

Jean-Paul Sartre's ***Being and Nothingness***

Class Lecture Notes
Professor Spade
Fall 1995

Table of Contents

Getting Started.....	4
Sartre: Life and Works	5
Program of Events	11
Two Main Influences on Sartre	11
Husserl: Life and Works.....	13
The Idea of Phenomenology.....	14
Kant	18
Review	25
The Two Stages of Husserl's Philosophy	26
The Idea of Phenomenology (Again)	27
The Phenomenological Reduction.....	31
The Eidetic Reduction	39
The Theory of Intentionality	46
Sartre	56
Sartre's Reaction to Husserl	63
Sartre's Metaphysics	72
Characteristics of Being-In-Itself	73
Being-For-Itself	80
Positional & Non-positional Consciousness, Reflective & Non-Reflective Consciousness.....	87
The Self-Love Theory	93
The Constitution of the Ego	96
The Magical	104
The Problem of Other Minds.....	114
The Origin of Negation	115
Hegel and Heidegger	123
The Origin of Nothingness.....	125
The Gambler	128

Vertigo	130
Bad Faith (Self-Deception).....	133
The Waiter.....	138
Belief	141
The Emotions.....	149
The Intellectual Theories	158
Sartre's Own Theory	160
The Magical World.....	163
False Emotions and the Physiology of The Emotions	165
Part II: Being-For-Itself	166
Presence to Self	171
Facticity	172
Lack.....	175
Value.....	177
Possibility	181
Time.....	183
Pure and Impure Reflection	191
The Existence of Others	197
Husserl.....	206
Hegel.....	208
Heidegger.....	209
Summary.....	210
The Look	211
Concrete Relations with Others.....	220
Examples of the First Approach	223
Examples of the Second Approach	224
Existential Psychoanalysis.....	225
Conclusion	234
Ethical Implications	238

This text looks at the other phenomenologists from Sartre's standpoint. For true accuracy, it is indicated to read the original texts of the other phenomenologists mentioned

1

Getting Started

The main textbook for this course is Sartre's *Being and Nothingness*, of course. But it will be quite a while before we actually get into that. There's a lot of build-up and background that you need to get a kind of running start on that book.

We are going to start with Edmund Husserl, *The Idea of Phenomenology*. I have not asked you to buy this book, but it is available on reserve. You should start reading that book immediately, and consult the outline included in the course packet.

The next main thing we will be reading is Sartre's *Transcendence of the Ego*. This is a difficult but extremely exciting book on the Philosophy of Mind. It introduces many of the main themes we will see in *Being and Nothingness*.

Only then will we be in a position to plunge into *Being and Nothingness*. We will start at the beginning and go as far as we can in one semester. Then, as we near the end of the semester, we will skip ahead to the section on "Existential Psychoanalysis" (near the end of the book), and the "Conclusion." They are important, and I want to be sure we do them.

Along the way, there are two books by Sartre on the imagination and one on the emotions. These are very interesting books, but for our purposes are subordinate readings. One of the books on the imagination, *Imagination: A Psychological Critique*, is now out of print. But there is a copy on reserve in the main library, and an outline included in the course packet. The other one, *The Psychology of Imagination*, contains one crucial passage that will be tremendously important. But, for the most part, that book is left for your own background reading. The same goes for *The Emotions: Outline of A Theory*. Don't neglect these two books, but they won't be centerpieces in the course.

We surely won't be able to get through the whole of *Being and Nothingness* in this one semester. Nevertheless, we should get far enough along that, by the time we are done, you will have the background to be able to read the rest of the book on your own — if you should wish to.

And you should wish to. In my judgment, *Being and Nothingness* is probably the single best piece of philosophy written in the 20th century. That is a strong claim, and I don't make it lightly. There is lots of good philosophy in the 20th century, but this book has a kind of sweep and scope that, as far as I know, no other work has in this century. There may be exceptions — for example, Heidegger's *Being and Time*, which I do not know well — but within the limits of my knowledge, *Being and Nothingness* stands out as without serious competition.

What are the alternatives? Husserl's *Logical Investigations*, for one, and his *Ideas*, for another. Heidegger's *Being and Time*, perhaps. Russell and Whitehead's *Principia Mathematica*, and Russell's *Principles of Mathematics*. Perhaps Whitehead's *Process and Reality*. Wittgenstein's *Tractatus* and *Philosophical Investigations*. Some people

Copyright © 1996 by Paul Vincent Spade. All rights reserved. Permission is hereby granted to copy this document in whole or in part for any purpose whatever, provided only that acknowledgment of copyright is given.

would nominate Quine's *Word and Object*, which is a work for which I have the highest respect.

But all these, in my considered judgment, are no deeper philosophically than Sartre's *Being and Nothingness* is, and are certainly less ambitious in scope. I hope to convince you of this during the course of the semester.

As I said, the later parts of *Being and Nothingness* are much easier than the earlier parts. This is not just because the earlier parts are presupposed by the later ones; the later parts are just plain easier. So, although we won't get through the entire book, you should be in a good position to complete it on your own.

Let me suggest some background reading before we get started:

Frederick A. Olafson, "Sartre, Jean-Paul," in *The Encyclopedia of Philosophy*. An OK article, but no great shakes.

Hazel Barnes' "Introduction" to *Being and Nothingness*. A pretty good overview, although it is rather difficult. It's good to read it early on, but don't expect to understand it until later.

Alisdair MacIntyre, "Existentialism," in *The Encyclopedia of Philosophy*. This is an excellent article, although people have raised questions about details of it.

Alisdair MacIntyre, "Existentialism," in Mary Warnock, ed., *Sartre: A Collection of Critical Essays*. This is not the same as the previous article, but is also excellent. This book is now, I think, out of print, but I have put a copy of the article on reserve in the main Departmental office. (It's about the whole movement, not just Sartre.)

Herbert Spiegelberg, *The Phenomenological Movement: A Historical Introduction*, Ch. 10. A fairly good account for those just getting started. Also, full of lots of lore and gossip about these people, and good pictures!

Sartre: Life and Works

Jean-Paul Sartre was born in Paris on June 20, 1905, and died there April 15, 1980. He studied philosophy in Paris at the École Normale Supérieure in Paris 1924–1928. After that he taught philosophy for a while in a number of *lycées*, in Paris and Le Havre (and perhaps elsewhere). He then went to Germany, to the Institut Français in Berlin. He had some kind of research assistantship there, but in any case during 1933–1934 he studied there under two giants of twentieth-century German philosophy:

- (1) Edmund Husserl, the father of modern phenomenology, who died in 1938.

Copyright © 1996 by Paul Vincent Spade. All rights reserved. Permission is hereby granted to copy this document in whole or in part for any purpose whatever, provided only that acknowledgment of copyright is given.

- (2) Martin Heidegger, who died in 1976. Heidegger was a student of Husserl's, and so in a real sense part of the phenomenological movement, although he went off very much in his own direction and was pretty much the originator of twentieth-century existentialism.

Sartre actually *met* Heidegger at one point, but always seems to have felt a closer intellectual kinship to Husserl, even as he came more and more to disagree with the master.

In 1935 he was appointed professor of philosophy at the Lycée Condorcet in Paris. The little biographical sketch on the back flyleaf of the English *Being and Nothingness* says he held this position until 1942. But Spiegelberg¹ says he resigned his position there in 1944. I do not know which is correct.

In any case, he didn't spend all those years from 1935 to 1942 (or 1944) teaching, because of course there was a big war going on. In 1939 he was mobilized and drafted into the French army, where in 1940 he was captured and held prisoner in a Nazi prison camp. He spent his time there writing and directing plays for his fellow prisoners. After nine months, he was released, in 1941, and returned to Paris and to his teaching.

But of course the war was still going on, and Sartre joined the French Resistance movement as a writer for various underground newspapers. You will see signs of Sartre's war-time experiences throughout his writings. They provide a rich source of *examples*, for instance.

All during this time, he published novels, plays, philosophical writings, essays, criticism, and so on. After the war he continued to do this right up to the time of his death, although he certainly slowed down toward the end. He was always involved in political and literary issues. In 1964 (the flyleaf to *Being and Nothingness* says 1965) he was awarded the Nobel Prize for literature, but declined it. (This just means he didn't take the money. He was and remains a Nobel laureate; you can't turn down the honor.)

Main Writings:

On Sartre's writings, you may want to look at Ch. 1 of Peter Caws book *Sartre*, the chapter called "A Conspectus of Sartre's Writings." There is a copy on reserve in the main library, and I have put a xerox copy of Ch. 1 on reserve in the main Departmental office. While I am not going to insist on your knowing all the grimy details, I am going to expect you to know the main facts about Sartre's writings when it comes time for the first quiz next Wednesday.

His earliest publications come from 1923, when Sartre was only 17 years old. These are two short pieces of fiction, with the intriguing titles "The Angel of Morbidity" and "Jesus

¹ Spiegelberg, 2nd ed., p. 450.

the Owl, Small-Town Schoolteacher.” (These are both on reserve in translation in Sycamore 026.)

There are other things as well from these early years, including an interesting fragment of a piece of philosophical fiction called “The Legend of Truth,” published in 1931. All of these have been translated, and I can give you the references if you want.²

But for the most part, Sartre’s *philosophical* writings can be divided conveniently into *three* main periods. In this course, we will be concentrating on the *first two* of them, and not on the third. (But the philosophy of his third period is fair game for your paper topics.)

I

The Phenomenological Period (1936-40):

Sartre’s earliest philosophical writings were very *phenomenological* in orientation, written very much under the influence of Husserl. They may be viewed as “in-house” writings within the phenomenological movement.

Among the earliest of his works, and the first *main* work we will be looking at in detail, is:

- (i) *Transcendence of the Ego*, published in either 1936 or 1937, depending on how you count it. You see both dates given. The cover of our paperback translation says 1937. But Barnes’ “Introduction” to *Being and Nothingness* says 1936, and this is confirmed by Caws (p. 10). The problem is that it came out in a journal, *Les Recherches philosophiques*, vol. 6 for 1936–1937. This is one of those journals where the division into volumes is out of synch with the calendar year. I think the correct date is 1936, but I haven’t really tracked this down, and don’t really care.

Some of Sartre’s main themes are already present in this work. It is immensely rich. In this work, he distinguishes his view of the nature of the “Ego,” the “I” or “Self” from Husserl’s later views. The book is basically a discussion of the nature of consciousness, self-awareness.

Sartre was also interested from the very beginning in *psychology*, partly because of his phenomenological background. As a result, he wrote:

- (ii) Two works on *imagination*. For Sartre, the fact that human beings have the peculiar ability to *imagine*, and so put themselves in some kind of mental relation to, things that *don’t exist* is very important.

² In Michel Contat and Michel Rybalka, eds., *The Writings of Jean-Paul Sartre*, Volume 2: *Selected Prose*, Richard McCleary, tr., (Evanston, IL: Northwestern University Press, 1974).

In these two early books, he explores and criticizes the psychological theories of his day, and sets out his own views.

The first of these two works is *L'imagination*, which appeared in 1936, and has been translated under the title: *Imagination: A Psychological Critique*. I was originally going to ask you to read this book for our course, but the translation is now out of print. There is a copy on reserve in the main library, and I have included an outline of the work in the course packet. It is an interesting book.

The second work is *L'Imaginaire*, translated as *The Psychology of Imagination*. It was published in 1940, and is an exceptionally interesting book. I have asked you to buy it for this course. Most of the book will be simply background reading, and we won't be dealing with it directly. But there is one passage that will be central to our understanding of a lot of things in Sartre. I'll deal with that when the time comes.

- (iii) Also during this early period, Sartre wrote a book on the *emotions*. This too is a very interesting little study, and I have asked you to buy it for this course. There is an outline of it in the course packet. It depends on how the course goes, but I doubt if we will be discussing much of this work directly in class. Nevertheless, there are some central notions that we *will* be discussing directly in class. We will not be reading it directly in this class, but I will have occasion to refer to it directly from time to time. The title is *The Emotions: Outline of A Theory*, and it appeared in 1939.

Also during this early period, there were a number of plays and novels. Probably the most important novel from this period (and probably his most important novel *of all*) is:

- (iv) *La nausée*, translated as *Nausea*. A very odd “philosophical” novel. Published in 1938.

All of these writings may be grouped together in Sartre's “early” or “phenomenological” period. (He was influenced by phenomenology for a long time, but this influence is perhaps strongest at the very beginning of his career.)

II

The Existential Period (1943-1952):

The second main period in Sartre's philosophical career might be called his “existential” period. It is marked by his *magnum opus*:

- (i) *Being and Nothingness* (1943). This is a huge work, of 800 pages or so. It is our main text for this course. It is *very* exciting — in my opinion, probably the best book of philosophy in the twentieth century — but also, as you will see, *very* difficult.

Basically, *Being and Nothingness* is an ontological analysis of human existence. It is a very uneven work. Parts of it can be readily understood without any special preparation. Part of it a jargon-laden and deliberately obscure. Parts of it are truly famous.

Everything else we will be reading this semester will be simply to elucidate or elaborate on the themes in *Being and Nothingness*.

Also, during this period, Sartre published a brief essay:

- (ii) “Existentialism Is A Humanism” (1946). In this essay (it was originally a public lecture), Sartre tried to set out for the general intellectual reading public in France the main themes of his “existentialism.” Because it is addressed to a non-technical audience, it is written in quite plain language and is quite easy to read.

(If you have not already read it, I am going to ask you to read this pleasant little essay in connection with this course. I have a discussion of the essay in the course packet. Pay particular attention to that discussion, because I am simply going to *presuppose* it in lecture when we get to that point.)

In the same year (1946), there also appeared an excellent essay:

- (iii) *Anti-Semite and Jew* (1946). This is a study of Anti-Semitism, which was a conspicuous problem in 1946, when France was just coming out of World War II and the Nazi experience. For our purposes, the interesting thing about this essay is that it amounts to a kind of “case-study” of what Sartre calls “Bad Faith” or self-deception. This notion of “Bad Faith” will be absolutely crucial to our study. The book is non-technical, easy reading, and — I think — a stunningly insightful essay.

Finally, also during this period, I should mention three other items:

- (iv) *No Exit* (1944). A short and *very* fine play with strong philosophical overtones. In effect, the play is a kind of dramatic presentation of Sartre’s theory of inter-personal relations. The theory is not a pretty one, but the play is excellent — in my opinion, Sartre’s most successful play. In fact, it is probably the

most successful attempt I know of to incorporate serious philosophical themes into fiction.

- (v) *What Is Literature?* (1948) A moderately short essay discussing the differences between poetry and prose, from a phenomenological point of view. A rather interesting discussion.
- (vi) *Saint Genet: Actor and Martyr* (1952). This is a kind of philosophical biographical study of Jean Genet, the famous French author. It's a big book, and I've not read it. But, from what I know about it, it is important for understanding how Sartre's thought developed between the time of *Being and Nothingness* and the next big period of his writings, to which we now turn.

III

The Marxist Period (1960-1980):

Finally, in Sartre's third main period, he moves to a kind of Marxism. I say "a kind of" Marxism, because Sartre was never a Marxist of the strict observance. (He could not accept Marxist materialism, for instance. In a late interview, he says he always thought materialism was ridiculous on the face of it.) The main work here is:

- (i) *Critique of Dialectical Reason*, vol. 1 (1960). There was a second volume, published posthumously. Some people describe this work as an *abandonment* of the existentialism of *Being and Nothingness*. But it is perhaps better regarded as just a kind of going beyond *Being and Nothingness* to consider themes that were not very well developed in that earlier work. These new themes concern the *social* order. (As you will see from your reading about Sartre, there is considerable controversy over just how to view this last main period of his writings in relation to his earlier "existentialist" period.)

When the *Critique* was published in 1960, it was preceded at the front of the volume by a more or less independent methodological essay that was been translated into English before the rest of the *Critique* was translated. You can find it under the title *Search for a Method* or *The Question of Method*. It was translated by Hazel Barnes (the translator of *Being and Nothingness*) in 1963. The *Critique* proper was translated by Alan Sheridan-Smith in 1976. *Search for a Method* was not included in that volume (since it had already been translated separately). In the original French, this introductory essay did not appear in print until the *Critique* as a whole was published in 1960. But it has been *written* somewhat

earlier. In any event, be aware that there is a close connection between those two works.

- (ii) *The Family Idiot*. This is an enormous multi-volume philosophical biography of Gustave Flaubert, the French author. The first volume of it was published in 1971. I have not read any part of this work, although it has been translated into English. It is Sartre's last main work. And it seems to be interminable!

In addition, we should remember that there were lots of articles, essays, interviews, plays, etc. that continued to appear throughout Sartre's literary career. We have only touched on some of the main ones. Once again, you may want to consult Peter Caws' Ch. 1.

Program of Events

Here is our plan of attack:

I will begin by talking a little about Descartes and Kant, to set the stage for Husserl, who was one of the main influences on Sartre.

Then we will look at Husserl's *The Idea of Phenomenology*. After that, we will turn to Sartre himself. It is at this point that you should familiarize yourself with "Existentialism Is A Humanism," if you have not already done so.

We will read *Transcendence of the Ego* (a crucial book), and then finally start on *Being and Nothingness*. So — be aware — we will spend a big part of the semester before we ever get to *Being and Nothingness*. That's part of the plan, not just a matter of getting behind. The preliminary material is *not* just a delay. As we'll see once we get to *Being and Nothingness* itself, it will go fairly quickly after we've done all the preliminary work.

Two Main Influences on Sartre

Sartre's early philosophy is strongly influenced by two streams of thought:

The Reactionary Stream:

A stream typified by Nietzsche (the first person mentioned *by name* in *Being and Nothingness*). In effect, this tradition is a *reaction* against the philosophy of the 18th century, with its unbounded confidence in the ability of *reason* to solve all our problems — philosophical, scientific or social. This tradition came to a kind of peak in Hegel.

I have to qualify that a bit. Scholars of Hegel himself will have a different point of view. But we're not really in disagreement. What I am talking

about is Hegel as *certain other people* viewed him, not Hegel as he regarded himself, and *certainly* not Hegel as *we* view him today.

Sartre's own attitude toward Hegel is perhaps a little strange to modern readers. Oddly enough, Hegel was almost totally unknown in France until after World War I, when Alexandre Kojève and Jean Hyppolite began to introduce Hegel to French intellectuals. And the main work they were interested in was Hegel's *Phenomenology of Spirit*, not the *Logic* and not Hegel's other writings.

Kojève's and Hyppolite's interpretations of Hegel are nowadays regarded as pretty unorthodox. Nevertheless, this is what Sartre knew. So, if you know something about Hegel on his own, don't expect it to conform necessarily with what Sartre says about him.

But before we get to Sartre, there was the nineteenth-century interpretation of Hegel, at least in certain quarters. And there he was regarded as a kind of arch-rationalist of all time. There was a reaction against this kind of thinking. The reaction included Kierkegaard and Nietzsche in the nineteenth century, and (although it was no longer perhaps especially associated with Hegel) the existentialists in the twentieth.

From this reactionary stream, Sartre inherited:

- (a) The view that traditional philosophy is *bankrupt*, that there is no future in old-style philosophy. We need to do something *radically new*. And furthermore, intellectual society *as a whole*, according to this view, has come to realize this. Thus, for example, we'll find Sartre forging a whole new terminology of his own, one that he feels is free of the connotations built into the old, traditional terminology.
- (b) An emphasis on the *individual*. The old-style philosophy tried to *categorize* everything in nice, neat *rational* pigeonholes. It tried to *systematize* everything in one complete theory of reality. It did this to such an extent that the *rational categories* came to be viewed as more interesting, more important, than the *individuals* that fit more or less *into* those categories.

We find this emphasis whenever we do *science*. The scientist is not interested in what happens to a particular specimen of a chemical in a test tube, or a particular culture in a petri dish. He is interested in this only insofar as it reveals something about the general laws governing *all* similar cases.

From a slightly different angle, the old-style philosophy emphasized the *state* at the expense of the individual citizen. For

example, Hegel, who had a great deal to say about the theory of the state.

The *reaction* against this switched the emphasis to the *individual*.

- (c) Going along with this emphasis on the individual, there is also an emphasis on *individual responsibility*. The individual cannot appeal to *general principles* or *universal laws* of human or social behavior to shift the burden of responsibility for his actions off his own shoulders. Remember, this reactionary tradition *downplays* all these *general* appeals.
- (d) Along with the emphasis on individual responsibility, there is a correlative emphasis on *human freedom*. (This theme is not so strong in all authors in this tradition. It is perhaps not so strong in Nietzsche. But it is there in Kierkegaard, for example, and it is *certainly* there in Sartre.)

All these features show up in Sartre's doctrine. They are most evident when Sartre is discussing the *ethical, moral* side of his philosophy.

The Phenomenological Stream:

The second main stream that influenced Sartre was *phenomenology*. This influence is most evident when Sartre is discussing the *metaphysical and epistemological* sides of his philosophy. It is this influence that I want to begin with in this class.

Sartre got this influence through Husserl, and also through Heidegger.

In order to see what is going on here, we must go back and look at Husserl, and at the *origins* of the problems Husserl was addressing.

Husserl: Life and Works

Husserl was born in 1859. He studied in Vienna (in part under the great Franz Brentano), and in Berlin. He died in 1938.

Husserl's philosophy developed through several stages. You should know about the following works, since I will have occasion to be referring to them:

- (1) *Logical Investigations*. The first part of this work appeared in 1900, so it's easy to remember.
- (2) *The Idea of Phenomenology*, which was done in 1907, although it wasn't actually published until 1950.
- (3) "Philosophy as Rigorous Science," an article from 1911.

- (4) *Ideas*, vol. I, which appeared in 1913. This is perhaps his main work.

There were also many later writings, and there remains a lot of unpublished materials. Husserl was a *tremendously* prolific writer.

From the later period, I should perhaps mention:

- (5) *Cartesian Meditations*, published in 1931 and based on a series of lectures Husserl delivered at the Sorbonne in Paris, in 1929. I do not know exactly how much Sartre knew about the material of these lectures. (And it is something I would *like* to know.) We do know that he was not himself *at* the lectures when they were given.

The Idea of Phenomenology

I want to look at *The Idea of Phenomenology*. This too was a series of lectures, given this time at Göttingen. While he was preparing the lectures, Husserl also wrote a kind of private outline to himself, which is included in the English translation under the title “The Train of Thought in the Lectures.” It is instructive to compare “The Train of Thought” with the actual lectures themselves, since *they don’t always agree*.

Husserl was in the middle of a major transition stage in his own thinking, and the lectures show his own unsettled state of mind on certain topics. I will want to discuss what it is a transition *from* and what it is a transition *to*. Both are important for understanding what Sartre is up to in *Transcendence of the Ego* and elsewhere. There is no reason to think Sartre knew anything about *The Idea of Phenomenology* at all. So I am not talking about it because it was *influential* on Sartre (there is no evidence that it was), but only because it is illustrative of things that *were* influencing Sartre.

The *problem* Husserl is addressing in these lectures is, as he puts it, “the possibility of cognition” (Lecture I, p. 15; “Train of Thought,” p. 1) — that is, the possibility of *real knowledge* of objective reality. So it an *epistemological* problem.

Here is how he puts the question in Lecture I (p. 15):

Cognition in all of its manifestations is a psychic act; it is the cognition of a cognizing subject. The objects cognized stand over and against the cognition. But how can we be certain of the correspondence between cognition and the object cognized? How can knowledge transcend itself and reach its object reliably?

This was hardly a new problem. It is already to be found in *Descartes* in the seventeenth century.

In his *Meditations*, Descartes was concerned with the problem of *error* and how to avoid it in his philosophy.

Why? Well, this is a long story. But in part, the reason is that Descartes had an ideal of philosophy as a *rigorous discipline*. Ideally, philosophy should have all the certainty and infallibility of mathematics (when mathematics is properly done). The fact that philosophers can never agree on anything, as mathematicians can, Descartes regarded as a *scandal*. And he thought the situation could be corrected.

This ideal of philosophy is a very old one. We find it, for instance, in Aristotle's *Posterior Analytics*, where Aristotle presents us with his picture of what a science is. After Descartes, of course, it is still to be found in Husserl's article "Philosophy as Rigorous Science." In fact, Husserl thought that philosophy should be a *presuppositionless science* that takes *nothing whatever* for granted.

(To call it "presuppositionless" is not supposed to mean that philosophy has no starting points that serve as the bases for everything else. Instead, it means that it should have no *unexamined* starting points.)

Now, as I said, Descartes thought the situation in philosophy could be corrected, and that philosophy *could* be put on a rigorous foundation, with the result that *errors could be avoided*.

How did he propose to do this, to avoid error, to reach the ideal of philosophy?

Basically, Descartes thought errors arose from what we might call "jumping to conclusions," from saying more than we really know. The basic problem, for Descartes (*Meditation IV*), is that we're in *too big a hurry*. Our *desire* for knowledge goes far beyond what we can actually know, and sometimes — driven by this desire — we allow ourselves to take shortcuts and hurry along, with the result that we end up affirming that we know something that we really are not in a position to know at all. Hence, we fall into error. (I think Descartes was absolutely right so far.)

It follows, therefore, that the way to avoid error is really a matter of *discipline*. We can avoid mistakes if we *refuse* to allow our desire for certainty to outrun our real ability to know, and so by refusing to say more than we strictly know. Or, as by Descartes puts it, by affirming only what appears to us

- (a) so *clearly* that there is no *obscurity* in it, and
- (b) so *distinctly* that there is nothing *confused* in it.

In short, Descartes thought we could avoid error by confining ourselves to those things that appear to us so *clearly and distinctly* that there is simply no *room* for error.

This notion of "clarity and distinctness" (and the opposites "obscurity" and "confusedness") becomes a kind of slogan, a catchword, in the Cartesian tradition. Husserl himself uses the phrase in a reference to Descartes in "The Train of Thought" (p. 6).

Well now, all this is fine, but what things are we aware of in this “clear and distinct” way? That is, what things can we affirm with complete safety?

In the end, Descartes (in agreement with a long tradition) thought we “clearly and distinctly” perceive the things we are directly aware of — without intermediary — the things that are, so to speak, present to the mind in person, not by proxy.

And what are these?

Well, first of all, I am aware of my own existence. This is summed up in Descartes famous phrase “I think, therefore I am” (= *Cogito ergo sum*.) This “cogito” is a famous notion. We will see it referred to time and again in Husserl and Sartre.

The “cogito” will always be a kind of funny case. As somewhat more typical cases of what Descartes has in mind, consider: the oar in the water (explain)..

In this case, the way things appear to me is not necessarily the way they really are.

