helplessness towards, and their feelings of identification with, their patients. That is why a simple sentence from the novel *House of God,* is almost universally amusing to physicians -- "remember, the patient is the one with the disease." Most importantly, such jokes allow physicians to tell each other how they feel, and thus to make empathic contact, without breaking the code of suffering in silence -- there is no apparent suffering, only laughter. Humor of this sort starts with one's first course in medical school-- gross anatomy, where for example, my own group named our human cadaver Earnest, because we were "working in dead Earnest." This allowed us a no-risk way to tell each other how strange it felt to dissect a person who was dead. Like curiosity, humor allows one to slacken the lines of one's identification with one's role, to note how strange the normal practice of medicine can be. But humor stops far short of curiosity in that it typifies the present experience rather than liberating one to imaginatively reconstrue and thus transform one's situation. For a discussion of the sociology of humor in medicine, see Rene Fox, "The Human Condition of Health Professionals," op.cit., pp.579-581.

226. Judith Ross, personal communication.


228. Consider, for example, a story told by William Carlos Williams about the impact of resonance emotion on a physician's anti-semitism. The doctor was visited by a Jewish European couple who asked him several anxious questions about their apparently healthy infant, and were mistrustful of his handling of the baby. They barely spoke English, and seemed in his eyes foolish and superstitious. Yet on learning that the woman had lost her entire family in Nazi-occupied Poland, the physician says he was "touched," and began to respect her protective attitude toward the baby. This shift in his perspective allowed him to communicate effectively, and help them for the first time. See "A Face of Stone," *The Doctor Stories,* New York, New Directions Books, 1984.

229. This is relevant to the current debate about how best to restructure cost-effective medical care. The values of medical practice are not well served by any direct commercialization of medical care. Rather, given the urgent need to restrict the cost of health care, the kind of physician-patient relationship argued for in this dissertation would be better grounded in a form of health care, such as National Health Insurance on the Canadian model, that emphasizes a long-term relationship with a primary care physician who is not a direct beneficiary of any limiting of care.

231. For a practical defense of the claim that the physician's moral reasoning needs to grow out of her moral experience in daily medical practice, see Edward Hundert, "A Model for Ethical Problem Solving in Medicine, with Practical Applications," *American Journal of Psychiatry*, 144:7, July 1987, pp. 839-846. Hundert's views are compatible with my account of the importance of perceptions of moral salience for the Kantian moral agent.


Freud, Sigmund, "The Interpretation of Dreams," Standard Edition, translated and


Herman, Barbara, "Integrity and Impartiality," The Monist, April, 1983.


Herman, Barbara, "Rules, Motives and Helping Actions," Philosophical Studies, 1984.


Rawls, John, Unpublished Lectures on Kant's doctrine: The Fact of Reason, given at the Johns Hopkins University Summer Program in August 1983.


Sartre, Jean-Paul, Being and Nothingness, Hazel Barnes, trans., New York: The Philosophical Library, 1956


**Books on the Doctor-Patient Relationship**


Hooker, Worthington, *Physician and Patient, A Practical View of Medical Ethics*,


Articles on the Physician-Patient Relationship


Psychotherapy, 6:39-64.


Fox, Renee, "The Evolution of Medical Uncertainty," 58 Milbank Memorial Fund Quarterly 1, 1980.


Pellegrino, Edmund D., "Toward a Reconstruction of Medical Morality: The