In general, with the exception of the self, which is always treated a special case, I am directly aware only of the way things appear to me — the appearances, the *phenomena*. I am not directly aware of the way they are *in themselves*.

Hence, we draw the conclusion:

I avoid all risk of error as long as I confine myself to a description of the phenomena, of the directly given.

Or, in other words:

The “safe” = the directly given = the phenomena.

(The first identity is a substantive claim, whereas the second one is merely a matter of terminological convention.)

Note: “*Describing*” the phenomena. Descartes doesn’t push this point himself (in fact he explicitly denies it), but Husserl will certainly push it later on. As soon as we begin to reason from the phenomena to something else — to argue from what is directly given to us to something that is *not* directly given to us, to *draw inferences* — we run the risk of error.

So far, what we have is a kind of rudimentary description of *phenomenology*. Husserl would accept everything we have said so far. Phenomenology, in Husserl’s sense, is not a *science* in the sense that physics or mathematics is a science. Phenomenology is not a matter of forming *inductive* theories to *explain* phenomena, and is not a matter of drawing *deductive* conclusions from them. Any such *going beyond* the directly given is risky and subject to error.

Phenomenology, then, does not *argue*; it *describes*. Husserl makes this point again and again. Phenomenology, for Husserl, is not a matter of learning to think clearly or to reason properly. It is a matter of learning to see all over again.

This “describing” of the phenomena is not a simple task. It involves discipline and training. Training in phenomenology is rather like the training a painter gets. The painter must learn to be sensitive to nuances that all of us in a sense *see*, although most of us don’t notice them.

As a result, phenomenologists often talk about the *inexhaustible richness* that is uncovered by the phenomenological method. There is a kind of aesthetic exuberance in much phenomenological writing. We will see some of this at its best in Sartre.

But now back to Descartes.

Descartes adds one additional principle that is important. He holds that the phenomena, what we are directly aware of, are one and all mental events: sense-impressions, direct pains, etc.

Recall the example of the oar in the water. My impression of the oar is a content of my mind, is *mind-dependent*, in a way that the real oar itself is not.

So, for Descartes we have a second principle:

The phenomena are all mental events, mental contents, mind-dependent.

(This too is a substantive claim, not just a matter of terminology.)

So it is as if we are in a kind of *mental movie-theater*. The *phenomena* are what we see on our movie screen, and those phenomena are *pictures, representations* of things and events going on in an “external” world out there beyond the movie-theater.

Given this, there is an obvious problem: How can we ever know *anything* about what is really going on outside the mental movie-theater? Or, in other words, how can we ever be sure that our phenomena are accurate pictures or representations of reality?

The threat here is *solipsism* — the view that I alone exist, I and the contents of my mind. Everything else is just a dream, a phantom, a product of my imagination.

Descartes’ theory then must answer this question: How are we going to rule out solipsism? How can we avoid the possibility that it might be correct? How are we going to be sure of *anything* outside my own mind?

This is *exactly* Husserl’s problem in *The Idea of Phenomenology*. As he says,

How can we be certain of the correspondence between cognition and object cognized? How can knowledge transcend itself and reach its object reliably?

Of course, given Descartes' two principles ("The safe = the directly given = the phenomena," and "The phenomena = mental contents"), the obvious answer is that we can't.

Descartes *tried*, by arguing that God exists and would not deceive us about such things. But most subsequent philosophers thought Descartes' dodge will not work. By what right can Descartes claim to be sure that God exists, if — *on his own principles* — all he has to go on is the contents of his own mind? Furthermore, if the argument *did* work, it would appear that we should *never* make mistakes. (God would not deceive us about the oar in the water any more than he would deceive us about other things.) But we obviously *do* make mistakes. In fact, this realization is what got Descartes going in the first place.

As he himself sets it up, Descartes' problem is *insoluble*. The only way we could ever be sure that our phenomena are accurate representations of external realities would be to look at the phenomena, on the one hand, and look at the external realities, on the other, and see whether they match up. But, *by hypothesis*, we can never look at the external realities. The only things, remember, we can be certain of, are what is directly given. (That's the first principle.) And on this theory, the external objects are never directly given; only the phenomena are. (That is Descartes' second principle.)

So, if Husserl is going to find a way out of Descartes' problem — and this is exactly the task of *The Idea of Phenomenology* — he is going to have to give up one or more of Descartes' two principles.

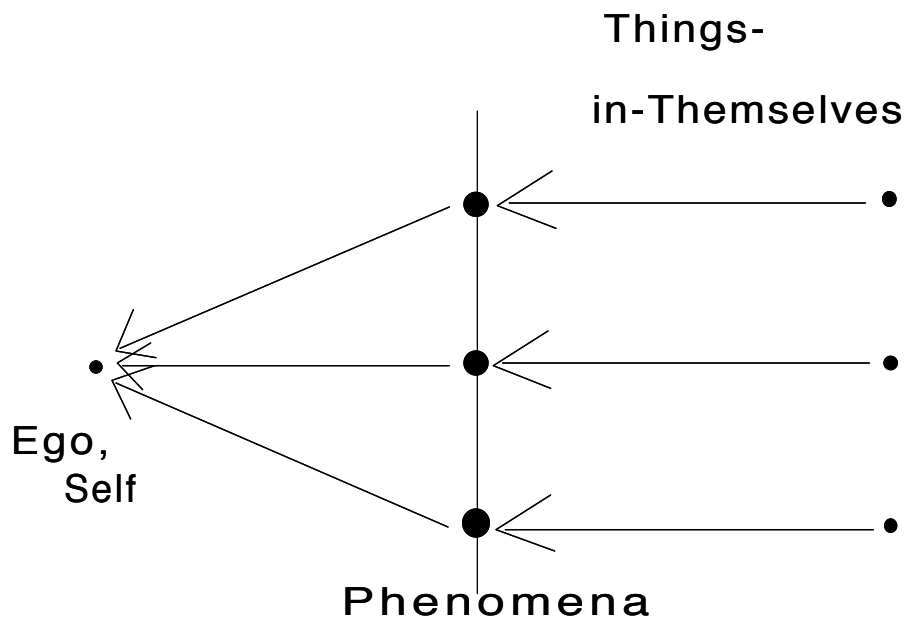
And he does.

But before we look at how he does this, I want to talk briefly at the subsequent history of Descartes' problem, up to Husserl's day, because many of important themes in Husserl and Sartre make their first appearance there.

Kant

Immanuel Kant realized what Descartes should have realized: that, given Descartes' two principles, it was hopeless to try to get any reliable knowledge of the realities behind the appearances — of what Kant called the "noumenon" (vs. the "phenomenon"), or the "thing-in-itself" (vs. the "thing-as-it-appears." We can *never* know the truth about the thing-in-itself.

But Kant went further than this, and he went further in two respects. In order to see what they are, let us diagram Descartes' theory:



Now Kant argued as follows: Descartes in effect assumed that the mind *contributed nothing* to the phenomena. All it did was *watch* them. For Descartes, the “self” or “ego” was simply a passive observer in its mental movie-theater. But, Kant claimed, that is not so. The mind in fact contributes a great deal to the phenomena.

For example (this is not Kant’s example), consider one of those “Gestalt” figures that can be seen now as a vase, now as two heads facing one another. In both cases, there is the same neutral *given*, the same geometrical figure consisting of a pattern of light and dark. But that same pattern can be seen in two different ways, depending on which (the light or the dark) is seen as the foreground and which as background.

What *determines* which way it is seen? That is, what determines how the figure *appears* to us — what determines which phenomenon I have? Obviously, the answer is that *I do*. That is, my mind does. My mind *organizes* the perceptual data in the one way or in the other, and interprets the data either as a vase or as two heads. So true is this that, with a little practice, I can learn to flip-flop from the one to the other at will.

In other words, in this instance consciousness is not altogether a passive observer of phenomena. It is active. It imposes a certain organization, a certain order on the raw data of sensation. The phenomenon, what in the end appears to me, is a product of two factors: the raw data of sensation, plus the interpretation imposed on those data by the mind.

This organizing and interpreting function of the mind is what is called Constitution — and it is very important. (The term ‘constitution’ is not Kant’s, but comes from the later tradition. But the doctrine is very much an authentically Kantian one.) The figure is

“constituted” as light foreground on a dark background, or “constituted” as dark foreground on a light background, and that “constituting” is done by the Ego.

Kant thought that the most *general* “categories” in terms of which we interpret the world — for example, notions like “causality,” “existence,” “substance/property” — are categories that come from us, categories that the mind *imposes* on the data. (These are the famous Kantian “categories.”) An Ego that behaves like this, an Ego that is not just a passive observer but an active constitutor of phenomena, is called a Transcendental Ego. (Get that notion down.)

Now Kant held that we have no right to think the “categories” apply to the noumena, to things-in-themselves, any more than we have a right, in the case of the Gestalt figure, to say that the light areas really *are* foreground and the dark ones really *are* background, in some ultimate “objective” sense.

In fact, this way of putting it leads us naturally into the second of the two ways in which I said Kant went beyond Descartes. (The first was in adopting the doctrine of “constitution.”)

Kant thought that not only could we *never be sure* that our representations, the phenomena, were accurate representations of the noumena or things-in-themselves — we could be quite sure they aren’t.

You can see this readily in the example of the Gestalt figure. It’s not just that we can’t be sure which one is “really” foreground and which is “really” background. Neither one is “in itself” — “absolutely” — foreground or background; the notions simply don’t apply at that “absolute” level.

That’s the basic idea, but let’s see how Kant puts it. In brief, his argument is this: He says that:

The “I think” must be capable of accompanying all our representations.

(Sartre refers to this claim at the very beginning of *Transcendence of the Ego* — on p. 32, after the translator’s introduction.)

What does the claim mean?

Basically, it means that whenever I am describing the phenomena, no matter what terms I use, no matter what concepts I employ, when I am done I could always in principle add the phrase ‘or at least that’s the way it *appears to me*’. All my descriptions — indeed, all my *thoughts* — are from a point of view, from a perspective — from my point of view, my perspective. (This need not be taken literally as a visual perspective.) Even if I do not explicitly make reference to that point of view or perspective, the possibility of doing so remains. This implicit reference to a point of view or perspective is *inevitable*. Without it, we could have no experience at all.

Why is this important? It is important because it means that all our concepts, and so too all our phenomena, which those concepts describe, carry with them an implicit reference

to ourselves and to our point of view. But, just as in the example of the Gestalt figure, that point of view or perspective is part of the mind's own contribution to the phenomena. (I view the figure from a "light is foreground" point of view, or from a "dark is foreground" point of view.)

Therefore — and here is the crucial move — it is contradictory to try to extend the use of our concepts to describe not just the phenomena but also the "things-in-themselves." It is contradictory to suppose that the phenomena are accurate representations of things-in-themselves. Things-in-themselves are whatever they are with no special reference to us; phenomena, on the other hand, necessarily involve a reference (even if only an implicit one) to ourselves.

The basic idea here is this: Suppose you say "I'm not interested in how things appear to me, from my own idiosyncratic point of view. I want to talk about how they are all by themselves, absolutely, how they are in themselves." Now consider what you are really demanding here. You are saying: I want to discuss how things are apart from any particular point of view or perspective. That is, I want to consider them apart from the very precondition under which alone I can have any experience or any concepts at all. In other words, I want to discuss how things are, in a condition under which — by hypothesis — I cannot discuss them or even think of them. And Kant's response is: What you are demanding is obviously contradictory.

Let's pause and make sure you see the point of this argument. People sometimes think it's a fallacy, and it isn't. People often feel that all this kind of argument shows is that you can't be sure (as though the problem were still just *Descartes'* problem). It's as if the argument were simply:

We always see things from our own point of view. (There's no other way to see things.) And so we are always biased. Now our biases may really be correct; they may accurately represent the way things are. But, because we are inevitably biased, we are never in a position to tell whether that's so or not.

I think this would be all there was to it, if Kant did not have the doctrine of constitution in the background of this whole argument. And here I think the example of the Gestalt figure illustrates the point quite clearly. (It's an illustration, not an argument.)

If you have a view that says it is the mind that determines which is foreground and which is background in that figure, then you cannot consistently go on to say, well, maybe the one really is foreground, quite apart from what the mind does, and the other really is background. You'll have to make up your mind; you can't say both the one and the other.

The point to see is that the doctrine of constitution doesn't just say (for example) the mind determines what looks like foreground and what looks like background in what we see. It says the mind determines what is foreground and what is background in what we

see. It's a theory about what it is to be foreground and background, and where that comes from. And what the theory says is: it comes from the mind. So of course, if that's your theory, then it makes no sense at all to wonder whether those notions apply to things apart from the mind's intervention. That would just amount to wondering whether your theory is right in the first place.

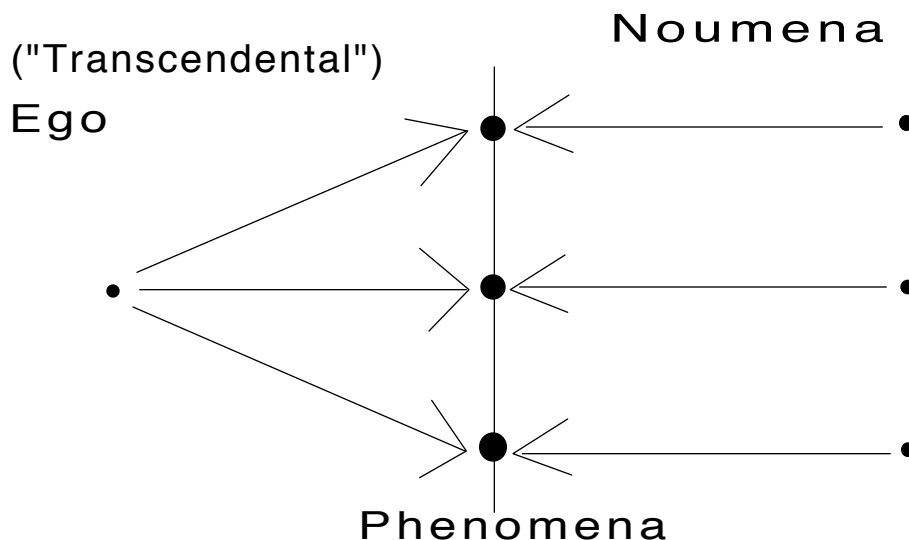
Now the theory of constitution may in fact not be right, but if it is, then Kant's conclusion about the inapplicability of the categories to things-in-themselves seems unavoidable.

I want to stress this now, because when we get into Husserl and Sartre you will probably find yourself wanting to resist this kind of move when you see what its consequences are really going to be. And I want to emphasize now that it's not easy to resist — unless you're simply going to miss the point or distort it. And that's what I want to prevent. (McCulloch, in *Understanding Sartre*, makes this mistake, as near as I can tell.)

Having said that, however, I must add that Kant was quite *certain* that there *were* such mysterious things-in-themselves out there. The whole Kantian picture is that, just as (with the Gestalt figure) what appears to us is a product of two things — a neutral datum caused in us by something in the external world, plus the mind's interpreting activity working on this datum — so too in general, phenomena are the products of raw data, disorganized and uninterpreted, which are caused in us by things-in-themselves, plus the mind's own organizing and interpreting activity.

Thus, there is our contribution, and there is the noumenon's contribution.

So the picture we get with Kant is like this:



Now of course there are obvious problems with this theory. First of all, I have drawn the picture as though there were *several* “things-in-themselves,” several noumena. But, on his own principles, Kant cannot know that. He doesn’t know whether there is one or many of them.

Second, although there is some controversy among Kant scholars about what Kant actually meant, it certainly *appears* as if he is saying that the noumenon causes the raw data of cognition in us. But I thought *causality* was one of those categories that we were forbidden under pain of contradiction to attribute to things-in-themselves.

Third — and there is the same scholarly controversy about Kant’s real meaning here — how can Kant even say that such things-in-themselves exist? “Existence” was another of those categories we cannot apply to things-in-themselves.

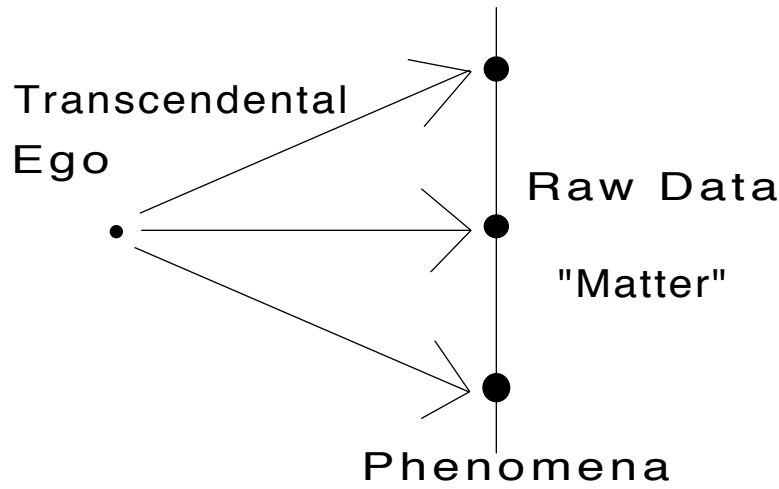
In short, the thing-in-itself became a kind of embarrassment for the followers of Kant. And eventually, people began to realize that if we can’t talk about such a thing-in-itself without contradiction, that’s a pretty good indication that there isn’t any such thing. (That’s what we call a reductio argument, after all.)

And so some post-Kantians came to the conclusion that we don’t need the thing-in-itself, that it is in fact impossible. If we can’t talk about it without contradiction, then we should just shut up about it.

All we really need are the *raw data* of cognition, the raw materials, together with the organizing activity of the mind. A kind of Aristotelian “matter”/“form” setup, with the mind providing the “form.”

We don’t have to ask — and indeed *cannot* ask — what “causes” the data to be there. Kant showed that that question is incoherent.

So now our picture is like this:



But notice something: Aren't we now back to *solipsism*, the doctrine that Descartes tried so hard to avoid? And in fact, the task of avoiding solipsism was what got this whole story going.

Answer: Yes, we are. The conclusion of this line of reasoning is that SOLIPSISM IS CORRECT!

This view has been called "*idealism*." It is the view that all reality is in some sense *mental*. It was a doctrine that had some currency after Kant, especially in Germany: in Fichte and Schelling, and (at least according to one interpretation — probably not the correct one) in Hegel. (We've already mentioned Hegel as part of the tradition against which Nietzsche and others reacted.)

Let's look at the situation a little more closely. It's not quite the situation I described a while ago, when we first talked about the threat of solipsism (in Descartes).

Let's think of the theater model again.

Descartes' problem was: Here I am in my phenomenal theater, looking at the world projected on the screen *from the outside*. (Perhaps it's better not to think of a *movie*-theater but rather of a *shadow*-theater where shadows are cast on the screen by objects on the other side.) How can I be sure that the *projector*, which is on the *outside*, bears any resemblance at all to what I see *on* the screen?

The answer, despite Descartes' best efforts, is that I *can't*. (That's what Kant showed.) Furthermore, Descartes can't *really* even claim there's a projector out there at all.

But now the story has changed. When Kant realized that the mind itself contributes to the phenomena, he in effect moved Descartes' dubious "projector," which caused Descartes so much worry, into the mind. That is, the source of the images on the screen is now inside the theater — and furthermore, it's me, the Ego. (Thus the "shadow"-theater model will no longer work; we're now talking about a *movie*-theater.)

This is the point of the doctrine of *constitution*. The whole phenomenal world I am aware of is simply a story the mind is telling itself — the mind itself is the cause of it. Kant still wanted to have some kind of thing-in-itself outside the theater, but the post-Kantians came to realize that such a thing-in-itself has absolutely no role to play — and is contradictory anyway.

Digression: Remember the raw data Kant was worried about, the uninterpreted data on which the mind imposed an order, the data that by themselves have no structure at all. In terms of our movie-theater model, these raw data are just the screen. By itself, the screen is completely featureless ("uninterpreted"). All content — whatever appears on the screen — comes from the Ego (the "projector"). Again, the screen functions a little like Aristotelian prime matter here.

Review

The "idealist" picture we have arrived at with the post-Kantians may strike you as implausible, as something you're not inclined to believe. So I think it will be useful to review how we got here, so that you will be able to see that, given certain philosophical starting-points, this solipsistic outcome is inevitable.

There are really three main premises that get us to the point we have arrived at:

- (1) We started with the Cartesian ideal notion of philosophy as infallible knowledge (the "quest for certitude"). Hence, as a methodological principle, we agreed to confine ourselves to what we are infallible about — that is, to use Descartes' phrase, to what we are aware of clearly and distinctly. Or, to put it in other terms, we confine ourselves to what is directly given, to the phenomena.

That was the first of Descartes' two principles I described above:
The safe = the phenomena.

- (2) We then added Descartes' identification of these clear and distinct, directly given phenomena with the contents of the mind. This was the second of Descartes' two principles: The phenomena are all mental, mind-dependent.

You put (1) and (2) together, and you get the result
that we can speak infallibly, without risk of error,

about the contents of our own mind, but *not* about anything else.

- (3) We add Kant's view that consciousness itself always and inevitably makes a contribution to the phenomena. It contributes a *perspective* or *point of view* that, by the very nature of the case, implicitly refers to the *observer*, to the *mind*. (The doctrine of Constitution.)

Therefore,

- (4) We get Kant's conclusion, that it is not only *risky* and *fallible* but positively *inconsistent* to try to talk about anything but the phenomena, our *mental representations*. It inevitably leads to absurdity and contradiction to try to talk about things-in-themselves, as opposed to things-the-way-they-appear-to-us.
- (5) And so — although Kant himself resisted this step — we conclude that there *is* no thing-in-itself, that it is contradictory to suppose there is, and that all there is in the mental movie-theater.

The Two Stages of Husserl's Philosophy

I've gone through the story above at such length because the development of Husserl's philosophy over his lifetime shows important connections with the various steps of this story.

In his early philosophy, the period of *Logical Investigations* and *The Idea of Phenomenology*, Husserl had a doctrine that promised to break out of this bind, to avoid the *idealism* that characterized a number of post-Kantians. In order to do this, of course, he had to reject one of the ingredients of the above recipe.

And he did. He rejected step (2).

But, as his philosophy developed, he worked himself more and more into a position that looks like the kind of *idealism* we ended up with above. This "turn" (Husserl's "transcendental turn") happens in his *Ideas* and his *Cartesian Meditations*. We can already see some glimmerings of what is to come in the later sections of *The Idea of Phenomenology*.

Now Husserl's *earlier* philosophy was the one that caught on and that influenced people at the time. They were attracted by the promise of the way out of the seemingly inevitable idealism, with its solipsistic consequences, that we've just gone through.

Therefore, when Husserl himself seemed to be turning more and more toward idealism in his later years, a lot of people felt *betrayed*, and they refused to follow Husserl into what they felt was a reversion to the old errors.

(Digression: Curiously, Husserl seemed remarkably incapable of explaining to his students the reasoning that led him to adopt his later idealism. Husserl himself *resisted* the doctrine, but came to think that certain philosophical considerations made it *unavoidable*. But he was never able to persuade others of this. For example, one of Husserl's more distinguished followers, a Polish philosopher named Roman Ingarden, wrote a book entitled *On the Motives Which Led Husserl to Transcendental Idealism*. The book begins with the astonishing sentence:

I have often asked myself why Husserl, really, headed in the direction of transcendental idealism from the time of his *Ideas* whereas at the time of the *Logical Investigations* he clearly occupied a realist position.

Then he goes on to dig around in Husserl's various writings to piece together a tentative reconstruction of what Husserl's motivations *must* have been.

(I call this opening sentence "astonishing," because the reader inevitably wonders, "Why didn't you just ask Husserl himself?" And of course, Ingarden and the others *must* have done so, but Husserl didn't seem able to explain himself to them very clearly.)

So there developed a *split* in the phenomenological movement. On the one hand, there was Husserl himself and some relatively few of his disciples (mostly, in my opinion, the weaker ones who believed anything Husserl said and didn't or couldn't take the trouble to think things out on their own). And then there were the others, who thought Husserl's "transcendental turn" was a disaster, and refused to go along.

This is the background to Sartre's own *Transcendence of the Ego*, in which he records his own personal split with the later Husserlian philosophy.

With this as background, then, let's look again at *The Idea of Phenomenology*.

The Idea of Phenomenology (Again)

After telling the above story, I think we can see that what goes on in *The Idea of Phenomenology*, at least at the beginning, is more or less *familiar*. It sounds a lot like Descartes.

In Lecture I, Husserl talks about what he calls the "natural attitude," as opposed to the "philosophical attitude." The "natural attitude" — which he will later call the "natural *standpoint*" in *Ideas* — is characterized as a practical, pragmatic attitude, in which the mind is caught up in the demands of external objects, and turns its whole attention to them.

This is the attitude the mind adopts when it is engaged, for instance, in the practical world — *including the world of science*.

When we adopt this attitude, Husserl says, there is a characteristic kind of *procedure* or *method* we employ, and a characteristic *assumption* we implicitly make.

(1) The *method* we employ is *reasoning*. (See p. 13.) We *observe* particular objects, particular facts, and we *generalize* on the basis of them, we form *inductive hypotheses*, general theories. This is *inductive* reasoning. But we also employ *deductive* reasoning, we draw logical consequences from the general theories we construct — in order, for instance, to predict *new* events that we can then test empirically. (This is just the familiar “hypothetico-deductive method” that used to be common in the philosophy of science.)

Of course, Husserl recognizes (p. 14), we sometimes make *mistakes* in doing this. But when we adopt the *natural* standpoint, we aren’t *obsessed* with mistakes the way Descartes seemed to be. We know how to handle them, to correct them as we go along: If it is a matter of *logical* error (in the *deductive* reasoning), then we just go back and do it over *more carefully*.

Sometimes too we make *inductive* errors — we infer general hypotheses from the particular data, and the hypotheses turn out to be *inconsistent* with one another (even though we may not realize this at first), or to be refuted by further observation. When this happens (and when we find it out), once again we go back and do it differently.

Ultimately, what we are *aiming* at in this *method* is a *coherent* theory that fits the observed facts.

Husserl himself describes all this in a very dry, abstract and pompous prose. (Husserl was a *terrible* writer!) But it is not hard to see what he is doing. He is in effect describing *scientific method*, as we usually understand it.

Of course (p. 15), one of the things we might study in this scientific way, from the “natural standpoint,” is the *mind* itself. In that case, we have the science of *psychology*, in which the mind adopts toward itself the same attitude of *disinterested objectivity* that it adopts toward any other object of scientific inquiry.

I mention this because Husserl contrasts psychology very sharply with his own phenomenology. For Husserl, psychology is a *science* that adopts the natural standpoint; phenomenology, as we shall see, is not. Yet the two disciplines are closely parallel. (It is an interesting point to keep track of these people’s attitude toward *psychology*.)

So much for the *method* we employ from the natural standpoint.

(2) There was also, I said, an implicit *assumption* we make when we adopt the natural standpoint. This is the assumption that, as Husserl puts it, “*cognition is possible*.” That is,

we implicitly assume that there is correspondence between our thoughts and what we are thinking about.

Look again at the passage we have already read from p. 15:

Cognition in all of its manifestations is a psychic act; it is the cognition of a cognizing subject [that is, me]. The objects cognized stand over and against the cognition. But how can we be certain of the correspondence between cognition and the object cognized? How can knowledge transcend itself and reach its object reliably?

It is exactly this possibility that we take for granted in the *natural standpoint*, the possibility of *getting at* the objective facts on the basis of which we then go on to construct our theories. The “objective facts” are not given by the theory; they are presupposed by the theory — they are what the theory is trying to explain.

When I am doing biology, for example, I may be worried about getting the facts straight, about controlling the laboratory conditions, and so on. But I am not worried about the general question how — or even whether — the mind can really get at any objective facts at all, about whether the mind is perhaps not a suitable instrument for this kind of inquiry. That’s the sort of question people leave to those “philosophy”-types.

Now of course the problem Husserl is describing here, the problem of the possibility of cognition is exactly the problem that bothered Descartes: How can we get at the realities behind the appearances? How can we break out of our own minds and get to anything beyond?

The natural standpoint takes all this for granted — that we can get at reliable, objective data, and not just at our own subjective biases.

(Digression: You might think that psychology is an exception, that in psychology, since what we are studying is the *mind* and its thoughts and contents, we don’t have to make this implicit assumption of a correspondence between thought and reality. But in fact, Husserl thinks, we do. There we implicitly assume we are able to get at objective and unbiased data about our own minds and their contents, on the basis of which we construct psychological theories — just as much as we assume in astronomy that it is possible to get at accurate and unbiased data about the stars. The fact that the *objects* we study in psychology are so close to us makes no difference — and in fact may make things harder! We still need to make sure somehow that we can get enough “distance” between ourselves and our object to allow an objective approach to it.)

This then is the natural attitude. Let us now contrast it with the “philosophical” attitude. If the natural attitude is characterized by an implicit assumption of the “possibility of cognition,” the philosophical attitude is characterized by the fact that it is there that we worry about precisely that possibility.

Once we *withdraw* ourselves from the business of pragmatic and scientific engagements and begin to *reflect*, the implicit assumption of the natural standpoint becomes a real problem for us. We begin to see that *we should not take it for granted*.

This is exactly what happened to Descartes.

And this realization requires a real change of attitude. It is not a small thing. It requires that we *put away* our pragmatic and scientific interests for a time, to look at their foundations. This is no small matter — it requires a complete shifting of mental perspective. It requires, for instance, a certain amount of leisure and freedom from external pressures. Descartes tells us at the beginning of his *Meditations* the kind of circumstances that are required for this peculiarly *philosophical* kind of attitude:

The present is opportune for my design; I have freed my mind of all kinds of cares; I feel myself, fortunately, distracted by no passions; and I have found a serene retreat in peaceful solitude.

The point of this is that the philosophical attitude and the natural attitude are *mutually exclusive*. You can be in the one or the other, but not both at once. They are incompatible. You either take the possibility of cognition for granted or else you don't.

On pp. 18–19, Husserl tells us that this “philosophical attitude” is what he means by *phenomenology*.

Now I've already told you that phenomenology is a matter of *describing the phenomena*, and that's true. But now we're saying it is a matter of inquiring into *the possibility of cognition*. We shall have to see how these two characterizations of phenomenology fit together in the end. That will be a long story.

Thus (and I'm paraphrasing the text now), the *problem* that phenomenology faces, as Husserl describes it here — that is, phenomenology as the “philosophical attitude” — is the problem of *the possibility of cognition*. That is what it must answer. It must investigate what the natural standpoint takes for granted. Phenomenology is therefore *a theory of knowledge; a critique of natural cognition*.

But how is it going to proceed? It cannot proceed the way the usual *sciences* do, by starting with particular data, and then proceeding to construct general theories to explain those data. It is the very possibility of *getting at* those particular data to begin with that is in question here.

So philosophy — or phenomenology, which is the same thing for Husserl — is not going to be just one science among many. (Some people have a view of philosophy that does think of philosophy this way: Philosophy as simply the most general and broadest of all the sciences.) Philosophy for Husserl is going to require an *entirely new method*. Philosophy, so to speak, goes off in a *completely new dimension*.

Philosophy (still paraphrasing the text) is going to have to try to answer the question of the possibility of knowledge — that is, of the correspondence between our thought and

the objects we are thinking about. It is going to have to investigate *what cognition is* and what it is to be an *object of cognition*, and then try to see what *correspondence*, if any, there is between the two.

Or, as Husserl himself puts it (p. 18), we are going to have to clarify the “essence of cognition and of being an object of cognition.” Just how this is going to work remains to be seen.

What I have just given you is a brief summary of Lecture I.

In effect, Lecture I sets up the problem that is going to be addressed in the rest of the book.

In the remaining lectures, there are three main points I want to focus on:

- (1) The “phenomenological reduction” — the main discussion of which is in Lecture II, but some implications of which are drawn out in Lecture III, pp. 33–35.
- (2) The “eidetic reduction,” or “eidetic *abstraction*” — the main discussion of which is in Lecture III, but some consequences of and observations on which are in Lecture IV.
- (3) The notion of “constitution” — which is obscurely the topic of Lecture V (although you might not recognize it), but is more clearly explained in Husserl’s summary of the lectures in “The Train of Thought in the Lectures.”

In the “Introduction” to the volume, George Nakhnikian discusses something further, called the “transcendental reduction.” I don’t want to treat that separately. Instead, I will discuss it under the heading of *constitution*.

Let’s look at these themes one at a time:

The Phenomenological Reduction

This is what on p. 31 he calls the “epistemological reduction,” and on p. 22, at the beginning of Lecture II, calls the “*epoché*” (ἐποχή = a Stoic term meaning “abstaining,” literally “holding off”). The term ‘phenomenological reduction’ is used on p. 7, in the corresponding passage of “The Train of Thought.” It is also the term used in Husserl’s later *Ideas*.

In all these cases, we are talking about the same thing, although the terminology is a little fluid.

What then *is* the phenomenological reduction?

Basically, it is the adoption of the policy of confining ourselves to what is directly given to us, to the *phenomena*, and *abstaining* (hence ‘*epoché*’) from any judgment about anything further.

It is a *reduction* in the sense that our judgments are confined, narrowed down, “reduced” to the phenomena. (There is a lot of talk about “reductions” in Husserl.)

In other words, once I adopt the phenomenological reduction, I no longer *infer* or *argue* on the basis of the phenomena to something further. I *stay* at the level of phenomena and simply describe them. In effect, this means I *reject the method* of the natural attitude, which — you recall — involved inference. Phenomenology is not an *argumentative* discipline; it is a *descriptive* one.

In effect, the “phenomenological reduction” is just Husserl’s name (or *one* of Husserl’s names) for the *first* of Descartes’ two principles we have already talked about. Husserl accepts it.

He describes this step in various ways in various places:

- (1) In his later *Ideas*, he describes it as “the suspension of the natural standpoint.” That is, as the *adoption* of the “philosophical attitude” he describes in *Lecture I* of *The Idea of Phenomenology*. It is the beginning of philosophy, the “critique of cognition.”
- (2) Also in *Ideas*, Husserl describes this move as the “bracketing” of existence. The term ‘bracketing’ is an important term in the Husserlian lexicon. The idea is that our job is simply to describe the phenomena on our mental movie-screen. It is *not* our job to try to decide whether the phenomena we see represent really existing objects out there. The question of “existence” is set aside, “put in brackets.”

Let me digress here for a moment. Sometimes in the secondary literature you see the claim made that Sartre *rejects* the phenomenological reduction. (For example, David Detmer’s book, *Freedom as a Value*.) On the other hand, there are lots of passages in Sartre that simply don’t make any sense if that is so.

In fact, this whole issue rests on some terminological sloppiness. If by ‘the phenomenological reduction’ we simply mean the resolve to *describe* and not to *infer*, to confine ourselves to what is directly given, then Sartre *accepts* the phenomenological reduction. What he *doesn’t* accept (and here he does break with Husserl) is the view that the phenomenological reduction in this sense requires you to “bracket existence.” Husserl thought the real existence of things is *not* a matter that is directly given to us, whereas Sartre thinks it is. (But, as we shall see, Sartre qualifies that so much that in the end the difference between him and Husserl on this point is not as great as it first appears.)

- (3) In *The Idea of Phenomenology*, Husserl describes the phenomenological reduction as the putting in question of “the

entire world of nature, physical and psychological" (p. 22). Once again, we make no claims about whether the natural world is real or not. It may all be an illusion — but I can describe it anyway. Here too we see the idea we touched on earlier: that psychology is no exception here. It is bracketed along with all the other sciences.

So Husserl starts off the way Descartes does. He accepts the first of Descartes' two principles.

What about Descartes' second principle? That, you recall, was the claim that phenomena — that is, the directly given, the "safe" — are to be identified with my thoughts, with mind-dependent events. Such thoughts are what Husserl calls cogitationes, which is just Latin for "thoughts," mental states, "thinkings."

Well, Husserl agrees with Descartes in part, but he disagrees in part.

He agrees that my thoughts (= cogitationes, singular cogitatio) are indeed directly given. They are included among the phenomena. And so they are fair game for phenomenology.

Nevertheless, he thinks Descartes made two important mistakes:

- (1) The first mistake is connected with the phrase 'and psychological' that I emphasized in the quotation just above.

Descartes had said "I think; therefore, I am." And Descartes thought he knew a fair bit about this "I" or "Ego" the existence of which he was so certain of.

In the end, Husserl thinks, Descartes in effect identified the Ego that we can still talk about after the phenomenological reduction with his own psychological personality or self. In other words, according to Husserl, Descartes thought that after adopting the policy of confining myself to the directly given, I am nevertheless still able to talk with certainty about the facts of my own psychology.

And this is where Husserl thinks Descartes made his first mistake. The "psychological ego" falls to the epoché, as Husserl in effect says on pp. 33–35, in Lecture III (skipping ahead a bit).

What exactly is the point here?

Well, Husserl thinks there is a sense in which the Cartesian *cogito* is correct. There *is* an "I" or "Ego" that I can be quite certain exists, even after adopting the phenomenological reduction. But that Ego is not the same thing as the self or Ego we talk about in psychology. Once again, for Husserl phenomenology is not psychology.

What is this Ego that Husserl thinks we can continue to be certain of after the phenomenological reduction? Well, think of it like this (This is important!):

Go back and think of the movie-theater model again. In this analogy, the phenomena are the pictures on the screen. But when I look at a scene on my mental movie-screen, there is something else I am directly given — in addition to the pictures on the screen.

Consider a John Wayne movie. John Wayne is crossing the Rio Grande with the wagon train. That is the picture on the movie-screen, the phenomenon. But that same scene will look different depending on the position of the camera when it is photographed. It will look one way if the camera is on the far bank, and what you see is John Wayne and the wagon train coming toward you. It will look another way if the camera is on the near side, so that John Wayne and the wagon train are receding. It will look different yet if it is photographed from the side, so that what you see is the wagon train passing across the screen in front of you. It will look different yet if it is photographed from above. And so on.

Furthermore, these differences in the position of the camera are things you can tell right off, without having to infer or argue at all. You may not know enough about the details of the area's geography to be able to describe the position of the camera in terms of map-coordinates or landmarks, but you can tell right off that "the camera is over here, and now it is moving over there," and so on. This is something I can be absolutely sure of.

This "position of the camera" is a good model for what Husserl is thinking of when he talks about the "Ego" that is left over after the phenomenological reduction. Let us use the term 'phenomenological Ego' for this kind of Ego. (This is my term, for convenience only, not Husserl's term.)

In slightly less metaphorical terms, this "phenomenological ego" can be regarded as simply a kind of point of view, a perspective on the phenomena. That point of view is not itself a phenomenon (the camera itself never appears on the screen — barring various sorts of "trick"-movies for the moment). But it is directly given to us, and therefore something I can continue to be certain of even after adopting the policy of confining myself to what is directly given. (Recall that I told you a while back that the Ego was going to be a special case. This is just as true for Husserl as it was for Descartes, although the details will perhaps be different.)

Notice something important here: We've introduced an important distinction, and with it we have to refine our terminology. Earlier, we said that the "directly given" = "the phenomena." But now we're saying, "The phenomenon here is John Wayne and the wagon train, but there is something else directly given too." So we have to make a distinction. We don't yet have the terminology to make the distinction clearly and non-metaphorically, but in terms of our movie-theater analogy, we can say that the phenomena are the images on the screen. They are "directly given," yes, but we now know something else is "directly given" too — so that we don't have to infer to know it — namely, in terms of the analogy, the "position of the camera."

So any complete discussion of what is directly given will not only have to describe the events going on on my mental movie-screen, but also contain a reference to a perspective or point of view (the "eye of the camera"). This perspective is not itself a phenomenon, but we must take account of it in any complete description of the phenomena, since the phenomena look different from different perspectives.

In effect, this is just Kant's point: "The 'I think' must be capable of accompanying all our representations." And this "perspective," this "vantage point," is the Ego that Husserl says remains certain to us even after the phenomenological reduction. It is what we have called the *phenomenological Ego*.

This phenomenological Ego is a bare "vantage point," an empty "point of view." But what else can we say about it? It is *individual*, in the sense that different movies involve different points of view, and in the sense that one point of view *precludes* all other points of view (at a given time) within a single movie. So Husserl will allow us to talk about an *individual* Ego after the phenomenological reduction, and we can be absolutely certain of its existence.

But this phenomenological Ego is *not* "personal," in the sense that we cannot talk with phenomenological certainty about this Ego as the seat of our psychological drives and impulses. The Ego in *that* sense (as the *psyche*) is an *object* that we can describe phenomenologically, but the real existence of which we "bracket."

Thus, as Husserl puts it in Lecture III, after the phenomenological reduction we can speak with certainty about *cogitationes* (= thoughts), but not about *my cogitationes*, if by that we mean to refer to a *personality* or *psyche*.

It is worth taking the trouble to get these points and distinctions straight now, because we will have occasion to return to them later on.

Let me summarize them:

- (i) Husserl agrees with Descartes on the policy of confining ourselves to what is directly given. (This is what we call the "phenomenological reduction.")
- (ii) For the most part this means confining ourselves to the *phenomena*. But Husserl agrees with Descartes that the Ego is a kind of special, exceptional case. For Husserl (what Descartes thought on this precise point is anyone's guess), it is directly given, and yet is not a phenomenon (not a picture on the screen).
- (iii) But Husserl *disagrees* with Descartes about what kind of Ego this is — in other words, about the Ego we can say *exists* with absolute certainty. Descartes (according to Husserl) thought it was the *psychological Ego*. Husserl thinks that is wrong. The *psychological Ego* for Husserl is an *object*, the existence of which is "bracketed" along with the existence of all the other objects of science and of our day to day experience.
- (iv) For Husserl, the Ego we can be certain exists is just a bare vantage point, a perspective — which we have called the "phenomenological Ego." (Again, that is *my* term.) It is *individual*, in the sense that different perspectives mean different Egos and

vice versa. But it is not *personal*, in the sense that this “phenomenological Ego” is just a kind of *geometrical point*. There’s nothing back there. We don’t want to think of this abstract “Ego”-point as endowed with a personality, drives, urges, wishes, hopes.

Note: This “phenomenological Ego,” as we call it, is not yet what we will later on call the *Transcendental Ego*. That’s a different story, and at this stage in his philosophical thinking, Husserl had not yet adopted the theory of the Transcendental Ego. At this stage, Husserl is thinking of the Ego as simply an *observer* of phenomena; it isn’t yet thought of as *contributing* to them in any way.

This then is where Husserl thinks Descartes made his first mistake: identifying the “phenomenological Ego” with the “psychological Ego.”

But he also thinks Descartes made a *second* mistake:

- (2) Descartes was confused over just what *is* and what is *not* a phenomenon for us.

Both agree that we are going to confine ourselves to what is directly given. Both agree that, with the special exception of the Ego, what is directly given to us are the *phenomena*. But what all do they include?

As Husserl puts it, what is it that is given to us “*with evidence*?” (‘Evidence’ in Husserl does not mean hints and clues. It means “*self-giveness*” — being “directly given.”)

Apart from the special case of the Ego, we have seen that *cogitationes* are directly given to us. But is that all? Descartes thinks it is (that’s his *second principle*, but Husserl thinks not.

So, in effect, Husserl agrees that: *cogitatio* → directly given. But does it go the other way around? Does: directly given → *cogitatio*? (Apart from the Ego, which as I said is always treated as something special.)

This is what is *really* going on in Husserl’s rather obscure discussion of the two senses of the terms ‘immanence’ and ‘transcendence’ in Lecture II. (Note: Husserl tends to drop this vocabulary later on — although it’s still there in *Ideas*, with some differences.)

Basically, the terms ‘immanence’ and ‘transcendence’ etymologically mean roughly “inside” and “beyond,” respectively. I have not so far put the issue in the technical vocabulary of immanence and transcendence, but rather in terms of what is *in* the mind as opposed to what is *outside* the mind. And of course Descartes didn’t use the ‘immanence’/‘transcendence’ talk either. But Husserl does. And we must understand that what is *really* at stake in all this talk is just whether we are going to accept the *second* of Descartes’ two principles we distinguished earlier:

that the *phenomena* are all “mental contents,” in the sense of being *mind-dependent* — things like sense-impressions, pains, etc. (In short, *cogitationes*.)

Husserl, recall, puts the whole question of *The Idea of Phenomenology* in terms of “transcendence” — recall the question in Lecture I (p. 15) in the passage quoted earlier:

How can cognition *transcend itself* and reach its object reliably?

And later on, in Lecture II, he tells us the phenomenological reduction gives us a *methodological principle* (p. 29 — this is not exactly a quotation):

Nothing “transcendent” can be used as a presupposition for our investigation. We must confine ourselves to what is “immanent” in cognition.

This is the way Husserl puts what I have described in other terms.

Now Husserl tells us in a *crucial* passage on pp. 27–28 that the terms ‘transcendence’ and ‘immanence’ are *ambiguous*.

- (a) On the one hand, there is ‘immanence’ in the sense of being a *mental ingredient*, a *mental content*, being *mind-dependent*. This is what Husserl calls “*real (= reell) immanence*,” sometimes translated as “*genuine immanence*.”

In this sense, something is *immanent* in another thing if it is “really in there,” if it entirely “inheres” in it — as, for instance, a *part does in its whole*. And something is “immanent” in *cognition* in this sense if it is a real *part* of that cognition, or a real *characteristic* of it. (This sense of immanence and transcendence is also the *hard* and obscure one, and we will have to refine it as we go along.)

For example, if I think about the planet Mars for ten minutes, then that thought’s *duration of ten minutes* is “immanent” in that act of thinking — it is a real feature, a real characteristic, of that act of thinking. And if I think about Mars *real hard*, then the *intensity* of my thinking is “immanent” in my act of thinking — it is a real characteristic of it.

In this sense, of course, what is “immanent” is *mind-dependent*.

The correlative opposite, ‘transcendence’, in this first sense, means: *not wholly contained in the mind, not really inhering in, not really a characteristic of, the mental act*.

So in this first sense, the pair of terms ‘immanence’/‘transcendence’ means roughly “in the mind”/“outside the mind.”

- (b) On the other hand, in the second sense Husserl distinguishes, ‘immanence’ means being immediately, or directly, given — “self-given,” as he puts it.

Something is “immanent” in this sense if it is present to the mind in person, in itself — rather than simply being represented there.

For example, if I think about the planet Mars again, and if we have a theory (which Husserl doesn’t, in fact, although Descartes does — but this is only an example) according to which what I have in my mind is some sort of concept or sense image of Mars — a representation, which is what I am directly aware of, on the basis of which I then infer certain things about the real Mars — if that is the situation, then the planet Mars is not “immanent” in my thinking in this sense. But the concept or sense-image is.

By contrast, something is “transcendent” in this second sense if it is not “immanent” in the second sense — that is, if it is not present to the mind directly and immediately, in person, but only at best represented there, so that I have to make an inference to get to it. I have to infer from what is directly present to my mind (the concept or sense image) to what is not (the planet Mars, say — if that is your theory).

The criterion or test of whether something is immanent or transcendent in this second sense is: Is an inference required before I can make a claim about this thing? If so, then it is not “immanent” but “transcendent.” If not, it is not “transcendent” but “immanent.”

These terminological points are confusing, but try to get them straight. They are important. And I think they will clear up some as we proceed.

Now what is the point of making these distinctions? Well, Husserl thinks it is a question that has to be asked: Whether these two senses of ‘immanence’ and ‘transcendence’ amount to the same. In other words, can something be “immanent” in one sense and yet “transcendent” in the other? If so, then we must keep these senses carefully separate, and anyone who confuses them will get into trouble. And that is exactly what he thinks happened with Descartes, and this is the second of the two ways in which Husserl thought Descartes went wrong. (The first one, recall, was the business about what kind of “Ego” was left after the phenomenological reduction.)

Husserl’s way of setting up this second point is pretty perverse. After all, this rather artificial distinction between the two senses of ‘immanence’ and ‘transcendence’ is a piece of Husserlian jargon. Descartes didn’t talk in these terms at all. He simply didn’t use the vocabulary of transcendence and immanence, so that it doesn’t make a lot of sense to say that he made a mistake by confusing two meanings of words he didn’t even use, and furthermore that he confused two meanings that no one but Husserl ever assigned to those words anyway!

But we don’t have to follow Husserl’s peculiar jargon here to see the point of what he is saying. Nothing rests on the words here. If you think about it a bit, it’s easy to see that what Husserl is really asking is just this:

Can something be “immanent” in sense (b) — directly present to the mind, present in person and not merely by representation or inference — and yet be “transcendent” in sense (a) — not mind-dependent in the sense of being a mental content or real part or real characteristic of the act of thinking?

In other words, does the mind ever come into direct contact with anything but its own mind-dependent ideas and impressions and their genuinely immanent constituents? That’s the real issue: Are we going to accept Descartes’ representational theory of cognition?

It is an important question because, you will recall, Descartes’ second principle that we distinguished above said that everything directly present to the mind — not only the Ego itself but also all phenomena — is mental, mind-dependent.

And it was that principle which, combined with his first principle (the phenomenological reduction), gave rise to Descartes’ insoluble problem of avoiding the possibility of solipsism. Thus, if this second principle should turn out to be a mistake, as Husserl thinks, then Descartes’ problem will have vanished!

Well, what about it? Are there any phenomena (things directly present to the mind in cognition) that are not themselves real parts or characteristics of mental acts?

Husserl cannot of course just take it for granted that there are such phenomena. Phenomenology is not supposed to proceed like that; it is supposed to be a “presuppositionless” science.

But neither can he argue that there are such phenomena. Phenomenology, remember, is not an argumentative science either — all we are allowed to do is describe and sort things out descriptively.

So if Husserl is going to answer his question in the affirmative (and say there are phenomena that are immanent in the one sense and transcendent in the other), and thereby break out of Descartes’ bind, he is going to have to examine his phenomena and see if can discover any that are like this.

Problem: How would you know when you had found one? Well, wait and see.

The Eidetic Reduction

Husserl finds such phenomena as the result of the second of the three main things I said I wanted to focus on in *The Idea of Phenomenology*: the eidetic reduction. This is discussed in Lectures III and IV.

Here is what he says (p. 40):

But can it be that absolute self-evidence, self-givenness in “seeing,” is realized only in particular mental processes and their particular abstract

aspects and parts, i.e., only in the “seeing” grasp of the here and now?
Would there not have to be a “seeing” grasp of other data as absolute data,
e.g., *universals*, in such a way that were a universal to attain self-evident
givenness within “seeing,” any doubt about it would then be absurd?
[Emphasis added.]

This is simply an extremely convoluted way of asking:

Is it the case that what is *directly given* is no more than the *particular*
thought, confined to a momentary instant, and the various particular real
parts and features *of* that momentary thought — and *that’s all*?

Husserl thinks *no*. Later on p. 40, he says:

To view the matter more precisely, in the subject-predicate judgments
which we make concerning them [that is, concerning the momentary
thoughts — for instance, “This, right here and now, is redness], *we have
already gone beyond them*.

There are two parts to the claim in this discussion:

- (i) *Universals* are among the things that are directly given to us —
among the phenomena. By ‘universal’ here, I mean things that can
recur and be *recognized* again as having been there before. In our
example, “redness” is a universal. I can say “This, right here and
now, is redness,” and then say “Here it is again.”
- (ii) The universal cannot be reduced to any one given momentary
phenomenon, *or to any (finite) collection of them*.

What is given to me here is *redness*, and *redness* is something that *goes beyond* any given
act of thinking about it or being aware of it. It *goes beyond* (= transcends) any given any
act of thinking about it, because I can think about it *again*, in a *new* and *second* act, and
there it is again. And in fact, no matter how many times I think about or am aware of
redness, I could in principle think about it again, so that it would then be the object of yet
another mental act.

Redness is thus never *exhausted* by my acts of thinking about it — either any *single* act
or any *combination* of acts. (Note: We are talking about a combination of mental acts that
I actually perform, not some infinite collection of mental acts I *might*, potentially
perform.) It can always come back again. It thus *goes beyond* — which is just the
etymological meaning of ‘transcends’ — any single mental act or any combination of
mental acts.

In that sense, redness is *not* a real *part* or a real *characteristic* of any single mental act or
any combination of mental acts; it is not *confined* to them.

Thus, what we have is something that is transcendent in the first sense of the term distinguished earlier (it goes beyond mental acts, it is not confined to them), but immanent in the second sense (we do not have to infer redness, it is immediately and directly present to the mind). And that is exactly what we were looking for.

Thus, Husserl says, we have broken out of Descartes' bind. We can be absolutely sure of certain objects that are not "mental" in the sense of being confined to our mental acts — namely, of universals. There may be other such "transcendent" things too that are directly given to the mind, so that we can be absolutely sure of them. That remains to be seen. But, in any event, universals are like that.

Let's pause for a moment at this point to make some observations and ask some questions. (There will be six such points.)

First of all, this is an extremely important move in Husserl. It is his answer to the question that threatened Descartes: how to avoid the possibility of solipsism, how to break out of the confines of the mind and establish that the Ego can get in touch with something besides itself and its own products.

But second, you may think this whole passage in Husserl sounds suspiciously like an argument, which is a little odd for a philosopher who says that the job of philosophy is not to argue but to describe. So isn't Husserl violating his own procedural strictures here?

It sounds as if what he is saying is:

- (1) Universals are directly given to us.
- (2) Universals are "genuinely" transcendent to acts of consciousness.
- (3) Therefore, we have broken out of Descartes' bind and solved his problem.

But that's not really what is going on at all. If you look carefully, what Husserl is really doing is saying: Look, see? Universals are directly given. And now look again, see? Universals are "genuinely" transcendent, aren't they? And finally, look once again. We've found our way out of Descartes' problem, haven't we?

You may think there's no real difference here, but there is. The whole point of Husserl's method is to get us to "see" certain things and facts directly, to "intuit" them. Any device that will help us do this is fair game — even arguments and inferences. So when we say that phenomenology does not argue, it only describes, we have to be careful. You will see what appear to be arguments all throughout the phenomenological literature. But the point is, while they can help produce insight, they can never serve as a substitute for it. (In effect, this is exactly what Descartes and Locke had in mind when they regarded demonstrations as simply chains of intuitions.)

But third, there is still something funny about all this. We have ended up saying that universals are transcendent to acts of consciousness in the first sense of 'transcendent' ("real" or "genuine" transcendence). They are not really or genuinely "immanent." But

earlier, when I was introducing the various senses of ‘transcendence’/‘immanence’, I said that to be “immanent” in this first sense (really or genuinely immanent) is a matter of being a mental ingredient, a mental content, of being mind-dependent. And I gave the example of thinking about the planet Mars for ten minutes; the duration of ten minutes was, I said, really immanent in that act of thinking. Or the intensity of my thought is really immanent in it if I think about Mars really hard.

And of course the fact of enduring for ten minutes, and a particular intensity of thought are universals in exactly the sense we have just been talking about: they can recur. Lots of things can endure for ten minutes, and lots of things — even lots of thoughts — can have that same degree of intensity. So the examples I gave of genuine immanence turn out to be universals, which in the present passage Husserl triumphantly declares are genuinely transcendent. So which is it? We can’t have it both ways.

I think that’s right, and I suspect what it means is that there is something wrong with my examples. But I am not sure how to fix them. If you look carefully at what Husserl says in *The Idea of Phenomenology*, he never (so far as I can find) gives any good example of genuine immanence. On p. 27, for instance, he says

the cognitive act, the *cogitatio*, has genuine abstract parts genuinely constituting it.

But he doesn’t say what those parts are. Again, on p. 40 (in the passage we are presently discussing), he says

But can it be that absolute self-evidence, self-givenness in “seeing,” is realized only in particular [that is, not universal] mental processes and their particular abstract aspects and parts ...?

But he still doesn’t say what those particular “abstract aspects and parts” might be.

So we have a lingering question for Husserl: Is there anything genuinely immanent in mental acts, and if so what? (This difficulty is perhaps one of the reasons he tends to stop talking in these terms in his later writings.)

Nevertheless, we can sharpen our understanding of this terminology by thinking about the present passage. Universals are said to be genuinely transcendent, and not genuinely immanent. Why? Because they can be repeated indefinitely. They cannot be confined to any one mental act or any series of mental acts.

So it would appear then that the genuinely immanent cannot be indefinitely repeated like this. Whatever it is, it has to be particular — it is exhausted in one single mental act, or at most in some finite series of mental acts. (Finite because, remember, we are talking about acts I actually perform.)

See p. 44:

Every genuine (*reell*) constituent of the cognitive phenomenon, this phenomenological particular, is also a particular...

This is important because it means we have to be careful about just casually identifying the notions of genuine immanence/transcendence with our loose talk about being “mind-dependent”/“mind-independent,” or “inside the mind”/“outside the mind.” The property enduring for ten minutes is “inside the mind” in the sense that my thinking about Mars really does have that property. That property is really in there. But it’s a universal, we now see, and so is not genuinely immanent but transcendent.

Similarly, the property being a mental act is obviously “mind-dependent,” in the sense that that property will only be found in the mental realm. Nevertheless, it’s a universal and so genuinely transcendent. Although it is “mind-dependent,” it does not depend on any particular mental act or any finite series of them. It could always in principle recur outside any particular act or series of acts.

So we can’t simply identify the genuinely immanent with the “mind-dependent” or with what is “inside the mind,” although those locutions are all right as rough approximations. In the end, when Husserl says that one thing is “genuinely immanent” in another, he means it is confined there. It’s like putting something in a box. It can’t both be wholly contained in the box and also found outside the box.

But once again, we still have the lingering question: What is it, if anything, that is genuinely immanent in a mental act? (Perhaps the instantiations of the universal?)

Fourth, note that the phenomenological reduction is still in force in the passage we are now discussing. Husserl is rejecting Descartes’ second principle, not his first one. Husserl is still confining himself to what is directly given to the mind. And as a result, there are still lots of things he cannot affirm.

For example, he still cannot say — any more than Descartes was willing to say — whether the oar is really bent or not, or even whether the oar really exists.

Those claims would require an inference beyond what is directly given to the mind, and we are not allowing ourselves to make inferences like that. That is exactly where the possibility of error arises.

In the end, Husserl is just not very interested in these inferences, or in the question of what “really” is going on out there. He just doesn’t think that question is very important — and he certainly thinks Descartes’ way of framing that question (in terms of whether there is a reality out there behind the appearances) is just wrongheaded. (We’ll see more on this later.)

Husserl is not interested in whether the oar is really bent in the water or not, but rather in what it is for an oar to be bent in water. He is interested, so to say, in “oar-being-bent-in-water-ness.” And he can get that from a mere description of the appearances.

Husserl is thus interested in essence, not so much in existence.

In the end, Husserl thinks I can do pretty much all I ever wanted to at the level of essences anyway:

I can describe the visual phenomenon of the bent oar.

I can describe the tactile phenomenon of the straight oar.

I can describe the apparent tension and opposition we feel between these two phenomena. I can describe how they relate to other phenomena. I can describe how they fit into scientific theory (without actually committing myself to that theory).

And so on. In short, I can answer all the important questions. What more do I want? For Husserl, the question of existence is not nearly so important as the question of essences.³

And essences are what we describe in terms of universals, which we now know we can have an immediate and direct knowledge of, even though they transcend our mental acts.

Now this process by which we look at a particular event (for example, a particular sensory event), and see in it the universal that is present there, is what Husserl calls EIDETIC ABSTRACTION or THE EIDETIC REDUCTION.

The term ‘eidetic’ comes from Greek εἶδος (‘*eidos*’ = idea). This is the term Plato used for the famous Platonic “Ideas” or Platonic “Forms.” Husserl uses the term ‘eidetic’ to suggest a connection between his doctrine and Platonism.

The connection is certainly there, but it is important to realize that Husserl is not a Platonist. We’ll see why a little later on.

Let me repeat: this eidetic reduction is a crucial move for Husserl. It is what shows him that Descartes’ second principle is wrong. It gives him the general notions in terms of which to describe phenomena.

There is yet another, fifth point that may be bothering you. Suppose I reflect on a case of imagining a unicorn, and I say with phenomenological certainty, “This is an act of imagining.” (I certainly *can* say that, according to Husserl — it is “directly given” to me.)

Now, notice that the predicate of that judgment is a universal term or concept. “Act of imagining” is a universal notion, since there can be lots and lots of acts of imagining. And, as Husserl says on p. 40,

in the subject-predicate judgments which we make concerning them [i.e., concerning my momentary, individual *cogitationes*], *we have already gone beyond them*.

³ Note: We don’t want a philosophy that tells us that we *can* infallibly decide whether the oar is *really* bent, or even whether it *really* exists. Because in actual fact we can’t infallibly tell about such things. So if our philosophy says we can, then it’s wrong. Phenomenology is not magic, after all.

In this case, it's the *predicate* that takes us "beyond" the *cogitatio*, since it's a universal notion.

But — and this is what may still bother you — it's hard to see how that universal ("act of imagining") is ever going to break out of the confines of the mind, since anything characterized as an act of imagining is going to *have* to be a mental act.

And in fact, if you think about Husserl's whole strategy here — starting with particular, individual *cogitationes* and discovering the universals *they* exhibit — it is hard to see how we are ever going to find anything *that way* that will inevitably lead us beyond the realm of the mental.

But that's not quite right. Let's ask ourselves: What is that universal characteristic being an act of imagination, for example? What is it to be that universal?

Well, since it's a universal, it's not confined to any one act of imagining. It can always recur, it can always be repeated again. Even if it doesn't recur in fact — if, by some freak turn of events, I only imagine once in my whole life — nevertheless, that universal would still in principle be repeatable. In other words, the reality of my one act of imagining does not exhaust what it is to be that universal.

And for that matter, the same point holds no matter how many acts of imagining I perform. What it is to be that universal is not exhausted even by the whole series of my imaginings.

Now, you may say, what if we think of this universal as just the infinite sequence, a kind of infinite summation of all my real and possible acts of imagining? That's all right, you can say that if you want. And Husserl himself sometimes talks about objects as infinite series of real and potential phenomena (as Sartre discusses in the "Introduction" to *Being and Nothingness*). But that doesn't affect the present point. The fact is, my mind doesn't engage in all the real and possible acts of imagining I might have; it only engages in the real ones. And what it is to be that universal is never exhausted by those, no matter how many they are.

So while in a sense the universal being an act of imagination may not lead us beyond the realm of the mental (since imagination is necessarily mental), so that the objection has some basis, that universal certainly does lead me outside my mind in the sense just described: what it is to be that universal cannot be exhausted by anything that actually goes on in my mind.

And for that matter, even what we have just conceded to the objection (that it won't lead me outside the realm of the mental in general) will fail for certain other universals I can find exhibited in my *cogitationes*. If I imagine a unicorn for a long time, let's say, the universal enduring for a long time is "evident" to me. And that universal is something that need not characterize only mental things; in principle it could belong to all sorts of other things too. All the less, then, is it "exhausted" by my mental acts.

Let me make one last point on all this. The way Husserl talks, it may sound as though the only transcendent things we have any reason to think are given with "evidence" are

universals I find exhibited by my individual *cogitationes*. But that's not right. As we'll soon see, all sorts of individuals can be directly given, with "evidence." It's just a question of *looking and seeing*.

But if that's so, then why does Husserl take this laborious detour, by focusing on *cogitationes*, and concentrating on the universals found in them?

Well, it's because we are coming at this whole business from a basically Cartesian starting point. We start with Descartes' problem, and try to work our way out of it without ending up in transcendental idealism. That was the historical context in which Husserl and others found themselves. So even though Husserl in the end thinks that lots of things are given to us directly, he realizes that we are probably not going to be willing to see that right away, and are probably going to have to be maneuvered into realizing the point. So he starts with something even the Cartesian would admit is given with "evidence," namely, *cogitationes*, and talks us through from there.

So don't think that *cogitationes* play any special or privileged role in Husserl's philosophy. They don't (at least not at this point in his development). But they do in Descartes' philosophy, and so they do in Husserl's book too.

The Theory of Intentionality

We've looked at two of the three things I wanted you to get out of *The Idea of Phenomenology*: the phenomenological reduction and eidetic abstraction.

The third is the notion of constitution.

But before I do that, I want to discuss the notion of intentionality, which will lead us into the theory of constitution.

The notion of intentionality is not something that Husserl discusses very directly in *The Idea of Phenomenology*. He does mention it briefly, at the very beginning of Lecture IV.

But if he doesn't discuss it explicitly here, it's not because it isn't important to him. It is. He had already discussed it at some length in *Logical Investigations*, vol. 2. The doctrine was one he got from Franz Brentano in Vienna.

Basically, the theory of intentionality can be summed up in the claim:

Every act of consciousness is always consciousness of something.

We never are just "conscious"; we are always conscious of something. We never just perceive; we always perceive something — whether it's real or not doesn't make any difference. We never just imagine; we always imagine something. We never just fear; we

always fear something. Even vague and free-floating fears have objects — vague and free-floating ones.⁴

Thus, every act of consciousness “aims at,” “reaches out toward” — and in that sense “intends” (= tends toward) — an object.

Left just like that, the claim looks pretty trivial and innocuous. But in fact there are several other, more substantive claims built into this theory of intentionality:

- (a) The relation of intentionality is irreflexive. That is to say, no act of consciousness is ever conscious of itself. The object of any act of consciousness is always something else.

If I perceive a table, the table is the object of my consciousness. My perceiving is not the object of consciousness. Of course, I can always make my perceiving an act of consciousness; I can always reflect on it. But that requires a second act.

Think about our movie-theater model. If I am “caught up” in the movie, I’m thinking about the events in the story. I’m not thinking about my watching the movie. In fact, if I do stop to think about my watching the movie, to that extent I have to “pull myself out” of the movie. There is an almost physical sense of pulling myself out of the one type of consciousness and putting myself into the other kind.

When we say that the relation of consciousness to its object is an irreflexive relation, do not confuse this with notion of reflective consciousness as just described. The latter is reflective in the sense that one act of consciousness takes another act of consciousness as its object. But no act of consciousness is reflexive in the sense of taking itself as its object. The terms are similar, but do not be confused by them.

(The terms ‘reflexive’ and ‘irreflexive’ as just described are taken from the logic of relations. They are not pieces of special phenomenological vocabulary in either Husserl or Sartre.)

Husserl himself is perhaps not altogether of one mind on this claim of irreflexivity. Some things he says in *Logical Investigations* seem to imply it (see *Logical Investigations*, vol. 2, p. 559 — I’ll quote it in a moment). But in *The Idea of Phenomenology*, p. 24, he tells us that

...every mental process, while being enacted, can be reflected on.

That of course doesn’t strictly imply a denial of ir-reflexivity. Perhaps we simply have two acts at once, the one reflecting on the other. It’s not clear to me whether this is what Husserl means here or not.

⁴ Gregory McCulloch, *Using Sartre: An Analytical Introduction to Early Sartrean Themes*, (New York: Routledge, 1994), Ch. 1, seems to take the notion of vague, free-floating emotions as somehow a *problem* for the theory of intentionality.

So there is some unclarity here on Husserl's part. But Sartre is quite clear about it. He strongly endorses the irreflexivity of consciousness.

As a second claim built into the thesis of intentionality, we have:

- (b) The object of an act of consciousness is transcendent to that act in the first sense of 'transcendence' we distinguished earlier — that is, it is not a real mind-dependent part or characteristic wholly contained in and confined to that one act of consciousness.

That is, the object is "genuinely" transcendent to the act.

Thus the theory of intentionality gets us out of Descartes' bind.

This second claim can perhaps be regarded as a kind of loose corollary of the first, although it might be a difficult task to say exactly how the one follows from the other. (And of course, we aren't interested in arguing from the one claim to the other anyway.)

We've already seen this second claim in the case of universals, in our discussion of eidetic abstraction. There we saw that universals are objects of consciousness and are "genuinely transcendent" to those acts.

But what about particulars or individuals as objects of consciousness? Well, the same thing holds there. They too are genuinely transcendent to any act of being conscious of them.

Now let's pause and head off some potential misunderstandings.

First of all, how did we get to the point of saying that the individuals or particulars I am conscious of are genuinely transcendent? I thought we agreed with Descartes that my particular thoughts or mental acts — my cogitationes, to use Husserl's lingo — are directly given to the mind and so are "immanent" in the second sense of the term, but that they are also "genuinely immanent." It was only when we discovered the realm of universals by eidetic abstraction that we discovered cases where these two senses of 'immanence'/'transcendence' diverge, we discovered objects of consciousness that are directly given (and so "immanent" in the second sense of the term) and yet transcend my act of consciousness by not being genuinely immanent in it.

But no. That's not what we agreed. At the time, it may be what we thought we were agreeing to, but if so we were confused. We are now in a position to see more clearly just we do and what we do not have to commit ourselves to.

We agreed — and still agree — with Descartes that my particular cogitationes are directly given to me, and so "immanent" in the second sense of the term. Now Descartes may have thought that they were also "genuinely immanent," and we may have thought so too, but nothing we said then committed us to that claim.

The point then of Husserl's talk about how universals are genuinely transcendent is not to suggest that individuals or particulars are somehow not genuinely transcendent. The point is rather that, in the case of individuals or particulars, it is easy to get confused about this.

But this confusion is much harder to make in the case of universals, once we realize that universals cannot be exhausted by any particular act or series of acts of thinking about them.

The significance of the eidetic reduction, therefore, is not that it is only there that consciousness reaches out to a genuinely transcendent object. That's not so, as we see in this second claim built into the thesis of intentionality (at least for Sartre). The significance of the eidetic reduction is rather that it shows us a case where we can see the point clearly.

But you may still have an objection. If the object of consciousness is always genuinely transcendent to the act itself, then what about illusions or hallucinations? Historically, one of the main reasons for postulating sense-data (mental contents) as objects of consciousness is to account for illusions. If what I'm seeing — the object of my consciousness — is illusory, then it obviously isn't "out there." It must therefore be "all in my mind." This was in fact one of the main motivations behind Descartes' representational theory of consciousness.

If the bent oar isn't really "out there," then it must be "in my mind."

No, Husserl says — that's a mistake. When I see the bent oar, it is the oar I see, and oars are not the kinds of things you can put inside a mind or consciousness. Oars are made of wood and metal and paint — in other words, of matter-stuff, not mind-stuff. (I can photograph the bent oar.)

And so Husserl just accepts the inevitable result of this: If the bent oar doesn't really exist "out there" (since the "real" oar is straight), then it doesn't exist at all. It in no way follows that it does really exist "in my mind" — as if we had to find someplace for it to exist!

This leads us to the third claim built into the thesis of intentionality (and this is very important):

- (c) The intentional object need not exist. I can imagine all kinds of things that don't exist, I can think of things that don't exist, I can certainly fear things that don't exist. All of these are ways of being conscious of things that don't exist.

In other words, Husserl is saying that the following kind of inference is a fallacy:

I am conscious of x; therefore, x exists.

It is this (fallacious) inference that is behind the tendency for us to say that things that don't exist "out there" in reality nevertheless do exist "in my mind," to find someplace for them to be. For Husserl, that is simply a mistake, as we saw just a moment ago with the bent oar.

(Note: Claim (c) is why Husserl can insist that, despite his maintaining that universal essences are genuinely transcendent but directly given to the mind, he is not a Platonist. He is not committed to saying that these directly given universals exist at all.)

Here is a passage from *Logical Investigations* that forcefully illustrates the point (*Logical Investigations*, vol. 2 of the English translation, pp. 558–559):

If I have an idea of the god Jupiter, this god is my presented object, he is ‘immanently present’ in my act [in the second sense: he is “directly given”], he has ‘mental inexistence’ in the latter, or whatever expression we may use to disguise our true meaning. [Husserl is hardly one to complain about this. Is this an attempt to be funny?] I have an idea of the god Jupiter: this means that I have a certain presentative experience, the presentation-of-the-god-Jupiter is realized in my consciousness. This intentional experience may be dismembered as once chooses in descriptive analysis, but the god Jupiter naturally will not be found in it. The ‘immanent’, ‘mental object’ is not therefore part of the descriptive or real make-up of the experience, it is in truth not really immanent or mental. [That is, it is not “genuinely” immanent.] But it also does not exist extramentally, it does not exist at all. This does not prevent our-idea-of-the-god-Jupiter from being actual, a particular sort of experience or particular mode of mindedness such that he who experiences it may rightly say that the mythical king of the gods is present to him, concerning whom there are such and such stories. If, however, the intended object exists, nothing becomes phenomenologically different. It makes no essential difference to an object presented and given to consciousness whether it exists, or is fictitious, or is perhaps completely absurd. I think of Jupiter [fictitious] as I think of Bismark [real], of the tower of Babel [mythical] as I think of Cologne Cathedral [real], of a regular thousand-sided polygon [real, or at least possible] as of a regular thousand-faced solid [impossible].

These so-called immanent contents are therefore merely intended or intentional [that is, immanent only in the second sense: “directly given”], while truly immanent contents [that is, “genuinely” immanent], which belong to the real make-up of the intentional experiences, are not intentional: they constitute [Note: This is not ‘constitute’ in the technical sense we have encountered in Kant and will soon encounter in Husserl.] the act, provide necessary *points d’appui* [= points of support] which render possible the intention, but are not themselves intended, not the objects presented in the act. I do not see colour-sensations but coloured things, I do not hear tone-sensations but the singer’s song, etc., etc.

In other words, if (to use Husserl’s expression) I “dismember” a thought of the god Jupiter, I won’t find Jupiter in there. What I will find is certain real ingredients that make up that act of thinking, but not that make up its object. These are what Husserl says are

necessary “points of support” for the thought. They are the things that enter into the make-up of the actual act of thinking. But *they are not what I am thinking about*.

Let me return to an earlier point for a moment. Earlier, I suggested that these ingredients might be things like the duration and the intensity of the thought. And while it is probably true that we could extract those features by “dismembering” a thought, they don’t seem to be what Husserl is talking about here. Remember, we said that because these features are universals, they are not “genuinely immanent” in the thought, and we then wondered just what was genuinely immanent in a thought. Well, in the passage I just read you, Husserl doesn’t seem to be talking about these universal features when he says ‘point of support’. In this connection, it is interesting that he goes on to say:

I do not see colour-sensations but coloured things, I do not hear tone-sensations but the singer’s song, etc., etc.

The context suggests that these color-sensations or tone-sensations are what are genuinely immanent in the thought, so that these are the answer to our question for Husserl.

There’s a lot to be worked out here, and I don’t want to pursue it any further since I am trying to set the Husserlian stage for what is coming up in Sartre — and as far as I can see, this point of Husserl-interpretation, although it’s interesting enough in its own right, doesn’t really lead up to anything in Sartre. Instead, let’s ask:

Why did Husserl think every act of consciousness was intentional in the sense we have just sketched under these three claims? He has no argument for it. (Remember, on his own principles he cannot argue in the sense that his point would depend on the argument.) Rather, he thinks that the eidetic abstraction of *what it is to be an act of consciousness* (the “essence” of consciousness) reveals that feature of it. (Recall, he said in Lecture I that he was going to have to nail down “the essence of consciousness” and “the essence of being an object of consciousness.” It’s the eidetic reduction that is going to give him these essences.)

This then is the payoff of the phenomenological method. Remember how Husserl said he was engaged in a “critique of cognition.” This is how it succeeds.

This notion of intentionality came as a breath of fresh air to many people. It was welcomed by lots of philosophers who saw in it the way to break out of Cartesian “subjectivism,” out of “idealism,” and to put the mind back again into direct contact with transcendent realities.

A nice example of this is in Sartre’s short paper, “Intentionality: A Fundamental Idea of Husserl’s Phenomenology” (1939). I have put a copy on reserve in the main Department office. (You shouldn’t look for too much theory in this paper. Sartre is being delightfully literary here.)

He begins by talking about the “traditional” (Cartesian-Kantian) theory according to which knowing something is a process that takes place *entirely within the mind*.

The problem with this kind of theory, for Sartre (and for Husserl too), is that we start off by wanting to know about, say, the tree, and yet we end up knowing only about our own *thoughts*. Sartre speaks of this kind of theory, in contemptuous terms, as a “digestive philosophy”: to know something is for the mind to make it part of itself, to “assimilate” it, to “devour” it (see also one of the paper topics in the course packet):

... we have all believed that the spidery mind trapped things in its web, covered them with a white spit and slowly swallowed them, reducing them to its own substance. What is a table, a rock, a house? A certain assemblage of “contents of consciousness,” a class of such contents. [Sartre is describing idealism here.] O digestive philosophy!

Against this doctrine, Sartre contrasts Husserl’s new theory of “intentionality”:

... Husserl persistently affirmed that one cannot dissolve things in consciousness. You see this tree, to be sure. But you see it just where it is: at the side of the road, in the midst of the dust, alone and writhing in the heat, eight miles from the Mediterranean coast. [The point is that you don’t see it in your mind.] It could not enter into your consciousness, for it is not of the same nature as consciousness.

Sartre then goes on to describe this new doctrine in the most rapturous terms.

This then is what people found so attractive in Husserl’s earlier philosophy.

But in his later work, Husserl seemed to throw it all away.

This brings us back to the notion of *constitution*, as discussed earlier. Recall our discussion of the Gestalt figure (two faces or a vase), and of the Kantian doctrine of the contribution of the Ego to the phenomena. (Also, Sartre on anger.)

Husserl began to worry about this point. (Not about the Gestalt figure in particular, but about the general point.) It’s as if he decided that he had not paid enough attention to the role of the Ego in constituting phenomena. And this represents a major turning point in Husserl’s thinking.

If you listen to what Husserl says when he is talking about the theory of *intentionality*, it sounds as if what he is saying is this:

When I perceive the “bent” oar in the water, there is a phenomenon present to my consciousness. That phenomenon has various features, which I can describe. Those features are *not* features of the *act* by which I am conscious of that phenomenon. This is just what have heard Husserl saying.

In short, when we think in terms of the theory of intentionality as Husserl presents it in his *early* theory, *all content comes from the side of the object*. The theory of intentionality gets rid of all “mental contents,” as

Sartre and others are fond of saying. In other words, the act of consciousness contributes nothing — it simply watches.

Obviously what we have here is an Ego that is not involved in any kind of constituting activity. It is simply a passive observer. In this respect, we are back with the Cartesian picture we started with. (Only in this respect, since we have gone beyond Descartes by rejecting the second of his two principles.)

But Husserl began to worry about this, in the same way that Kant began to worry about Descartes' notion of the Ego as a purely passive observer. And, the more he thought about it, the more Husserl began to think that consciousness did contribute something to the phenomena after all. In other words, he began to adopt the notion of a "Transcendental Ego," and Ego that constitutes phenomena.

Husserl doesn't discuss this very clearly in Lecture V, where it is on the program, and is supposed to be discussed. He does say some cryptic things about it there, but they are not very clear.

There is a somewhat more enlightening discussion in the corresponding passage of "The Train of Thought," which differs interestingly from what Husserl actually said in Lecture V. Listen to what he says in "The Train of Thought" (p. 9 — my emphasis):

At the lowest level of reflection, the naive level, at first it seems as if evidence [recall, this just means "self-givenness," being "immediately given"] were a matter of simple "seeing," a mental inspection without a character of its own, always one and the same and in itself undifferentiated: the "seeing" just "sees" the things, the things are simply there and in the truly evident "seeing" they are there in consciousness, and "seeing" is simply to "see" them. [Consciousness then would be something like a "searchlight," lighting up phenomena.] Or, to use our previous simile: a direct grasping or taking or pointing to something that simply is and is there. All difference is thus in the things that exist in themselves and have their differences through themselves. [Don't put too much emphasis on the 'exist' here.]

The last sentence is the crucial one. It means, of course, that the differences among things do not come from the constituting activity of the mind.

Husserl says it seems this way at first, at the "lowest level of reflection," the "naive" level. But now, in the very next paragraph, we see Husserl beginning to move away from this notion and toward a theory of constitution (pp. 9–10 — again my emphasis):

And now how different the "seeing" of things shows itself to be on closer analysis. Even if we retain under the heading of attention the notion of an undifferentiated and in itself no further describable "seeing," it is, nevertheless, apparent that it really makes no sense at all to talk about

things which are “simply there” and just need to be “seen.” On the contrary, this “simply being there” consists of certain mental processes of specific and changing structure, such as perception, imagination, memory, predication, etc., and in them the things are not contained as in a hull or vessel. [*I.e., they are not “genuinely” immanent.*] Instead, *the things come to be constituted in these mental processes*, although in reality they are not at all to be found in them. [*That is, they are “genuinely transcendent.”*]

Earlier, we took the example of the Gestalt figure. Husserl himself (earlier on p. 9) takes the example of the *auditory perception of a melody*. There the discrete tones are organized by the mind into a whole, into a melody.

The more he thought about it, the more Husserl came to think consciousness contributed to the phenomenon — exactly as happened with Kant. But Husserl never subscribed to the limited list of Kantian categories. For him, the mind eventually came to contribute all content to the phenomena.

So Husserl’s theory of what we have called the “phenomenological Ego” (our term, not Husserl’s) — which we have already seen is not to be identified with the *psychological* personality, as Descartes implicitly assumed — turns out as Husserl’s thought develops to be a lot more than the austere, pure “point of view,” “eye of the camera” that we thought it was in the early stages of Husserl’s philosophy. It has a bigger job to do than that. It is no longer regarded as an *empty perspective or point of view* — but as a kind of *projector* in the theater of the mind.

This *developed* notion of the impersonal Ego (it’s still not the “psyche”) that remains after the phenomenological reduction is what Husserl came to call the “Transcendental Ego.” (Kant had already used that term.) The Transcendental Ego *organizes* the raw data the same way the Ego did for Kant. (This Transcendental Ego is still not the same as the *psychological* Ego or psyche.)

Here is a list of jobs the Transcendental Ego is supposed to do:

- (1) A *constituting* job. It *organizes* the raw data of consciousness, which by themselves have no structure. (By themselves, they are like the featureless *screen* in the movie theater; all *content* comes from the *projector*.)
- (2) A *unifying* job. It *ties* the phenomena together into a coherent picture, a coherent “movie.” In Husserl’s example of the perception of a melody, for instance, what we have is not just an act of consciousness of *one* tone, followed by an act of consciousness of the *next* tone. Rather, the Ego links all these momentary acts together into a single consciousness of the *melody*. To take another example, we have not just an act of perceiving, an act of fearing, an act of imagining, but rather first *I* perceive, and

then I fear, and then I imagine. The Ego is what ties these things together.

- (3) An individualizing job. This is not the same as (2). Husserl himself doesn't use this terminology of 'unifying' vs. 'individualizing' to designate these two distinct functions of the Ego. But Sartre does (in *Transcendence of the Ego*), and so we might as well adopt his terminology here.

If the unifying job of the Transcendental Ego ties certain acts of consciousness together into a coherent story, the individualizing job is what makes one such unified story distinct from another. It is what makes one mind distinct from another (if there are any others). Or to put it another way, if the unifying job of the Ego is what ties certain acts of consciousness together, the individualizing job is what excludes other acts from that bundle.

The last two jobs of the Ego, its unifying and individualizing jobs, were performed by the Ego even before Husserl began to think of it as a Transcendental Ego, as a kind of projector. (Sartre makes this point in *Transcendence of the Ego*.) The Ego "unifies" and "individualizes" in virtue of the fact that it is a kind of perspective or eye of the camera. You can't be taking two points of view at once; one perspective excludes all others.

I made this point earlier, when I was talking about what we called the "phenomenological" Ego: The Ego that remains after the phenomenological reduction, while it is not the personal Ego that we study in psychology, is nevertheless an individual Ego.

Important: In his earlier theory of the Ego, all content came from outside, and the Ego was just a passive observer (as for Descartes). In this later theory, as a result of the doctrine of constitution, it is just the reverse: All content comes from the Transcendental Ego.

Question: Why go all the way? In our example of the Gestalt figure, it was clear that the mind contributed part of the structure we saw in the phenomenon; it was what decided which region was going to serve as foreground and which as background, the light area or the dark area. But not all structure, not all content, comes from the mind in that example. (At least we don't yet have any reason to think it does.) The basic configuration of light and dark comes from the outside — comes from sensation. What reason is there in anything we have seen so far — in Kant or Husserl or anyone else — for going all the way with this idea and supposing that the mind contributes all content?

The answer to this is not at all clear in Husserl — at least not to me. On the other hand, there will be some theoretical motivation for this step in Sartre, as we shall see. And it is an crucially important question to ask.

Sartre

Let's turn now to Sartre.

The first passage I want to discuss from Sartre is a short passage from *The Psychology of Imagination*: Part I, § 3 (pp. 8–14). Remember that this was Sartre's second book on the imagination, written in 1940, after his *Imagination: A Psychological Critique* (1936). I think it's a crucial passage. It will help to explain some things in Husserl, and will clarify some obscure points at the beginning of Sartre's "Introduction" to *Being and Nothingness*. We will be constantly returning to this passage, so get completely familiar with it.

My purposes in discussing this passage are threefold:

- (a) To illustrate the relative concreteness and vividness of Sartre's writings, in contrast to Husserl's. Sartre can write obscurely too, but he doesn't always. Here we have a fine example of phenomenological description at its best.
- (b) To shed some light on what Sartre says at the beginning of his "Introduction" to *Being and Nothingness* about Husserl, namely, that Husserl had succeeded in getting rid of a lot of troublesome dualisms that had haunted traditional philosophy for centuries. For instance, the dualism of: phenomenon vs. noumenon, appearance vs. reality, etc. But he had not succeeded in getting rid of all such dualisms. All he has done is to reduce them to a single dualism or dichotomy: the dichotomy between the finite and the infinite. What does Sartre mean by this?
- (c) To explain how Husserl (and Sartre) can continue to draw the distinction between the subjective and the objective, the real and the illusory, imagination and perception, fantasy and reality, while remaining within the phenomenological standpoint.

In the passage, Sartre says there are three basic types of consciousness by which the same object can be given to us, three ways the phenomenal thing can "come on to us." According to this way of dividing them up, there will be three and only three types of consciousness. (There may also be other ways of dividing things up, of course.)

The three are: perception, imagination, and mere conception (that is, mere abstract thinking as distinct from visualizing). I may perceive a cube, for instance, or I may imagine one, or I may be merely thinking of one (without any accompanying mental imagery), in the way a mathematician might.

(In the passage, Sartre treats these in the order: perception, conception, and imagination. For my own purposes, I will treat them in a different order.)

What are the differences among these three? Let's see what he says:

Perception:

When I perceive a cube, I see at most three sides of it at once. (I may see fewer, depending on where I am situated with respect to the cube. But I will never see more than three at once.) I see the cube only "in profile," as Husserl had put it. A "profile" in German is an Abschattung (= shadowing off). Thus we will find Husserl saying that the cube appears in Abschattungen. (The cube example Sartre got from Husserl himself.)

But there is something odd about these three sides. In a narrow sense of the term, I see only them. But, insofar as I see them AS three sides of a cube, they carry with them a kind of promise. The promise is that there are three other sides around in back. The three sides I see "promise" three more I don't see. There is the promise of "more to come."

Thus, I can be said to perceive "the cube," but I do not perceive all of the cube. I see the three sides facing me, and I see them AS part of a larger whole, the rest of which I do not see at the moment.

This feature is characteristic of perception. Perception is always a matter of perceiving only part of a larger whole, and getting only a kind of "promissory note" for the rest.

Note that this is not just a matter of visual perception. For example, "hearing a footstep" promises that there is a foot to be seen, even if I am not in fact seeing it now.

If I turn the cube around, I will of course now be in a position to see the remaining three sides. But in the process, the three sides that originally faced me have disappeared from view. Now they are only "promised."

Let's go back now to our original situation, and look at the first three sides, which we perceive as three sides of a larger whole cube. And now suppose I turn the object around, and the three remaining sides aren't there after all. What I perceive "as" a cube is really not a cube at all, but only a kind of hollow facade. In that case, of course, there is still no denying that I originally perceived those three sides as sides of a cube, and therefore as promising the three other sides around in back. But those other three sides are not there, and the "promise" made by the three sides facing me is an empty promise; the object fails to keep its promise. It is not what it seemed to be.

What is to prevent this? Nothing at all. The point is, there is always something tentative and dubious about perception. Perception can always be corrected in the light of further perception.

In fact, if you think about it, those same three sides of the cube that face you make a number of other "promises" as well. Insofar as you perceive the object as a solid cube (rather than, say, some sort of light-cube created with mirrors), the implication — the "promise" — is that if I put my finger on one of the surfaces, it will encounter a resistance, and not just pass right through the empty space.

And so on. It is not hard to see that, in the long run, there is an *infinity* of “promises” made by those three faces seen *as* three sides of the cube. And of course, I would never be able to check out *all* those promises.

For practical purposes, it is not hard to check out enough of them to be able to say, “Well, I’m going to treat this thing as a cube and be done with it.” But that’s just a matter of practical expediency (the “natural standpoint”). In principle, it is always possible that the object would *fail* the next “cube”-test I put it to.

As a result, in the case of perception, we always — to use Sartre’s apt phrases — have to “learn” the object, “make a tour” of the object, “serve an apprenticeship” to it. The outcome is always in doubt, and I can always *learn* something from putting the object to the test.

Thus (let us diagram this):

	<u>Makes a promise</u>	<u>Promise can fail (there is a “risk”)</u>
Perception	Yes	Yes

There should really be nothing surprising or shocking about any of this. All Sartre is saying in effect is that the notion of the “objectivity” of perception always carries with it the notion of *testing*, of *experiment*, of the *tentativeness* of scientific theory, etc.

The phenomenon of the *perceived cube*, therefore — that is, of the *three sides perceived as part of a cube* — presents us with a kind of *reference* to something *else* besides what we directly see.

But — and this is the crucial point — *notice* what it is that is being referred to in his way. The reference is *not* to something *hidden behind the phenomena*, some “thing-in-itself,” some mysterious *noumenon* behind the *phenomenon*, but to *other potential phenomena*. In fact, to an *infinity* of other potential phenomena.

This is what Sartre means when he says (in the “Introduction” to *Being and Nothingness*) that Husserl has replaced the Kantian dualism of *phenomenon* and *noumenon* by the dualism of *finite* vs. *infinite*. For Husserl (at least as Sartre interprets him), the cube just *is* the infinite sum of these actual and potential phenomena. There is nothing *more* to the cube than that. (This will not be Sartre’s view.)

Thus, a phenomenon is a *perceptual* phenomenon — that is, it comes on to us in the “perceptual” way and makes an “*objective*” claim about reality — if the phenomenon refers in this *tentative* way to *an infinite series of phenomena*.

In our example, the phenomenon makes the “objective” claim “*I am a cube*,” with all the infinity of implications that involves.

Such an “objective” claim is “objectively *true*” if all the implications *hold*, if all the “promises” come true. That is something we can never finally be sure of, to be sure. And

of course, as phenomenologists, we are not even really concerned with whether those “promises” actually come true or not. That would be to go beyond what is directly given to us and to infer something further. The phenomenological reduction prohibits our doing that.

Nevertheless, as good phenomenologists, we can describe what the “objective” claims are when we perceive the three sides as three sides of a cube. What we cannot do is try to validate those claims.

In the “Introduction” to *Being and Nothingness*, Sartre also says that for Husserl, the essence is the “principle of the series,” and is itself a phenomenon. What does this mean?

Well, that the essence is itself a phenomenon we already know — from Husserl’s discussion of eidetic abstraction. To say that the essence is the principle of the series just means that it is the essence of the cube that determines what is promised by the three sides directly facing me.

That is, when I see the three sides as three sides of a cube, I see the three sides and also see the essence cube. I see the three sides, so to speak, as exhibiting the essence cube. And because I see the three sides as exhibiting that essence, I see them as promising what they do, the whole infinite series of other potential phenomena.

This is what it means to say the essence is “the principle of the series” of phenomena.

Note that the essence of a cube can be a phenomenon in this way, even if we don’t really have a cube at all. That just means that the promises are risky, and might turn out false, as we have already seen.

Note also that the essence of a cube can be a phenomenon in this way, even though we may come to learn about cubes, to think in terms of cubes and other geometrical figures, through social conditioning or other relative factors. There’s nothing in anything Husserl or Sartre has said so far that is in any way incompatible with all sorts of relativistic theories about the way we perceive things. What we are talking about now is not how we come to see things the way we do, but rather what’s going on when we see them that way.

We are now in a position to answer a question that may have been bothering you ever since we talked about the post-Kantians, who insisted that there is not and cannot be anything beyond the realm of phenomena, that the Kantian notion of a world of noumena was contradictory and absurd. The question that may be bothering you is: “What happened to the real world?” If everything takes place at the level of phenomena, how can we any longer make the distinction between reality and illusion?

Well, Husserl’s answer — and Sartre’s answer — is “the real world is still there, and we can still distinguish it from cases of illusion.” Think: How in practice do we tell the real from the illusory?

Suppose you take a jewel to a jeweler, and ask: “Is this a diamond or not? It looks like one to me, but I’m no expert. Is it a real diamond or a fake?” How does the jeweler proceed? He checks it to see, for example, whether it can cut glass. He checks to see if it has the right refractive index, specific gravity, and so on. In principle, there is an infinity of tests he could run, although in practice we are satisfied after only a few of them.

Now notice what the jeweler is doing. He is proceeding entirely at the level of phenomena — the phenomenon of what appears to be a diamond, and the further phenomena that are “promised” by the fact that it appears to be a diamond. He is testing those promises, and checking to see if the promised further phenomena actually show up.

What he is not doing is checking to see if there some kind of real diamond (which we don’t see) out there hiding behind the apparent diamond (which we do see). When we ask him whether our apparent diamond is a real one or not, this is not what we are asking him, and this is not the question he tries to answer. Furthermore, no one thinks it is.

In other words, the way to distinguish reality from illusion, the real from the fake, is just to check out the promises, to perform the tests.

So, to the question “What happened to the real world?,” the answer is: “It’s still there, just as it always was.” Phenomenology doesn’t do away with the real world at all. What it does is to do away with a bad theory about what we are talking about when we talk about a “real world” — the Cartesian theory that put the whole business in terms of a “correspondence” or lack of “correspondence” between the phenomena and something else. We reject this “correspondence” theory in favor of what is sometimes called a “coherence” theory.

You may have another question. You may say: How can we get by with saying that the Cartesian picture is false? How can we allow ourselves the luxury of saying there is no “thing-in-itself”? I thought we were adopting the epoché and were not going to allow ourselves to pass judgment on what does and does not exist. What about it? What happened to the epoché? What happened to the “bracketing of existence”?

Just as before, the answer is: Nothing has happened to it. It’s still in force. We haven’t violated it in the slightest. We still do not pass judgment on what does and what does not exist. But we do pass judgment on a bad theory of what it means to say that a thing “exists,” or is “real.” We reject the Cartesian theory of what this means. That theory is incoherent. But the epoché is still in effect. We still do not try to decide what does and what does not exist. This means only that we do not actually perform the tests, we do not actually check the promises of “more to come,” we do not “take the risk” of actually committing ourselves to their being fulfilled. We can describe what those promises are, but as phenomenologists we are not interested in whether they are fulfilled or not.

Imagination:

Now contrast all this with what happens in the case of imagination. Instead of actually looking at our apparent cube, let us now just shut our eyes and imagine a cube.

Just as before, we imagine it *in profile*, from a perspective. We imagine it as presenting to us *at most* three sides at once. The cube is presented in *Abschattungen*. So imagination is *like* perception in that respect.

But now take the three sides that face you in imagination, and turn them around, just as we did with the perceived cube. In this case, *is there any danger at all* that what we imagined as the three sides *of a cube* will turn out not to have the three more sides in back? *Is there any possibility at all* that what we are imagining would *fail* any of the tests it would have to pass in order to be a cube, if we should perform those tests in our imagination?

No, of course not. *I'm* the one imagining this, after all, and if I imagine it *as* a cube, then it's a cube.

Thus, in imagination, unlike perception, the “promises” made of “more to come” are *guaranteed*. There is no way they could fail.

In this sense, we don't have to “serve an apprenticeship,” “make a tour,” “learn” the objects we imagine. There is an important sense in which we never really *learn* anything new from imagining. There's nothing to be learned there except what we have already built in. Thus we have:

	<u>Makes a promise</u>	<u>Promise can fail (there is a “risk”)</u>
Perception	Yes	Yes
Imagination	Yes	No

This difference *makes* all the difference between the “objective” and the “subjective” — in the sense of an objective or a subjective *claim*. Perception is “objective” in the sense that in perception the phenomena always make a claim of “more to come,” and that claim may turn out to be false. Imagination is “subjective” in the sense that there the phenomena always make a claim of “more to come,” and that claim is *guaranteed* not to turn out false, because the answer is built in from the very beginning.

Note once again: In this usage of the terms, ‘objective’ does not mean *true*. It means “testable,” the proper kind of thing on which to perform *experiments*. Thus, a *sensory hallucination* is objective but false. But it is not a case of *imagination*. A mirage is a perception, not a case of *imagination*, because I can test it, I can be *fooled* by it — I can even *photograph it!* (Sartre emphasizes this point in *Imagination: A Psychological Critique*, that imagination is *not* just “faint perception,” as Hume had thought of it.)

In short, don't confuse *imagination* with *false perception*. They are totally different things.

Question for future reference: How can this way of distinguishing perception from imagination be reconciled with the theory of constitution taken in the strong sense in which the mind contributes all content to what it is conscious of? If the reason I cannot be surprised by imagination is simply that I am the one who put all the content into the imagined object in the first place, so that there is really nothing tentative about it — well, isn't that also true for perception, and indeed for any act of consciousness? The answer is: Yes, it is possible to reconcile a strong theory of constitution with this way of distinguishing perception from imagination. That will be a long and important story. But begin thinking about it now.

Conception:

Now contrast both these first two cases with the case of conception. Here we are talking about the kind of thing an abstract mathematician might engage in, without resorting in any essential way to imagination.

In the case of conception, the cube is presented to me all at once. It is not presented to me “in profile.” It doesn't make any sense to talk about conceiving of a cube “with the three other sides around in back.” That kind of talk may go with perception or imagination, but not with conception. When I am conceiving of a cube, I am thinking of it simply as a Euclidean solid figure with six square sides, etc. (There may be some imaginative visualization that *goes along* with this, but that's not what we're talking about now.) Thus, if you say you are conceiving a cube, it makes no sense to ask from what angle you are conceiving it. There aren't any sides around in back; they are all equally presented at once in the concept.

Thus, unlike the two preceding cases, in conception there is no “promise of more to come.” And since there is no promise, there is of course no danger that the promises might not come true.

Just as with imagination, therefore, I do not have to “serve an apprenticeship” to the concept, I do not have to “make a tour” of it. In short, I can't learn anything from a concept alone, any more than I can from an image.

(In a sense, of course, we certainly do learn things from concepts. That's what mathematicians do when they prove new theorems. But that is a matter of inference and reasoning. Sartre's point is that I cannot learn anything from them by simply inspecting the phenomena.)

Thus, to complete our table, we have:

	<u>Makes a promise</u>	<u>Promise can fail (there is a “risk”)</u>
Perception	Yes	Yes

Imagination	Yes	No
Conception	No	No

Thus, we see how Husserl and Sartre can continue to draw the distinctions between appearance and reality, between the subjective and the objective — and in fact, to do it just the way people ordinarily do it. And we have seen the role of the duality of finite/infinite in Husserl.

Finally, note why there can be three and only three ways of being conscious of an object according to this way of dividing things up. There is no fourth possibility, with ‘No’ in the middle column and ‘Yes’ in the right column, since if an object makes no promises of “more to come” in the first place, there are no promises that can fail and so no “risk” involved.

Keep this passage from *The Psychology of Imagination* constantly in mind as we proceed. It will be absolutely crucial.

Sartre’s Reaction to Husserl

So far, we have talked mainly about Husserl, and about Sartre insofar as he agrees with Husserl. How does Sartre disagree with Husserl? What does he accept and what does he reject of Husserl’s theories?

The key to this is to remember that Sartre was influenced by that other stream — not just phenomenology, but also Nietzsche, etc.

That tradition emphasized the individual, we said, as opposed to the universal or general. Sartre accepts that emphasis.

As a result, he will downplay the role of Husserl’s eidetic reduction. He doesn’t deny that universals are directly given to us. (Some commentators say he does, but I think that is not right, and I can prove it.) But universals do not play the same role in Sartre’s philosophy as they do in Husserl’s.

Husserl, recall, thought that the most important feature of a thing was not that it exists, but rather what it is, and that by “bracketing” the existence of the thing, he was not in the end bracketing anything very important about it. What was important for Husserl was those universal essences in terms of which you could describe what the individual thing is phenomenally. (Remember how Husserl announces his eidetic reduction as the result of thinking about the subject/predicate judgments we make about phenomena.) Any kind of theoretical discussion is going to have to proceed in terms of these essences. For Husserl, you can do pretty much everything you really want to do at the level of universal essences.

Sartre will have none of that. He thinks the individual is primary, and that you cannot get at what is really important and interesting about an individual by thinking of it as in effect nothing more than the intersection of a bunch of general principles. So, for Sartre, while there are these essences given in eidetic abstraction, just as Husserl said there were, they are not where the emphasis lies in Sartre.

There is a striking passage late in *Being and Nothingness* that illustrates this point very forcefully. It occurs on pp. 713–715, near the very beginning of the Chapter on “Existential Psychoanalysis.” Sartre is talking about the common practice of writing psychological biographies of people, of organizing your subject’s life around certain Freudian themes, for example, and in practice thinking of your subject as though he were simply a collection of universal principles. He quotes a passage from one such biography, a biography of Gustave Flaubert. Here is what he says:

... A critic, for example, wishing to explain the “psychology” of Flaubert, will write that he “appeared in his early youth to know as his normal state, a continual exaltation resulting from the twofold feeling of his grandiose ambition and his invincible power The effervescence of his young blood was *then* turned into literary passion as happens about the eighteenth year in precocious souls who find in the energy of style or the intensities of fiction some way of escaping from the need of violent action or of intense feeling, which torments them.

Then Sartre goes on:

In this passage there is an effort to reduce the complex personality of an adolescent to a few basic desires, as the chemist reduces compound bodies to merely a combination of simple bodies. The primitive givens will be grandiose ambition, the need of violent action and of intense feeling; these elements, when they enter into combination, produce a permanent exaltation. Then — as Bourget [the author of the biography Sartre just quoted] remarks in a few words which we have not quoted — this exaltation, nourished by numerous well-chosen readings, is going to seek to delude itself by self-expression in fictions which will appease it symbolically and channel it. There in outline is the genesis of a literary “temperament.”

Now in the first place such a psychological *analysis* proceeds from the postulate that an individual fact is produced by the intersection of abstract, universal laws. The fact to be explained — which is here the literary disposition of the young Flaubert — is resolved into a combination of *typical*, abstract desires such as we meet in “the average adolescent.” What is concrete here is only their combination; in themselves they are only possible patterns. The abstract then is by hypothesis prior to the concrete, and the concrete is only an organization of abstract qualities; the

individual is only the intersection of universal schemata. But — aside from the logical absurdity of such a postulate — we see clearly in the example chosen, that it simply fails to explain what makes the individuality of the project [understand: “person”] under consideration. The fact that “the need to feel intensely,” a universal pattern, is disguised and channeled into becoming the need to write — this is not the *explanation* of the “calling” of Flaubert; on the contrary, it is what must be explained....

At each state in the description just quoted, we meet with a hiatus. Why did ambition and the feeling of his power produce in Flaubert *exaltation* rather than tranquil waiting or gloomy impatience? Why did this exaltation express itself specifically in the need to act violently and feel intensely? Or rather why does this need make a sudden appearance by spontaneous generation at the end of the paragraph? And why does this need instead of seeking to appease itself in acts of violence, by amorous adventures, or in debauch, choose precisely to satisfy itself symbolically? And why does Flaubert turn to writing rather than to painting or music for this symbolic satisfaction; he could just as well not resort to the artistic field at all (there is also mysticism, for example). “I could have been a great actor,” wrote Flaubert somewhere. Why did he not try to be one? In a word, we have understood nothing; we have seen a succession of accidental happenings, of desire springing forth fully armed, one from the other, with no possibility for us to grasp their genesis. The *transitions*, the becomings, the transformations have been carefully veiled from us....

I cannot imagine a more persuasive rejection of the primacy of the universal.

Furthermore, insofar as *knowledge* proceeds in these universal and general terms (as Husserl realized), Sartre’s rejection of the primacy of universals means that *we must abandon the primacy of knowledge*. This is one of the main themes in the Sartre’s “Introduction” to *Being and Nothingness*. We are not going to be able to grasp the real individuality of the individual by *thinking* about it. (Remember this. It will be important.)

(Sartre thinks there are other ways of grasping this individuality — for example, in what he calls certain “privileged emotions.” In particular, in what he calls “nausea” and describes in his novel by that title.)

The reactionary stream that includes Nietzsche and that influenced Sartre not only emphasized the individual at the expense of the universal, it also emphasized the notion of *human freedom*, the absence of any *general principles* of a moral or metaphysical nature to determine what we are and what we ought to do. This is what Sartre brings out so strongly in his essay “Existentialism Is A Humanism” in the phrases ‘Existence precedes essence’ and ‘Man makes himself’ and in the famous example of the student in that essay. (If you have not yet read “Existentialism Is A Humanism,” do so now, along with my notes on that essay in the course packet.)

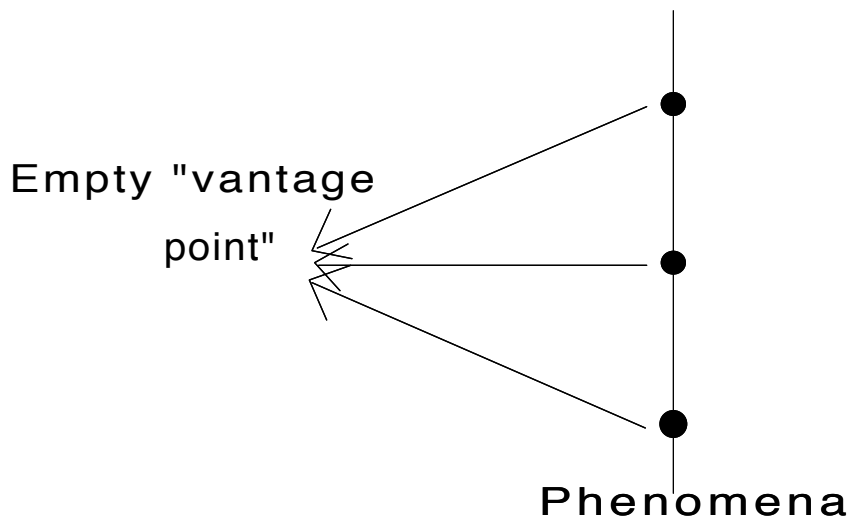
Thus, in Sartre too we find a strong emphasis on human freedom and the lack of moral or metaphysical absolutes. One important question we shall have to ask is why Sartre is so sure human beings are free in his sense.

But that's for later on. For the present, the point I want to make is that Sartre thinks human freedom is incompatible with the later Husserlian doctrine of the Transcendental Ego. Sartre's reasoning here frequently baffles students, and we will have to return to it several times. But basically, Sartre thinks the Transcendental Ego would play the role of a kind of deterministic human nature, a kind of preprogrammed "projector" in our phenomenological movie theater. In short, Sartre thinks such a Transcendental Ego would get in the way of the spontaneity of human freedom.

Exactly why that should be so is the point that always baffles students at first. So for the moment, let's just grant Sartre his point for argument's sake and see what consequences this has for his theory of consciousness.

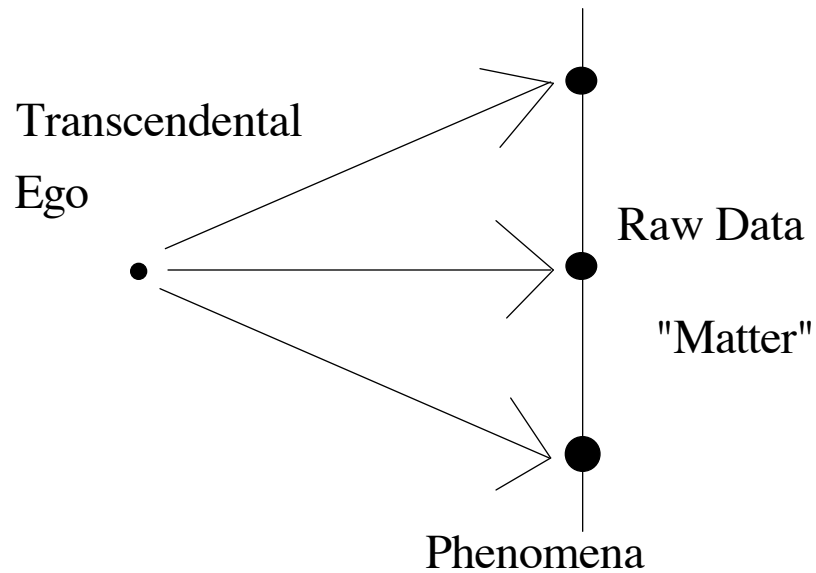
The Transcendental Ego thus, we suppose, gets in the way of human freedom. But Sartre is not willing to sacrifice human freedom. So, from this standpoint, Husserl's earlier view is more congenial to Sartre than Husserl's later theory.

The earlier theory, recall, looked like this:



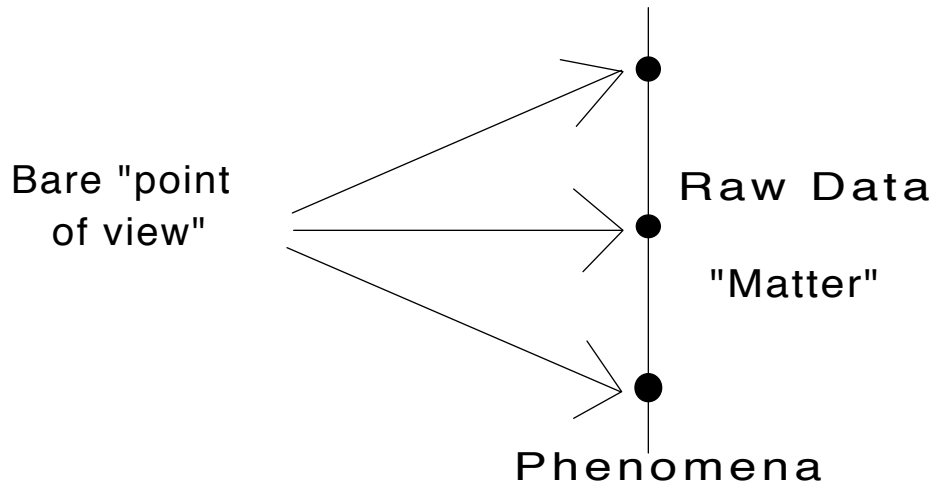
In this theory, there was no projector in the mental movie-theater. The Ego is a bare "vantage point," a kind of geometrical limit — with nothing back there. All content comes from the object, remember, on this earlier theory. The Ego is just an observer.

Still, Sartre takes seriously the later Husserlian doctrine of constitution. This, you will remember, looked like this:



But of course that later theory, as it stands, has the objectionable Transcendental Ego.

What Sartre does in this situation is to combine Husserl's two theories. He retains the notion of constitution, but allows no constitutor — no Transcendental Ego. Thus Sartre's theory of consciousness might be diagrammed like this:



(Note the directions of the arrows in these diagrams.)

In other words, in the later Husserl's theory, our phenomenological movie-theater is showing a movie cast on the screen by the Transcendental Ego, which serves the role of the "projector." But in Sartre's theory, our phenomenological movie-theater is showing a movie cast on the screen by *nothing at all*. There's *no projector* in this theater. The "light rays" just spontaneously emerge from a common "vantage point" and hurl themselves at the screen. That's *real* freedom!

Sartre does not use the term 'Ego' for this empty "vantage point" — not even the term 'phenomenological Ego', which we used for Husserl's earlier theory. And he *certainly* doesn't use the term 'Transcendental Ego'; the whole point of this theory is to *do away* with the Transcendental Ego. (The early Husserl used the term 'Ego' for his "bare vantage point." *We* called it a "phenomenological Ego" to distinguish it from the later doctrine of the Transcendental Ego, but Husserl himself did not do that.)

There is one other point I want to bring out here to contrast Sartre's theory with Husserl's.

In Husserl's later doctrine, there is a theory of *levels* of constitution. The doctrine is obscure. But as Sartre interprets him (I think plausibly), Husserl holds that the Transcendental Ego generates *everything* else in our diagram. It not only generates ("projects") the various *acts* of consciousness that play the role of "light beams" in our analogy. It also generates ("projects") *the screen!*

What can that mean? Well, go back and consider our example of the Gestalt figure that can be viewed as either two faces or as a vase. We used that to illustrate the doctrine of constitution, the mind's own contribution to building up the phenomena.

Now in that example, the mind (or "Ego," if we want to talk that way) was what decided whether it was going to be the light areas or the dark areas that served as foreground in the figure. So that much of the structure of what we see in the phenomenon came from us. But the actual spatial arrangement of light areas and dark areas — that did not come from us. That was given. (Or at least we have not yet seen any reason to think otherwise.) So the result — the phenomenon — is a combination of two factors: the raw data that are not our responsibility, do not come from us, and the organization of those data, which does come from us.

But now what if it turned out that what we have been calling the raw data are also the result of an organizing activity by the mind, an earlier and "deeper" (more "primordial") level of "constituting" — so that the configuration of light and dark areas is the result the mind's organizing and arranging of prior data, and then the mind goes on to organize that result yet further by deciding that the one will serve as foreground and the other as background?

In that case, the raw data given to the mind would come at a deeper level, and there would be two levels of organization by the mind, two levels of constitution.

But what if it turned out that there was yet a deeper level of constitution, and a deeper one yet, and so on?

In short, what if it turns out that there are no ultimate raw materials for the mind to go to work on, but rather what we have are ever deeper levels of constitution without any stopping point?

In that case, the mind would constitute absolutely everything in its phenomenological movie-theater. It is not as if there would be to begin with a raw, blank, uninterpreted movie-screen on which the mind then projected an organized structure (a "show"). Instead, the mind would — so to speak — project not only the movie but also the screen. (At this point, perhaps, the movie-theater analogy is no longer the best one to picture what is going on.)

This is exactly the doctrine Sartre attributes to the later Husserl. Husserl talks about something he calls "hyle" (= ὕλη), which is just Greek for "matter." And Husserl does talk in some places as if the Transcendental Ego constitutes its own "matter" — its own "raw materials."

The doctrine is referred to obscurely in the translator's "Introduction" to *Transcendence of the Ego*. (On the whole, I do not think that "Introduction" is to be relied on very much.)

The upshot of all this is that, for the later Husserl, absolutely everything besides the Transcendental Ego itself depends on the Ego — is a product of the Ego.

And in that quite *strong* sense, the later Husserl is an *idealist*.

It is a puzzling question (for me) why Husserl thought he had to go so far with this theory of constitution. What prompted Husserl to come to the view that there cannot be any mind-independent “givens”?

Well, whatever Husserl’s reasons were, Sartre disagrees with Husserl on this point. For Sartre, the *screen* in the movie-theater model — the raw material that is organized by acts of consciousness — is not itself a product of consciousness. The screen has a reality of its own, independent of consciousness.

In fact, Sartre thinks the theory of *intentionality* — the theory of consciousness as reaching out beyond itself to something *else* (recall the “irreflexivity” of the relation of intentionality and the “real transcendence” of intentional objects) — is simply *incompatible* with this later Husserlian idealism. That is why Sartre often accuses Husserl of having *thrown away* his earlier doctrine of intentionality by adopting his later theory.

Sartre is sometimes a little careless about how he puts this. In the “Introduction” to *Being and Nothingness*, for instance, he says that Husserl’s doctrine of *constitution* conflicts with his doctrine of *intentionality*. But that is not quite right. In fact, Sartre himself accepts *both* doctrines — constitution and intentionality. What he rejects is the Husserlian view that consciousness constitutes the *whole* object of consciousness — that consciousness doesn’t just “make a contribution,” it *does it all*.

For Sartre, the “movie screen” — the raw uninterpreted data on which consciousness goes to work — is not dependent on consciousness at some deeper level. Thus, Sartre is NOT an idealist, at least not in the *strong* sense we have just described. (He may be in other senses.) This is one of the main points he tries to establish in the “Introduction” to *Being and Nothingness*.

(It is worth stressing this point, because some commentators speak as if Sartre *were* an idealist, as if Sartre’s theory of constitution amounted to idealism. That’s not so. The theory of constitution does a lot of work for Sartre, but it does not *do it all*. There is always something *independent* of consciousness for Sartre.)

So, to summarize what we have said so far, the main differences between Sartre and the later Husserl are two:

- (1) The question of the Transcendental Ego. Husserl accepts it; Sartre rejects it.
- (2) The status of the “screen” — the materials on which the constituting function of consciousness goes to work. For Husserl, at least as Sartre interprets him, all those materials are themselves the products of consciousness at some deeper level. Not so for Sartre. The “screen” is independent of consciousness.

While we're at it, let's just review some other points on which Sartre agrees or disagrees with Husserl:

- (3) Does Sartre accept Husserl's phenomenological reduction? Yes, if you take the phenomenological reduction to be nothing but the policy of confining ourselves to a description of what is directly given. No, if you take the phenomenological reduction to involve also "suspending" the existence of the objects of consciousness. Sartre won't buy that. The "screen" really exists — there's no "bracketing" that!
- (4) Does Sartre accept the theory of constitution? Yes, if you take that theory as saying only that consciousness makes a contribution to the phenomena. No, if you take the theory of constitution as saying that consciousness contributes everything.

We are now in a position to ask ourselves: What is really real for Sartre? What, on Sartre's theory, is really inside that phenomenological movie-theater? And we can now see that the answer is: two things. (So Sartre is a dualist.) The two are:

- (1) the bare, neutral, "uninterpreted" screen; and
- (2) the acts of consciousness.

The phenomena — the "movie" — are not some third kind of ultimate reality. They are what appear on the screen. They are the product of the two ultimate realities: the uninterpreted screen, and the interpreting activity of consciousness. In other words, John Wayne and the wagon train are not really in the movie theater; that's simply the "show."

So, in an important sense, Sartre thinks the world of experience, the world we encounter, has only a derivative and secondary kind of reality. It is the result of an absolutely neutral, featureless, inert reality (the "screen") and the interpreting activity of consciousness. And that's all.

It's important to understand here that, for Sartre, consciousness is not some thing in the theater that acts, some thing that constitutes the phenomena, so that we would have:

- (i) the thing, consciousness;
- (ii) its acts; and
- (iii) derivatively, the product of those acts, the phenomena.

No. For Sartre, consciousness is not a thing that acts; it is the act itself. You can refer the act if you wish to a "vantage point," a "point of view," as in Husserl's early philosophy. But there is nothing at that vantage point. The "vantage point" is in no sense the source and origin of the mental act. The act has no source or origin. It is completely spontaneous.

Sartre's Metaphysics

So for Sartre, there are two basic kinds of realities: the bare, neutral, undifferentiated “screen,” and the “light beams,” the “rays” cast on the screen.

The screen is passive, inert, as dead and featureless as anything could be. The light rays are consciousness; they are “alive,” flickering, lively.

The former — the screen — is what Sartre calls “being-in-itself” in the last section of the “Introduction” to *Being and Nothingness*. The latter — the “light rays” — he calls “being-for-itself.”

The terms ‘being-in-itself’ and ‘being-for-itself’ are derived first of all from Hegel’s *Ansichsein* and *Fürsichsein*. But, only slightly more remotely, the term ‘being-in-itself’ is an obvious allusion to Kant’s notion of the “thing-in-itself” — absolutely independent of our viewpoint. The screen of course is just the way it is, regardless what movie is projected on it. The movie in no way affects the screen!

And yet there is an important difference between Kant’s thing-in-itself and Sartre’s notion of being-in-itself. For Kant, the appearances, the phenomena, stand between you and the thing-in-itself, they hide the thing-in-itself, which is some mysterious and (some say) contradictory entity masked by the phenomena. For Kant, the fact that all we are directly conscious of is phenomena ensured that we are never able to get at the thing-in-itself. (In terms of our model, what Kant calls the “thing-in-itself” is not the movie-screen but rather something hidden behind the screen. For Kant, the screen plays the role of raw, uninterpreted sense data, caused by the thing-in-itself.)

For Sartre, by contrast, the phenomena don’t hide or mask being-in-itself; they reveal it. (This in effect is what the theory of *intentionality* says for Sartre.) The movie doesn’t make the screen invisible. On the contrary, the movie lights up the screen, so you can see it.

Of course, you never see the screen just by itself — with no picture on it at all. (At least not until the shows are over for the evening and they turn on the theater lights. But let’s ignore that for present purposes.) And yet every picture you do see on the screen reveals the screen. What you see is the screen lit up in a certain way, being-in-itself interpreted and processed in a certain way by consciousness. (That last part is what the theory of *constitution* is all about, according to Sartre.)

So for Sartre, being-in-itself is not something “hidden” from us; we don’t have to infer it beyond our phenomena. On the contrary, we are put in direct contact with being-in-itself in our every conscious act. In Husserlian terms, it is “immanent” to consciousness in the second of Husserl’s two senses of that term.

The difference between Kant’s thing-in-itself (at least in the “cartoon”-version of Kant) and Sartre’s being-in-itself is the same as the difference between Locke’s notion of matter (material substance) — in his famous phrase, a “something I know not what” hidden behind what I see — and Aristotle’s notion of matter (material substance), which

is certainly not “hidden” at all. (The statue doesn’t hide the bronze; it just is the bronze — “processed” or “interpreted” bronze.) And in fact, it is perhaps useful to think of Sartre’s being-in-itself as a kind of “matter” in the Aristotelian sense, so that when Sartre divides all reality into “being-in-itself” and “being-for-itself,” he is effectively dividing it up into matter and thought.

This is good to a first approximation, but we shall see that some adjustments will need to be made in this picture.

Thus, Sartre, unlike the later Husserl, is a dualist.

This “being-in-itself” is what in the “Introduction” to *Being and Nothingness* calls the “being of the phenomenon.” It underlies and supports the phenomenon, as the screen in a sense supports the movie picture. In a way, as Sartre agrees, it is not itself a phenomenon (“the being of the phenomenon is not itself a phenomenon”) — that is, we never see the screen just all by itself. And yet it is not hidden by the phenomenon; on the contrary, it is revealed in aspects and ways by the phenomenon.

When you watch a movie in the theater, the screen isn’t hidden from you. On the contrary, when you see John Wayne leading the wagon train crossing the Rio Grande, what you are seeing is really just the screen, lit up in a certain pattern. Of course, what you don’t see is the plain screen — the screen unprocessed and uninterpreted.

At the end of the “Introduction” to *Being and Nothingness*, Sartre gives us some general characteristics of being-in-itself.

Be careful. Sometimes (as in the passage we are about to discuss) Sartre speaks loosely, and says simply “being,” when he really means “being-in-itself,” which is only one kind of being. There is also being-for-itself. In fact, it is one of the main questions of *Being and Nothingness* to ask just how these two “regions” of being are related and connected with one another.

Characteristics of Being-In-Itself

First characteristic: “Being is in itself.” (That is, being-in-itself is in itself.)

The point here is basically a metaphysical claim. Sartre is in effect saying that being-in-itself has no cause. If it did have a cause, it would not be “in itself” but “in its cause.” (We sometimes speak of a cause as “containing” its effect implicitly, so that the effect is viewed as arising out of the cause.)

So the term “in itself” is meant to suggest something like “self-contained.”

In effect, then, Sartre is saying that being-in-itself (or matter) has no cause.

In part, this is a consequence of Sartre’s atheism. (See my discussion of “Existentialism Is A Humanism” for a fuller treatment of this.)

Why is Sartre an atheist? There are several reasons — or several levels of reasons. But one of them we can look at now. Basically, he thinks the existence of God would be incompatible with human freedom. So once again, Sartre still owes us an explanation of why he is so sure human beings are free. Note that, earlier, we saw that Sartre rejects the *Transcendental Ego* for basically the same reason: it would get in the way of human freedom.

Why is the existence of God incompatible with human freedom? In “Existentialism Is A Humanism,” Sartre gives a simple analogy: the analogy of the letter-opener.

The letter-opener is an artifact. It is designed by someone, and manufactured in accordance with that design. The design — a kind of blueprint or plan — in effect sets the limits to what the letter-opener is and what it can do. This plan or blueprint gives us the essence of the letter-opener. (Recall: The essence is the “principle of the series” of phenomena.)

This essence exists in a sense in the mind of the designer or artisan of the letter-opener before the actual object is produced or manufactured. And in that sense, for the letter-opener: *Essence precedes existence*.

And the same thing will obviously hold for anything that is designed or produced in accordance with a plan in this way. So not only does essence precede existence for the letter-opener in particular, essence precedes existence for all artifacts in general.

Now if God exists, as he is traditionally conceived, then he is a kind of divine artisan of the whole of creation — he produced the whole world as a product of his divine creative act. By the same token, therefore, there must exist in God’s mind a divine plan for creation and for everything in it.

This divine plan — traditionally known as “providence” — sets limits to what goes on in creation. Nothing happens fortuitously, nothing happens by chance. It was all foreseen by God.

In effect, now we’re squarely up against the good old traditional philosophical problem of how to reconcile divine foreknowledge with human free will. If God knows what we are going to do before we do it — and he really knows it, and is not just making a good guess — then how can we be said to be free in the matter any more? In other words, how can God exist as traditionally conceived (which includes his omniscience) if human beings are free?

Historically, there have been lots of attempts to explain this, to reconcile these two things. But Sartre thinks they all fail, and that the enterprise is impossible on principle. In the end, Sartre thinks, you can’t have both: If God exists, then human beings are not free. Conversely, if they are free, then God does not exist. But of course, Sartre is absolutely convinced human beings are free, for reasons he still owes us. And so, Sartre has to be an atheist.

In the actual essay, “Existentialism Is A Humanism,” Sartre goes on to argue the other way too: Not only is it the case that if God does exist, then human beings are not free. It

also goes the other way around: If God does not exist, then human beings are free. In the essay itself, the point of the illustration is to prove human freedom on the basis of atheism. In the essay, Sartre just assumes atheism, and argues on that basis for human freedom. But the way he sets it up, it will work the other way around too.

(For details, see my discussion of “Existentialism Is A Humanism” in the course packet.)

Now, to come back to being-in-itself: If God does not exist, he could not have created being-in-itself. In fact, nothing created being-in-itself. It is “self-contained” — “in itself.” This doesn’t mean it produced itself. It means it wasn’t produced at all. Being-in-itself is, more or less, eternal (although we shall see soon that this is not the correct way to speak about it). It is not caused.

This self-containedness is what Sartre’s first characteristic means, which he sums up in the slogan: Being is in itself.

Second characteristic (the third in Sartre’s own numbering): “Being is.” (That is, being-in-itself is.)

Sartre often puts the point here by saying that being-in-itself is “too much,” it is “superfluous.” That is, there is no good reason why it should be there, rather than not being there.

This is closely related to the previous characteristic: “Being is in-itself.” That first characteristic amounted to saying that being-in-itself has no cause. This second claim amounts to saying that being-in-itself has no explanation.

Whereas the former is a metaphysical claim, this one is more epistemological. An “explanation” is an account that satisfies the mind. It may be a causal explanation, or it may be some other kind of explanation. But in any case, this second claim says that being-in-itself has no explanation.

Thus being-in-itself is a violation of the “Principle of Sufficient Reason.” The “Principle of Sufficient Reason” says that there is a sufficient reason or explanation for everything. There must be some reason why things turned out the way they are, rather than some other way. Even if we don’t know what that reason is, there is one.

Sartre is in effect denying this. There is no sufficient reason for the existence of being-in-itself.

A “sufficient reason” would be a necessitating reason, one such that, given the reason or explanation, the existence of being-in-itself would necessarily follow. (Otherwise it wouldn’t be “sufficient,” would it?) But there is no such reason, so that the existence of being-in-itself is not necessary. It is contingent. It is just a brute fact.

Sartre sometimes expresses this by saying that the existence of the in-itself is absurd. To call something “absurd” for Sartre doesn’t mean it is contradictory (as ‘absurd’ meant for Kierkegaard), and it doesn’t mean it is just incongruous or ridiculous (as it meant for

Camus and sometimes in common parlance). It means it is “without ultimate explanation.”

In fact, anyone who admits that human beings are free must admit that there are some things that are contingent, and so must admit that there are exceptions to the Principle of Sufficient Reason. But Sartre is going further here. He is saying that not only are human actions free, and therefore contingent, and therefore violations of the Principle of Sufficient Reason. The existence of matter — being in itself — is another such exception to the Principle of Sufficient Reason.

Once again, this second characteristic may be regarded as a corollary of Sartre’s atheism. If God is not around to cause being-in-itself, then he is not around to provide any other kind of explanation for it either. And the idea seems to be that if there is any ultimate explanation, it is somehow going to have to involve God — or something that plays the same role as God, which comes down to the same thing.

It is worth pointing out here that Sartre’s atheism to some extent plays the same role in his philosophy as Nietzsche’s plays in his. When Nietzsche proclaims that “God is dead,” he doesn’t just mean a declaration of atheism in some narrow, technical sense. He means that everything the notion of God traditionally stood for, all the roles and functions — philosophical, psychological, sociological, etc. — the notion of God has played for us, all that is canceled, no longer applies. Sartre’s atheism to some extent works the same way. When he seems to be saying that any ultimate explanation would have to involve God somehow, he doesn’t just mean God in the theological sense, but a broader notion — God in the loose sense in which that concept has been used throughout Western thought.

Third characteristic: “Being is what it is.” (Once again, this means being-in itself is what it is.)

This one is perhaps the hardest one for us to grasp. Sometimes Sartre puts it by saying that being-in-itself is opaque, it is solid. Roughly, the point of all these metaphors is to say that being-in-itself is through and through positive. There is nothing negative about it.

What Sartre has in mind here is a doctrine that goes back to Parmenides the Presocratic, as commonly interpreted. (The relevant text is Parmenides’ Fragment 8.)

Parmenides began his philosophy with what he regarded as a principle that was absolutely demanded by reason: Being is — not in Sartre’s sense of this as a slogan, the “Second characteristic” we’ve just looked at — (and non-being isn’t). Being is what it is (and nothing else). That is, for Parmenides, reality is completely affirmative. There is nothing negative about it.

What Parmenides is saying is that there is something paradoxical about negative notions, something about them that is hard — perhaps impossible — to understand.

In effect, this means that it ought to be possible to describe the whole of reality without ever once resorting to the little negative word ‘not’, or to any other negative word that

implicitly has ‘not’ built into it. It ought to be possible simply to strike the word ‘not’ from our vocabulary.

Parmenides drew some conclusions from this line of thinking — conclusions *remarkably* similar to Sartre’s own conclusions about being-in-itself, as we shall see. Parmenides concluded that:

- (1) There is nothing negative about being.
- (2) There is no change in being, in reality. In order for something to change, it would have to change from what it is to what it isn’t, or vice versa. And that of course involves negation.
- (3) There is no coming to be. There is never anything new, no generation. Being could only come to be from non-being, and there isn’t any such thing. Being could not come to be from being, since it is already being — that would not be a coming to be at all.
- (4) By the same token, there is no destruction, no annihilation. Destruction would be a kind of change from being to non-being, and there’s no such thing as non-being.
- (5) There is no time. Time is made up of past, present and future. But the past doesn’t exist — any longer. And the future doesn’t exist either — not yet. They would be non-being. But there isn’t any such thing. And the present is just the limiting point separating the past and the future.
- (6) There is no differentiation in being. Being is not divided up into this being and that being, so that this being is not that being. If x is not y, then that is a kind of non-being — not-being-y, and non-being is not.

Thus, for Parmenides, reality is just one unchanging, timeless, undifferentiated, featureless blob of being. And that’s all.

Here is part of what Parmenides says (Kirk and Raven translation):

... it is uncreated and imperishable, for it is entire, immovable and without end. It was not in the past, nor shall it be, since it is now, all at once, one, continuous; for what creation wilt thou seek for it? how and whence did it grow? Nor shall I allow thee to say or to think, ‘from that which is not’; for it is not to be said or thought that it is not. And what need would have driven it on to grow, starting from nothing, at a later time rather than an earlier? How could what is thereafter perish? and how could it come into being? For if it came into being, it is not, nor if it is going to be in the future. So coming into being is extinguished and perishing unimaginable. Nor is it divisible, since it is all alike; nor is there more here and less there,

which would prevent it from cleaving together, but it is all full of what is.
So it is all continuous....

For Parmenides, the appearances to the contrary — the appearances of change, time, differentiation, generation, destruction — all those are just illusions. We shouldn't pay attention to them. They are the "Way of Deception."

Sartre is saying exactly the same things — but note that he is talking about only one-half of reality, what he calls being-in-itself. (There is also being-for-itself.) For example, he says on p. 29 in the "Introduction" to *Being and Nothingness*:

Transition, becoming, anything that permits us to say that being [that is, being-in-itself] is not yet what it will be and that it is already what it is not — all that is forbidden on principle.... It is full positivity. It knows no otherness; it never posits itself as *other-than-another-being*...it is not subject to temporality.

Sartre is in effect just listing here the same conclusions Parmenides came to. I cannot read this passage without thinking that Sartre had Parmenides in mind. (Actually, he probably more immediately had *Hegel* in mind, but I'll bet he thought of Parmenides too.)

But while Parmenides thought the appearances to the contrary were just illusions, and should therefore be ignored, Sartre has a more complicated view. Change, time, etc., are indeed appearances (phenomena), just as Parmenides said. But that doesn't mean we can ignore them. On the contrary, as phenomenologists, we must describe them carefully. But if we are going to do it correctly, we must, in view of the previous considerations about being-in-itself, introduce the notion of non-being into the picture. Where does this come from?

How do we account for the appearances of change, time, all these things that involve negativity? They cannot come from being-in-itself, as we have just seen.

You can guess what the answer is going to be: They will have to come from consciousness, from being-for-itself.

What Sartre is dealing with here is a classical philosophical problem of negation. Parmenides worried about it. And it is still with us. Bertrand Russell, in his lectures on Logical Atomism, for example, worried about the status of negative facts.

If I open the refrigerator and see a milk carton there, that is a positive, affirmative fact: There is a milk carton in the refrigerator.

But if I open the refrigerator and find no milk carton, then does that mean that there is something called "milk-carton absence," "non-milk-carton" there? Russell thought there was something bizarre about that, and then went on to give his own theory of what such negative facts amount to. Sartre is in effect worrying over the same point (although his answer is going to be completely different).

We can perhaps get a better idea of this third characteristic of being-in-itself if we return to our theater model. Being-in-itself is the screen, recall. But the screen is — all by itself — completely featureless; it is blank. It does not change — change is what we see in the movie projected on the screen. There is no time, in a sense. Time is what we see unfolding in the story on the screen.

There is no differentiation. The screen is completely homogeneous. It is not divided up into John Wayne and the wagon train and the Rio Grande. All that is what appears on the screen.

Now that we have the notion of being-in-itself, I must warn you about Sartre's terminology. Sometimes Sartre talks about being-in-itself as though it were "the world." When he talks that way, it sounds as if the dichotomy between being-for-itself and being-in-itself just amounts to the dichotomy between consciousness and the world, in the sense of non-conscious things: trees, stones, automobiles, pencils, etc. In fact, sometimes Sartre calls such things "beings-in themselves" (in the plural).

But how can he do that, given what we've just seen him say? In order to speak about "beings-in-themselves" in the plural, we would have to allow that one being-in-itself is distinct from another. But didn't he just say that being-in-itself is undifferentiated?

Besides, the things he calls "beings-in-themselves" certainly don't look like being-in-itself as he has just described it. Automobiles, for example, are conspicuously "differentiated" things — they have lots of little parts in them. Moreover, they change, break down, depreciate with time, etc.

So is Sartre just guilty of equivocation here? Is he going back on what he has said about being-in-itself?

Not exactly. Recall that another way to think of being-in-itself is on an analogy with Aristotelian matter.

Aristotelian matter is the kind of neutral stuff on which you impose a form, yielding a product. For instance, think of bronze as a kind of matter. (The example is Aristotle's own.) It turns out that our own use of the word 'bronze' in English parallels Sartre's use of the expression 'being-in-itself'. Here's how it goes:

You never see just plain bronze — just as you never see pure being-in-itself. You always see bronze that is shaped (it has some shape or other), a bronze coin or a bronze statue, or even just a bronze lump — a product of matter and form. So too you are conscious only of being-in-itself that has been processed by consciousness (a picture has been cast on the screen).

We can describe bronze — that is, bronze just all by itself. All by itself, bronze has no particular shape. The shape is always something added on. By itself, bronze is just a certain kind of stuff.

Nevertheless, when we see a coin or a bronze statue, we can (and do) say "This is a bronze." And if we see two of them, we can say "Here are two bronzes." (For example,

The Royal Ontario Museum in Toronto has an excellent display of Shang “bronzes” from ancient China.)

That is, what we are talking about when we talk about such “bronzes” is surely bronze. It’s not cellophane, after all, but bronze. But what we have is not just pure bronze; it’s not bronze all by itself. It’s bronze shaped and molded in a certain way.

This is exactly the way Sartre talks about being-in-itself. He can describe pure being-in-itself, so to speak. (That’s what we just saw him do with his three characteristics.) But he will also allow that we can talk about beings-in-themselves — tables, automobiles, etc., which are not pure being-in-itself, but being-in-itself that has been molded and processed in various ways.

Being-For-Itself

Let’s now turn to being-for-itself — that is, to consciousness, the peculiarly human kind of reality.

We may as well get used to it now: Sartre identifies human beings with consciousness. Consciousness is not something human beings have; it is what they are. Of course, that doesn’t mean he thinks of human beings as pure intellects (like angels). On the contrary, for Sartre consciousness goes much further than the intellect. Consciousness is also our fears, hopes, wishes, desires, emotions, memories — even our bodies. We will see more of this later on. But get used to it now.

So while Sartre is a dualist in general, he is not a dualist about human beings. Historically, most dualisms start off with a dualism about human beings — they are souls and bodies, minds and bodies — and then extrapolate from that to a dualism about their ontology at large. Sartre doesn’t proceed that way.

For the present, we can say that consciousness or being-for-itself has the opposite of two of the three characteristics we have just described for being-in-itself. (The remaining one will require a little more discussion.) Let’s run through those three once again.

First characteristic: Being-(in-itself) is in itself.

But Being-for-itself is not in-itself. In the case of being-in- itself, this meant it was not metaphysically caused by anything, it did not depend causally on anything else. Thus, to say that being-for-itself is not in-itself means that it is caused, it does depend on something else.

And what does it depend on? Well, what else is there? It depends on the in-itself. Sartre describes the for-itself as “arising” out of the in-itself, it “surges up,” as he says in “Existentialism Is A Humanism.”

In short, consciousness depends on matter. Without matter, there would be no consciousness. This doesn’t mean that consciousness is itself a material process, or that it

Copyright © 1996 by Paul Vincent Spade. All rights reserved. Permission is hereby granted to copy this document in whole or in part for any purpose whatever, provided only that acknowledgment of copyright is given.

can in any way be reduced to matter. But it remains the case that without matter there would be no consciousness.

Why does Sartre think this is so? Well, in the end, he thinks it follows from the notion of consciousness as intentional. Since consciousness is always consciousness of something other than itself, that something else is going to have to be being-in-itself. There just isn't anything else.

Of course, consciousness brings a certain contribution of its own to what it is conscious of — it constitutes the object in part (not entirely, as for Husserl). What an act of consciousness does in every case, then, is to take up being-in-itself and mold it into a particular kind of object, interpret it as a tree, a table, etc. Similarly, the sculptor takes up the bronze or clay and molds it into a statue. Similarly too, to use our movie-theater model again, the light rays in a sense grab hold of the screen (“take it up”) and transform it into a screen epic. The point is that it is the screen that is operated on here — being-in-itself.

In terms of the movie-theater model, the point can be put quite clearly: No screen, no movie.

Thus, consciousness by its very essence — which we get by eidetic abstraction, and which requires that consciousness be intentional — requires something other than consciousness; it requires being-in-itself.

This move is what Sartre somewhat mischievously calls his “ontological argument” in the “Introduction” to *Being and Nothingness*. (The original “ontological argument” was of course Anselm’s argument for the existence of God from the definition of God. So too here: We can see from the very definition of consciousness that it requires the existence of being-in-itself. The connection with the original ontological argument is pretty strained, to be sure.)

Problem for Sartre: In “Existentialism Is A Humanism,” Sartre argued that for human beings — that is, for the for-itself — “existence precedes essence.” That is, consciousness has no definition or “essence” given to it in advance. On the contrary, “Man makes himself” — he defines himself only in the process of living. Consciousness is free, not confined to the limits of a definition or nature.

But now we see Sartre claiming that the for-itself is not “in-itself,” and claiming this on the basis of the essence or definition of consciousness in terms of intentionality. What about this?

This will be a recurring puzzle for us. Sartre strongly emphasizes human freedom, with the correlative lack of definition or essence for consciousness. And yet he goes on to give us all kinds of general principles about consciousness.

There are two ways to regard this. We can say that Sartre is just being inconsistent, and let it go at that. In that case, we can all go home and not pay more attention to this man.

Or we can look more deeply. Why would a reasonably intelligent man proceed in this way that appears so obviously inconsistent? If we are going to make sense out of this, we are going to have to rethink the terms of the problem. This seems to me the only honest way to proceed — provided our point is genuinely to understand Sartre, and not just to dismiss him with quick “refutations.” (We don’t have to hold Sartre in any special awe to do this.)

We will be returning to this problem repeatedly throughout this course. For the present, let’s just observe that none of the general laws Sartre is laying down for consciousness must be allowed to get in the way of human freedom. Exactly how this is going to work out we will have to wait to see.

Back to the point: Unlike being-in-itself, consciousness is not in-itself; it is, so to speak, “in-another.”

Consciousness depends on being-in-itself. But we should not think of this as though being-in-itself somehow produced consciousness. Being-in-itself doesn’t do anything, remember. It is completely inert.

So how do we get from being-in-itself to being-for-itself? Sometimes Sartre talks in almost mythical terms about this. He speaks of consciousness as surging up. At several points he speaks of a kind of explosion of being-in-itself. But none of this can be taken at face value.

In fact, it is wrong to think of this in temporal terms at all, as though first there is only being-in-itself, and then, later on — bingo — there is being-for-itself too. This is wrong, because for being-in-itself there is no “first” and “later.” Being-in-itself is not subject to time, remember.

Sartre struggles with this problem. It is one of the recurring themes throughout *Being and Nothingness*. How to express this relationship adequately and not misleadingly?

Second Characteristic: Being-in-itself just is.

This is the second (the third in Sartre’s presentation) characteristic of being-in-itself. We have seen what this means: Being-in-itself is just a brute fact. There is no reason why matter should exist rather than not exist. It violates the Principle of Sufficient Reason.

Now this is where I said that the for-itself has the opposite of two of the three characteristics of the in-itself, but not of the third one. The for-itself, like the in-itself, just is. There is no ultimate “sufficient reason” for the existence of consciousness any more than there is for the existence of being-in-itself. Consciousness too is “absurd” for Sartre.

On the other hand, the situation with consciousness is not exactly the same as we had with the in-itself. Since consciousness arises out of the in-itself somehow (as we just saw), it is not altogether without any reason or grounding. It is not totally contingent. There is a certain limited kind of necessity forced on consciousness from the outside.

This kind of “limited necessity” is what Sartre calls “facticity.” (In “Existentialism Is A Humanism,” he calls it — or something more or less *like* it — the “human condition.”) It will be an important theme of *Being and Nothingness*.

In part, the basic idea behind “facticity” is that, while I am free to choose this or that, I am not free not to choose. I didn’t ask to exist, no one consulted me. But I do exist anyway, and so have to choose. Even if I choose to commit suicide, so that I will not have to choose any more, that is still a choice. As Sartre puts it in “Existentialism Is A Humanism,” “Man is condemned to be free.”

This “facticity” is going to be something Sartre will try to account for in terms of the fact that the for-itself depends on the in-itself. This will be a long story.

The Third Characteristic: Being-in-itself is what it is.

The third characteristic of being-in-itself (third in our own ordering, second in Sartre’s) is that the in-itself is what it is, and is not what it isn’t. By contrast, the for-itself, Sartre says, is not what it is, and is what is it not. This is ABSOLUTELY CRUCIAL.

The claim is of course deliberately paradoxical. Some people (for example, Arthur Danto in his book *Jean-Paul Sartre*) think Sartre is just being perverse here, and that he doesn’t literally mean what he says. They think what Sartre has in mind can be explained more perspicuously in terms that are not out and out contradictory.

I think that is wrong. I think Sartre means this to be out and out contradictory. And is absolutely crucial to see why.

To say that consciousness — the “for-itself” — “is not what it is and is what it is not” means in the end that consciousness — the “for-itself” — has negation, non-being, “nothingness” (note the title: *Being and Nothingness*) all through it. As he says at one point, “consciousness is its own nothingness.”

But why should that commit us to outright contradiction? Well, it is because Sartre accepts the Parmenidean view that there is something contradictory about the little word ‘not’. Parmenides, recall, started from the principle required by reason — what is is, and what isn’t isn’t, and never the twain shall meet — and drew the conclusion we have seen: there is no change, no time, no differentiation. Change, time, differentiation, are all contradictory. They violate the principle required by reason. They all mix being with non-being in various ways.

Now in a sense, Parmenides is obviously right: Philosophers have known for centuries that change, time, the fact that things are differentiated from one another — all these are very mysterious things. The fact that they are altogether familiar makes them no less mysterious as soon as we begin to look carefully.

There have been many philosophical attempts to get around the Parmenidean reasoning, to find some way to make change consistent, to make the notion of time consistent.

For example, Aristotle gave an analysis of change in terms of what he called “potentiality” and “actuality.” (And he did this precisely in order to avoid the Parmenidean problem.) The acorn changes into the oak tree, develops into the oak tree. For Aristotle, this is analyzed by saying that the acorn is actually just an acorn, but potentially an oak tree. The fact that the acorn changes into the oak tree is a matter of that potentiality’s being actualized. On the other hand, the acorn is not potentially a carrot — which is why when you plant an acorn, you get an oak tree and not a carrot.

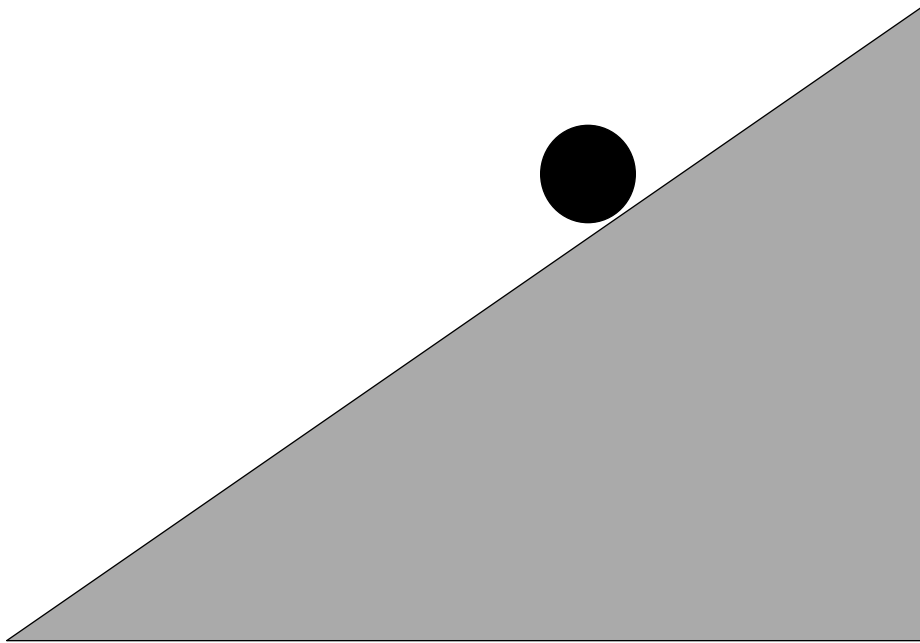
But isn’t that just to say that the acorn really isn’t an oak tree — not actually — and yet in a sense it really is an oak tree — potentially. What are the words ‘actually’ and ‘potentially’ doing here? They don’t really avoid the inconsistency Parmenides thought he saw in the notion of change. They just disguise it. What you are saying is that the acorn both is and is not the oak tree. The qualifications ‘actually’ and ‘potentially’ just cover over this blatant contradiction — because, after all, we don’t have an independent account of what the difference between actuality and potentiality is supposed to be. It’s as if we were saying:

There isn’t any contradiction, because we have to distinguish two senses of being the oak tree: sense A and sense B. In sense A the acorn is an oak tree; in sense B it isn’t.

In the absence of any other information about what the two senses A and B are, this is not a very successful way to avoid the contradiction. It’s just to pretend we’re avoiding it, because we don’t like it. And, on Sartre’s view, that is exactly the sort of thing we have with Aristotle’s theory of potentiality and actuality — and with all other philosophical attempts to avoid these paradoxes.

Sartre thinks that happens on principle — that there is no consistent account of these things. They really are contradictory.

Another way to look at this was suggested to me once by a remark one of my colleagues made (Norman Kretzmann) in a totally different context. Suppose you have a sphere rolling down an inclined plane:



And suppose you take an instantaneous snapshot of it, as above. Then you show it to someone else and ask him to describe what he sees there. If he just describes the inclined plane, and the sphere, and the position of the sphere at the moment the picture was taken, then everything he says is correct — as far as it goes. He's described what's there in the picture, all right. But what he's not described is the motion, the fact that the ball is rolling down the plane.

In order to describe the ball as rolling, you can't just talk about where the ball is right now, when the snapshot was taken. You have to make some kind of reference to where it was a moment ago, or to where it is going to be a moment from now. In other words, in order to describe the situation in terms of change and process, you have to talk not just about what it is, but about what it isn't as well. The situation, the rolling down the plane, then, is not just what it is, but also what it isn't.

All of this is just to say that Parmenides was right! Change, time, motion, differentiation, all involve a mixture of what is and what isn't, and are contradictory in this way.

Now back to Sartre. Since, as we saw in discussing being-in-itself, being-in-itself cannot be the source of these positive-negative contradictions (being-in-itself is purely affirmative, remember, there is nothing negative about it), then they must somehow be accounted for in terms of being-for-itself, in terms of consciousness. That's what the theory of *constitution* says, after all. (There isn't any third alternative, since these two are the only things that are, for Sartre.)

Now this tells us something important about consciousness. If consciousness too were completely positive, it would be just like being-in-itself in that respect, and consciousness could no more account for these negative, contradictory things than being-in-itself can.

Copyright © 1996 by Paul Vincent Spade. All rights reserved. Permission is hereby granted to copy this document in whole or in part for any purpose whatever, provided only that acknowledgment of copyright is given.

Therefore, consciousness is not purely positive and affirmative. It must somehow have negation running all through it. And that means, of course, that consciousness (like change, motion, etc.) is going to be contradictory too.

And of course, in a sense that is something else philosophers should have realized a long time ago. People have always known there is something very mysterious about consciousness. No one has succeeded in giving a coherent account of it.

Now *here is where Sartre's phenomenological method pays off*. If he were trying to argue, to proceed according to what Husserl called "the natural standpoint," to construct a "scientific" theory of consciousness and of the world and of the relation between them, then he would be in a bad fix at this point. He would start from certain claims about consciousness and the world, and then deduce the consequences from them. But, as everyone knows, if your initial premises are contradictory, then you can deduce anything at all from them. It is a standard theorem of logic: Anything follows from a contradiction. (Barring various paraconsistent logics, relevance logics, and other exotica. It is perhaps interesting to speculate about the usefulness of such logics in Sartre's own enterprise.)

So if Sartre were proceeding from "the natural standpoint," if he were constructing an argumentative theory, why then anything would go. His theory would commit him to saying absolutely everything. In short, his whole enterprise would break down.

But of course Sartre is not arguing; he is describing. As a phenomenologist, he cannot argue for his views. But now, if change, becoming, differentiation, etc., are all contradictory, that doesn't mean I cannot describe them anyway. I can describe contradictory things. I can describe what a square circle would be like, for example. (If I couldn't, how could I ever decide there aren't any such things.) Of course my description of a square circle is going to be contradictory. But that doesn't matter.

And so too with change, time, differentiation — and with consciousness. They are all contradictory, but I can nevertheless describe them. And there is nothing wrong with that procedure, provided I don't try to infer anything from these descriptions in an argumentative manner.

So what happens to logic, for Sartre? After all, whether Sartre himself actually wants to draw the inferences or not, it would seem to be just a simple law of logic — the Law of Identity — that says that things are what they are, and are not what they aren't. So when Sartre says the for-itself is not what it is and is what it isn't, isn't he just denying a basic principle of logic?

Yes he is. But that doesn't mean that logic has no legitimate role at all for Sartre. In the "Introduction" to *Being and Nothingness* he explicitly describes the Law of Identity as what he calls a "regional principle." That is, it applies to only one region of reality — to being-in-itself. It does not apply to the for-itself.

So here is where Sartre parts company with Parmenides. Parmenides thought that negation is contradictory, and so should be abhorred and avoided — ignore all the appearances of negative things, they will only mislead you. Sartre agrees that negation is

contradictory, but does not agree that we can therefore avoid it. On the contrary, we must be especially careful, must pay close attention to it, to make sure we describe it accurately.

Furthermore, we should not think that calling something contradictory means it doesn't exist for Sartre. We have just seen him call consciousness, the for-itself, contradictory, and it certainly does exist. For Sartre, there are real contradictions (and they are us). As a terminological convenience, therefore, let us for future reference distinguish the contradictory, which might in fact exist, from the impossible, which can't exist.

Question for pondering: How are we supposed to tell which is which, which contradictions are “good contradictions” and can exist anyway, and which ones are “bad ones” and cannot exist? I suspect this is going to have to have something to do with eidetic abstraction, but I'm not sure of the details.

Positional & Non-positional Consciousness, Reflective & Non-Reflective Consciousness

So far, we've explained how the for-itself is like and how it is unlike the in-itself with respect to the three characteristics of the in-itself Sartre listed at the end of his “Introduction.”

But there is another aspect of consciousness we have yet to consider. It is what is suggested by the peculiar phrase ‘for-itself’. What is “for-itself” about consciousness?

This brings us to one of the main themes of Sartre's *Transcendence of the Ego*, and of the “Introduction” to *Being and Nothingness*. (The discussion is clearer in *Transcendence of the Ego*, pp. 43ff.)

There are two pairs of terms discussed there, and it is important to get them straight:

Positional (or “thetic”) vs. non-positional (non-thetic) consciousness.

Reflective vs. non-reflective (unreflective, pre-reflective) consciousness.

The first thing you have to know about these two distinctions is that they are not the same distinction. The secondary literature often treats them as though they were the same distinction. For example, Danto's book *Jean-Paul Sartre*. As I understand it, this is a mistake — and an important mistake.

Let's start with the second pair first: reflective vs. non-reflective (unreflective, pre-reflective) consciousness.

Suppose I am reading a book, a gripping murder mystery. In that case, what is the object of consciousness? (There must be one, as required by the theory of intentionality.)

Answer: The story, the murder. That entirely occupies my consciousness; it and it alone is my object. In fact, the more I am “caught up” in the story, the more I tend to lose any

awareness of my environment. I ignore the fact that the refrigerator is buzzing, I filter out the traffic noises outside the window, etc. All I am conscious of is the story.

In particular, *in no sense* is the object of my consciousness in this case me. On the contrary, to the extent that the story is a good one, we say that I have “lost myself” in the story; I am “absorbed” in the story.

In other words, the object of consciousness in no sense involves me.

This is what Sartre calls “non-reflective” (pre-reflective, unreflective) consciousness.

But now suppose I suddenly say to myself: “I’m really enjoying this book.” What has happened there? All of a sudden, “I” appear. Now I’m thinking about myself and my enjoyment of the book. Now the object of my consciousness is me — or at least it’s a situation of which I am a part. And that wasn’t so for my earlier consciousness.

In fact, there is a sense of almost physical “wrenching” involved in passing from the one state of consciousness to the other.

This second act of consciousness, the one in which I am thinking about myself, or about some situation in which I am involved, is what Sartre calls “reflective” consciousness. The previous act, in which the object of my consciousness was just the story and in no sense me, is what he calls “non-reflective” consciousness (sometimes called “pre-reflective” or “unreflective” consciousness).

Just what kind of “me” it is that appears in reflective consciousness — whether it is the psyche (the “psychological ego”) or the Transcendental Ego, or what — is something we will have to look at later on.

On pp. 48–49 of *Transcendence of the Ego*, Sartre gives some examples of this non-reflective consciousness:

... When I run after a streetcar, when I look at the time, when I am absorbed in contemplating a portrait, there is no I. [That is, there is no “I” as part of the object of consciousness.] There is consciousness *of the streetcar-having-to-be-overtaken*, etc., [I am leaving out some words that refer to a distinction we have not yet discussed.] In fact, I am then plunged into the world of objects; it is they which constitute the unity of my consciousnesses; it is they which present themselves with values, with attractive and repellent qualities — but me, I have disappeared; I have annihilated myself. There is no place for me on this level. And this is not a matter of chance, due to a momentary lapse of attention, but happens because of the very structure of consciousness.

But now (Sartre doesn’t go on to make this point, but we can do it for him), suppose I stop running after the streetcar, and say “Oh, I’m not going to catch it.” At that moment, I am no longer aware of just the streetcar-having-to-be-caught; the “I” makes its appearance. At that moment, I move to the reflective level.

Important: Reflective consciousness and pre-reflective consciousness are mutually exclusive. They are two distinct acts of consciousness. No one act can be both.

On the other hand, every act of consciousness is either the one or the other. And how do you tell which it is, whether any given act of consciousness is reflective or non-reflective? You check what the object of that act of consciousness is. If it involves the “self,” the act is reflective; if not, it is non-reflective.

Thus, we get a Great Law of Consciousness:

Every act of consciousness is either reflective or non-reflective, but not both.

This much is easy. Let’s now look at the second pair of terms: ‘positional’ vs. ‘non-positional’. The first half of this pair is easy too:

The notion of positional consciousness is just an aspect of the doctrine of intentionality. Every act of consciousness is consciousness of something. Every act of consciousness takes an object, posits an object. And in that sense Sartre calls it “positional” consciousness.

(The term ‘positional’ here has nothing especially to do with space or location, at least not literally. It has to do with “positing.”)

Sartre also calls this notion “thetic” consciousness. ‘Thetic’ is just Greek, where ‘positional’ is Latin. In Greek, a ‘thesis’ is a “putting.”

So every act of consciousness is “positional” in this sense; it satisfies the doctrine of intentionality and has an object. Thus positional consciousness is not just one kind of consciousness, as opposed to other kinds, in the way reflective and non-reflective consciousness are two different kinds of consciousness, two distinct acts. No, every act of consciousness is positional.

So when we now turn to talk about something called “non-positional” consciousness, we are not talking about a new and different kind of act of consciousness. We are instead talking about a new side of consciousness that goes along with its being positional.

In other words, every act of consciousness will indeed be positional. But that is not the end of the matter. To say it is positional does not exhaust everything you can say about an act of consciousness.

As it will turn out,

Every act of consciousness is both positional consciousness of some object or other, and also non-positional consciousness. (And, of course, depending on what the object is, it will be either reflective or non-reflective.)

This may be called a Second Great Law of Consciousness.

But what is this “non-positional” side of consciousness? This is where we get the notion of what is “for-itself” about consciousness. The non-positional side of consciousness is self-consciousness. And by that we of course don’t mean being shy, but rather the awareness that each act of consciousness has of itself.

But now of course we have to be careful. This doesn’t mean that every act of consciousness is consciousness of itself as an object. It isn’t. And in fact, no act of consciousness is conscious of itself as an object. We saw that when we were discussing the theory of intentionality: Intentionality is irreflexive. Every act of consciousness is conscious of something else besides itself. In some cases that something else will be a different act of consciousness, and in those cases we have reflective consciousness. But the fact remains, no act of consciousness is consciousness of itself as an object.

So every act of consciousness is aware of itself — but not as an object. Thus there must be some other way in which an act of consciousness is conscious of itself. How are we to understand this?

Go back and think of an example we used before, the case of reading a good book. In such a case, I am “absorbed” in the book. I “lose myself” in it, as we say.

Yet, I am not “absorbed” in the story to the extent that I literally become the story. There is still a distinction between me and the story. The story is the sole object of my attention; it absorbs my attention entirely. And yet there is always a certain distance between me and the story. I am not literally the story; I am an onlooker, a spectator, an outsider. I adopt a vantage point with respect to the story. In fact, a skillful author will be able to manipulate the vantage point you adopt.

If the story is a third-person narrative, you adopt the vantage point of an invisible spectator who is present secretly at everything that is done in the story. But you are not doing what is being done. You are just a spectator.

If the story is a first-person narrative, things are different — but the main point still holds. You are still a spectator. You are not the person who is telling the story, the one speaking in the first person. You take up a vantage point inside his head, behind his eyes.

This is true in every case of consciousness. There is always a distinction between consciousness and its object. (Remember the irreflexivity of intentionality.)

Now there is a sense in which you are perfectly aware of this all the time, throughout the entire story. But of course it’s not part of the story. It doesn’t occupy your attention the way what is going on in the story does.

This is very much like the old notion we saw in Husserl — the notion of the phenomenological Ego, the pure “vantage point,” the “point of view” on an object. We never see the point of view, the “eye of the camera,” but it is always given (we never have to infer it) in everything we do see.

This way in which consciousness is always aware of its own point of view on the object is what Sartre calls non-positional consciousness. I am positionally aware of the object, and

non-positionally aware of my point of view on the object, my standpoint with respect to it.

And now we get an important move: Sartre thinks of consciousness as just this taking of a point of view. This is probably the most revealing way of thinking of consciousness for Sartre.

Consciousness is the taking of a point of view on an object. Not something that takes a point of view on an object; it is the taking itself. (There is no Transcendental Ego back in there, remember.)

As a result, since every act of consciousness is always positionally aware of the object and non-positionally aware of taking a certain point of view on the object, and since the taking of that point of view just is the act of consciousness, we get the following important statement:

EVERY ACT OF CONSCIOUSNESS IS SIMULTANEOUSLY
POSITIONAL CONSCIOUSNESS OF AN OBJECT AND NON-
POSITIONAL CONSCIOUSNESS OF ITSELF.

Sartre sometimes expresses the notion of non-positional consciousness by saying that every act of consciousness is consciousness (of) itself — putting the ‘of’ in parentheses to indicate that it is not consciousness of itself as an object. In English, we don’t have to resort to this. We can say “self-consciousness,” and don’t have to use the perhaps misleading construction “consciousness of self.”

Sartre thinks this non-positional side of consciousness is just as essential as the positional or intentional side of consciousness is.

If there could be such a thing as an act of consciousness that was purely intentional or positional, and was not non-positionally aware of itself, this would mean that there could be an act of consciousness that is consciousness of some object, but doesn’t take a point of view on that object, doesn’t separate itself as an onlooker at a distance from its object.

In that case, the consciousness would not only be absorbed in its object, in the sense that the object occupies its entire attention. It would literally become the object of consciousness. The distance, the separation, between consciousness and its object would vanish.

When I perceive a tree, I would no longer be conscious of the tree; I would literally be the tree.

Sartre thinks such an act of consciousness is nonsense. We can’t have a consciousness like that. It would be an unconscious consciousness, which is preposterous. An unconscious consciousness is no consciousness at all.

Note: This tells us already that Sartre is going to reject the Freudian notion of the unconscious mind. Sartre thinks that is utter confusion. This will be an important theme.

There is nothing unconscious about consciousness for Sartre. Consciousness is consciousness through and through. It is completely transparent, like light.

Note that Sartre is not really arguing here, although it may sound as though that's what he is doing. He cannot argue like this. He cannot say that the notion of an unconscious mind is contradictory and therefore must be rejected — because Sartre himself thinks the notion of consciousness is itself contradictory, as we have seen, and yet he does not reject it for that reason. No — Sartre is simply describing. And he thinks the Freudian notion of the unconscious is simply a misdescription of the way things are.

We have then two realms of being for Sartre: consciousness, the for-itself, and inert being-in-itself. The latter is the opposite of consciousness in some respects. It is unconscious. And in fact it is the only thing that is unconscious for Sartre. The attempt to find an unconscious part of consciousness is an attempt to insert a little dab of the in-itself right into the heart of the for-itself. And it won't work.

Sartre thinks Husserl's Transcendental Ego is like this. The Transcendental Ego is not really conscious; it is what produces conscious acts. It is out of place in consciousness. It is an opaque object in this transparent medium, "like a stone at the bottom of a pool," as he says in *Transcendence of the Ego*.

Sartre thinks it is wrong and contradictory to try to combine the in-itself and the for-itself in this way — and worse, it is simply a misdescription of the facts.

This is also one of the reasons Sartre is an atheist (although not the main reason). The notion of God is the notion of something that is simultaneously being-in-itself and being-for-itself.

God is an in-itself insofar as he is traditionally thought of as eternal, changeless, simple (having no parts), outside time. Yet he is also for-itself insofar as he is traditionally thought of as conscious, as having a providential plan, as having knowledge and will, as caring about creation.

Sartre thinks this is contradictory. And it's not just contradictory. (After all, consciousness itself is contradictory.) It's contradictory in a way that can't happen. Basically, Sartre's view is that by investigating the nature of consciousness and the nature of being-in-itself, he can tell that they just don't go together. And where does he find out about these natures? By the eidetic reduction. The eidetic reduction is still important for Sartre, although it does not play the same role that it did for Husserl.

The Freudian unconscious, the Husserlian Transcendental Ego, the traditional notion of God — Sartre rejects all these things for basically the same reason. They all would involve an impossible combination of being-in-itself and being-for-itself. In fact, Sartre sometimes talks as if the Transcendental Ego would be like a little "god" back there in consciousness, generating our acts of consciousness and responsible for our actions. Just like God (in "Existentialism Is A Humanism"), the Transcendental Ego would get in the way of our freedom.

The Self-Love Theory

Now what is the importance of this distinction between positional and non-positional consciousness, and between reflective and non-reflective consciousness?

Well, Sartre thinks a lot of philosophical mistakes arise from getting these distinctions confused, from thinking they are the same distinction. (As Danto in effect does, in my opinion.)

As an illustration, let's discuss the theory Sartre describes as the "Self-Love Theory." Sartre treats this in *Transcendence of the Ego*, beginning on p. 54, in the section entitled "The Theory of the Material Presence of the Me." (We'll talk later on about the odd distinction Sartre tries to make between the "I" and the "Me.")

The theory he is discussing here is a common one, and you've no doubt heard it. It goes like this:

My friend Pierre is in need in some way. Say he's slipped and fallen on the ice. And so I go to help him up. At first this looks like a nice, friendly, virtuous thing to do. After all, we're supposed to help our friends on such occasions, aren't we?

But the Self Love Theory would say this superficial appearance is not right. What's really going on is not virtuous at all. Seeing Pierre in need produces in me a certain feeling of distress and anxiety, and that's of course unpleasant. Furthermore, if I go to Pierre's aid, this will reduce my feeling of distress, and may even produce in me a certain sense of self-satisfaction for being so helpful, and that would be pleasant. And I anticipate these things if I go to Pierre's aid.

Now the Self Love Theory claims that my real motivation for going to Pierre's aid is not to help him out, but rather to reduce my distress and give myself the satisfaction of being so helpful.

In short, the real reason I go to Pierre's aid has nothing to do with him; it's a deed that's entirely self-centered. It's selfish.

If you reply that I certainly don't seem to be acting selfishly in that case, that I'm certainly not aware of being prompted by a selfish motive, the Self Love Theory would reply that that doesn't prove anything. You still were acting selfishly, even if only unconsciously.

Now of course the case of Pierre's need is only an example. The general point of the Self Love Theory is that we always act for selfish motives in this way; there's no avoiding it.

Notice the move here. This theory argues as follows:

- (1) Every action is of course directed at some goal or other. Without some purpose, we would never act at all. (Random twitchings do not count as actions.) And furthermore, that goal must always involve the object of consciousness — reaching it, changing it,

etc.

(ACTION → GOAL INVOLVES INTENTIONAL OBJECT)

- (2) Now (this theory says) every action of ours is directed at the Self or Ego. It's goal is always a selfish one. (The "Ego" involved here would be the psychological Ego, the seat of our desires and pleasures, of our distresses and satisfactions.)

(GOAL → SELF)

- (3) Therefore: It appears to follow that every action of ours involves a reflective act of consciousness — an act of consciousness that has the Self or Ego as its object or as part of its object.

(∴ ACTION → REFLECTIVE)

- (4) But it doesn't seem that this is so. We have the examples of helping Pierre, chasing the bus, etc.

(DOESN'T SEEM SO.)

- (5) Therefore: Since I'm not consciously being selfish, my selfishness in those instances must be unconscious.

(∴ UNCONSCIOUS)

Notice how the notion of an unconsciousness is being appealed to here in order to save a theory that is not supported by the facts of consciousness. (I'm not being consciously selfish.) The data of consciousness seem to conflict with the Self Love Theory. But, rather than rejecting the theory, its proponents just move the whole thing into the unconsciousness.

In the above argument, there are three premises: steps (1), (2) and (4). (Steps (3) and (5) are conclusions drawn from the preceding steps.)

Now Sartre will accept the premises in steps (1) and (4). But why on earth would anyone accept the premise in step (2) — the main thesis of the Self Love Theory?

Sartre thinks the plausibility of step (2) rests on a correct insight combined with a mistaken assumption.

The correct insight is that the sight of Pierre in need does produce a feeling of distress in me. And I am quite aware of that feeling of distress. (After all, a feeling I'm not aware of is like a pain I don't feel — it's no pain at all.)

Furthermore, it is correct that when I go to Pierre's aid, my action reduces my distress, and perhaps even enhances my feeling of self-satisfaction. And again, this reduction and this enhancement are things I'm quite aware of.

Again, it is no doubt correct that there is a sense in which I anticipate this reduction of my distress and this enhancement of my self-satisfaction if I go to Pierre's aid. After all, I'm not surprised to find I'm no longer so distressed when I help Pierre up. And again, this anticipation must be something I'm aware of.

So Sartre grants that I am aware of all the things the Self Love Theory says I'm aware of.

But he thinks these correct points are combined with a mistaken assumption. The assumption is that the only WAY I can be aware of all these things is POSITIONALLY.

In short, the Self Love Theory gets its plausibility from ignoring the non-positional side of consciousness, and thus overlooking the distinction between positional and non-positional consciousness. It thinks that since I am aware of all these things that involve myself, and since (it implicitly supposes) the only I can be aware of things is positionally, it follows that what I am doing in these cases is reflecting. And since this theory thinks that all my actions are motivated by selfish considerations in this way, I am constantly on the level of reflection.

But of course that's wrong, and the mistake comes from not recognizing that there is another way of being aware of myself besides reflection. And that is: non-positional consciousness. And of course we are always aware of ourselves that way. (Recall: Every act of consciousness is both positional consciousness of an object and non-positional consciousness of itself.) But that doesn't mean we are always on the reflective plane.

Here is what Sartre says on p. 55:

Now the interest of this thesis [the Self-Love Theory], it seems to us, is that it puts in bold relief a very frequent error among psychologists. The error consists in confusing the essential structure of reflective acts with the essential structure of unreflected acts. It is overlooked that two forms of existence are always possible for consciousness. Then, each time the observed consciousnesses are given as unreflected, one superimposes on them a structure, belonging to reflection, which one doggedly alleges to be unconscious.

Sartre's own view is that I am positionally and unpleasantly aware of Pierre in need, and non-positionally aware of being unpleasantly aware. I am not reflecting on my displeasure. (Of course I always could reflect on it, but there's no theoretical reason why I have to.)

In short, for Sartre, the object before my mind in this case is Pierre in need. It's in no way me. The goal of my action is to change the situation, to get Pierre out of need. The goal is not to reduce my displeasure, even though that will happen as a result, and even though I am quite aware that it will happen.

Note: In Sartre's account, there is nothing unconscious about any of this. Once we recognize the distinction between positional and non-positional consciousness, we don't need to postulate an unconsciousness.

Question: Is this anything more than a mere terminological point? Is what Sartre calls "non-positional" consciousness really just what others call an unconscious mind? Couldn't one say that, far from rejecting the notion of an unconsciousness here, Sartre

has really just given us a very elegant way of understanding how the unconsciousness works?

Well, perhaps. But, as we'll see in the chapter on "Existential Psychoanalysis," the notion of an "unconsciousness" is linked to a certain theory — Freud's theory, of course — that Sartre thinks is just dead wrong. For Sartre, it's not just a matter of terminology.

But we can begin to get a handle on this even now. One problem with the theory of the unconsciousness is that it tries to split each overt act of consciousness into two acts — an explicit, conscious one, and a hidden, unconscious one that provides the real motive for the conscious one, and so must take any blame that may be involved.

Notice how any kind of blame is taken off the explicit act of consciousness and shifted elsewhere. For Sartre, the theory of the unconscious mind is just a way of trying to avoid responsibility.

But of course, for Sartre, responsibility has to go along with human freedom. Whether we like it or not, we are free and so absolutely responsible for what we do. In effect, therefore, by trying to avoid responsibility, the theory of the unconsciousness is in the end trying to deny human freedom. And so of course Sartre is going to reject it.

Notice how absolutely central the thesis of human freedom is for Sartre. He uncompromisingly rejects everything that would get in its way. So far we have seen Sartre reject:

- (1) the Self-Love Theory;
- (2) the notion of the unconsciousness;
- (3) the Transcendental Ego;
- (4) God

on these grounds.

The Constitution of the Ego

I now want to turn to Part II of *Transcendence of the Ego*. (Beginning on p. 60.)

This section will give us an opportunity to apply some of the general themes we have been investigating to a particular case. In my opinion, this is an exquisite passage, an excellent example of phenomenological description.

Part II of *Transcendence of the Ego* is entitled "The Constitution of the Ego." Now, remember, for Sartre there is no Transcendental Ego. And in fact, he doesn't even speak of a phenomenological Ego. The only Ego allowable in Sartre's theory is the psychological Ego — the psyche, the personality, the seat of character traits, the "real me."

So what Sartre is describing in this Part II is how I come up with such an idea. Now of course this Ego only appears to me at all in cases of reflection. So what we are going to be doing in this Part II will be a reflective enterprise.

Let's begin on p. 60.

In the section *Notes on Transcendence of the Ego* in the course packet, I remarked that at least at the beginning of the book Sartre distinguishes the notion of the Transcendental Ego (which he rejects) from the notion of the psychological Ego by means of a little terminological convention:

The Transcendental Ego = The "I"

The Psychological Ego = The "Me"

That is why the discussion of the Self-Love Theory is called "The Material Presence of the *Me*." We were talking about the psyche there, the source of selfish motives and drives.

The terminological device is pretty artificial, to be sure, but there is some rationale to it. The pronoun 'I' is in the nominative case, and so stands as the grammatical subject of verbs. Thus the "I" is the Ego as something active, which certainly fits the notion of the Transcendental Ego. Conversely, the pronoun 'me' is in the objective (accusative) case, and so functions as the grammatical direct object of verbs (among other roles). Thus the "me" is the Ego as something passive. And this fits the notion of the psychological Ego, at least in part. We'll see this in a bit, and we'll also see that Sartre's picture of the psychological Ego is really much more nuanced than this initial distinction would suggest, and that there is an active side to it as well.

In any case, that is the distinction between the "I" and the "Me" at the beginning of the book: the distinction between the Transcendental Ego and the psychological Ego. But of course there is no Transcendental Ego for Sartre, as we've seen. So by the time we get to the end of Part I of the book, in which he argues *against* the theory of the Transcendental Ego, the point of observing this nice terminological distinction has been largely lost. And so, by the end of Part I, we begin to see Sartre using the term 'I' in a rather different way: for the active side of the psychological Ego! The term 'Me' then comes to be reserved more or less for the passive features of the psychological Ego.

Thus there is an unannounced terminological shift going on in the book. The "I" and the "Me," which started off as two quite distinct things, now become just two sides or two distinct roles of one and the same thing, what he formerly called the "Me."

Thus, he says on p. 60, in the penultimate paragraph of Part I:

We begin to get a glimpse of the fact that the *I* and the *me* are only one.

And again, in the last paragraph of Part I (p. 60 again):

The *I* is the ego as the unity of actions. The *me* is the ego as the unity of states and of qualities. The distinction that one makes between these two aspects of one and the same reality seems to us simply functional, not to say grammatical.

The terminological distinction between the “I” and the “Me” is perhaps not very important in the long run. Nevertheless, there are several important notions and terms introduced in the passage I just read. The question of the whole of Part II is: How does the psychological Ego, the personality or “self,” appear to us when we reflect? (‘Appear’ doesn’t of course mean visually appearing here. It just means being any kind of object of consciousness.)

In the passage just quoted, Sartre says this Ego appears to us as a unity of several things. Of what things? Well, of actions (on the active side), and of states and qualities (on the passive side). So we will have to ask what those are exactly.

But there is something else here we have to note. If we look back again to the penultimate paragraph of Part I (p. 60), Sartre says:

We are going to try to show that this ego, of which *I* and *me* are but two aspects, constitutes the ideal and indirect ... unity of the infinite series of our reflected consciousnesses. [The emphasis in the last phrase is mine.]

(Note: He says ‘reflected’ consciousnesses, not ‘reflective’ consciousnesses. A “reflected” consciousness is the object of a “reflective” consciousness.)

Now among other things in this passage, note the word ‘indirect’ here. As he goes on to say in the first sentence of Part II (still on p. 60):

The ego is not directly the unity of reflected consciousnesses. [My emphasis again.]

This suggests, then, that something else is going to be the direct unity of reflected consciousnesses.

Now, what is all this about?

Well, recall the passage from *The Psychology of Imagination*, where we talked about the difference between perception, imagination, and conception. There we said that when we perceive something as a cube, we perceive at most three sides facing us at once. We see it only in profile, in an Abschattung. We see those three sides as belonging to a configuration that has six sides in all, which are arranged in such and such a way, result in a solid physical object, and so on.

We see all those extras as implied, as promised, by the three sides that are presented directly to us.

So too in the case of consciousness. Pierre walks into the room, and — to use Sartre's example (p. 62):

I feel a sort of profound convulsion of repugnance and anger at the sight of him.

It is this experience we are going to reflect on in this Part II of the book, and we are going to try to be as phenomenologically sensitive as we can about it.

So, how does this feeling of repugnance appear to me when I reflect on it?

Well, that momentary repugnance is present to me, it is directly before my mind. But, just as with the three sides of the cube, the momentary repugnance comes on to me as only part of the story. When I reflect on that repugnance, I see it as simply a profile of a larger whole, as a kind of instantaneous window that provides a glimpse of a more stable and enduring fact: I HATE PIERRE!

(Note: Sartre is not saying that my momentary repugnance must appear to me in this way on reflection. This is just an example. The same thing holds, of course, for the three sides of the cube. I can just see them as three sides.)

Here is what he says (p. 62 again):

But is this experience of repugnance hatred? Obviously not. Moreover, it is not given as such [any more than the three sides come on to me as being — all by themselves — THE CUBE.] ... If I limited it to what it is, to something instantaneous, I could not even speak of hatred anymore. I would say: "I feel a repugnance for Peter [it's "Peter" in *Transcendence of the Ego*, but "Pierre" in *Being and Nothingness*. The French is "Pierre" in both cases.] at this moment," and thus I would not implicate the future. But precisely by this refusal to implicate the future, I would cease to hate. [Emphasis added.]

Hatred, in other words, is serious business. It involves much more than a momentary twinge of repugnance. In the case of hatred, as in the case of the cube, there is more than meets the eye. Hatred involves a commitment. There is a kind of stability and permanence to hatred that isn't there in my fleeting momentary feeling of repugnance all by itself.

Now you may well ask here, "Just what is the object of my consciousness here?" I am engaging in a reflective enterprise, but what is the object I am reflecting on? I thought it was my momentary repugnance for Pierre, but now it turns out to be more than that, to be some kind of long-term project of hating Pierre. Which is it?

Well, it's both. And there should be no confusion caused by that fact if we are careful. When I look at the cube, what is the object of my consciousness? Is it the cube or is it only the three sides directly facing me? Well, obviously, the answer is both. I see the three sides, and in the same act, I see the cube through them.

When you look at this chair, what do you see? Do you see the chair, or only the front of the chair? The obvious answer is, “I see the chair BY seeing only one side of it. That’s just what we mean by ‘seeing it from one side’.

So too (p. 62, bottom):

Now my hatred appears to me at the same time as⁵ my experience of repugnance. But it appears through this experience. It is given as precisely not being limited to this experience.

And later (on p. 63):

Hatred, then, is a transcendent object. [In fact, it is transcendent in both of the senses of that word distinguished by Husserl in *The Idea of Phenomenology*.] Each *Erlebnis* [= experience] reveals it as a whole, but at the same time the *Erlebnis* is a profile, a projection [= *Abschattung*]. Hatred is credit for an infinity of angry or repulsed consciousnesses in the past and in the future. It is the transcendent unity of this infinity of consciousnesses. [Emphasis added.]

Recall that the cube in our example is just the summation of the phenomenon of the three sides facing me, plus everything that is not directly facing me but is implied or promised insofar as I see the thing as a cube. The cube is just the sum total of all that. And the essence of the cube is the “principle of the series”; the essence, which I can extract by eidetic abstraction, is what determines what all is included in that sum total.⁶

So too, the hatred in the present passage is just the sum total of the momentary repugnance I am now reflecting on, plus all those past and future repugnances that are, so to speak, “around in back.”

Furthermore, the promises made by my momentary repugnance for Pierre, the determination that I will continue to hate Pierre until my dying day — there is no guarantee those promises will come true. I may change my mind. I may decide Pierre is really an OK guy.

So, just as in the case of the cube, we have a promise of “more to come,” and that promise is not guaranteed.

⁵ The translation is ambiguous here, and might cause confusion. Don’t read it as: “...my hatred appears to me at the same time as being my experience of repugnance.” Read it as: “...my hatred appears to me at the same time as does my experience of repugnance.”

⁶ Recall that Sartre said Husserl had reduced all the traditional dualisms to one: the duality of finite vs. infinite. For Husserl, at least as Sartre interprets him, the object was just the sum total of an infinity of phenomena. For Sartre, however, there was more involved; there was being-in-itself, which provided the being of the phenomenon. So it’s not quite true to say that for Sartre the cube just is the sum total of all the phenomena; that’s Husserl, not Sartre. But the analogy between the cube and anger still holds, and the technical refinement should not distract us for now.

All this means that hatred is a perceptual object, just as in the case of the cube. It's not a sensory object, of course, in the way the cube is. Hatred is not something we see or hear. But, in accordance with the schema we set up earlier, hatred has to fall under the heading of perception. (We certainly aren't just imagining or merely conceiving the hatred.)

Obviously what we are doing here is broadening our notion of perception, so that it includes more than just sensation.

Another way to look at it, so that this broadening perhaps doesn't seem so arbitrary, is to recall that in that passage from *The Psychology of Imagination*, we said the notion of making an objective claim went with perception, whereas imagination and mere conception were subjective. Perceptual claims were objective in the sense that they were something I can be wrong about.

So too here. I can be wrong about my hatred for Pierre. Hatred is not just a matter of making a resolution to hate Pierre; it is a matter of keeping that resolution. And I can never be sure I will do that. Thus, in that sense, to say I hate Pierre is to make an objective claim. And that means that hatred belongs under perception.

Now of course hatred is not yet the Ego, not the psyche or self. It is what Sartre in this Part II of *Transcendence of the Ego* calls a state — a mental state.

A while ago we saw Sartre say that the Ego was going to be an indirect unity of our reflected consciousnesses, implying then that something else was going to have to be the direct unity. Well, now we have it. The state is the direct unity here. It's direct in the sense of being immediate and not indirect. It's the first unity we come to beyond that momentary feeling of repugnance, the first transcendent unity that is foreshadowed by that momentary repugnance — the first one we come to “around in back,” “over the horizon.”

Now, before we go any further, let us note that states in this sense are not the only kinds of such direct psychological unities for Sartre. There are also what he calls actions. Sartre discusses actions on pp. 68–69, but he doesn't say much about them.

There is really no need for him to say very much about them. The idea is basically the same as for a state; the differences between them are relatively minor.

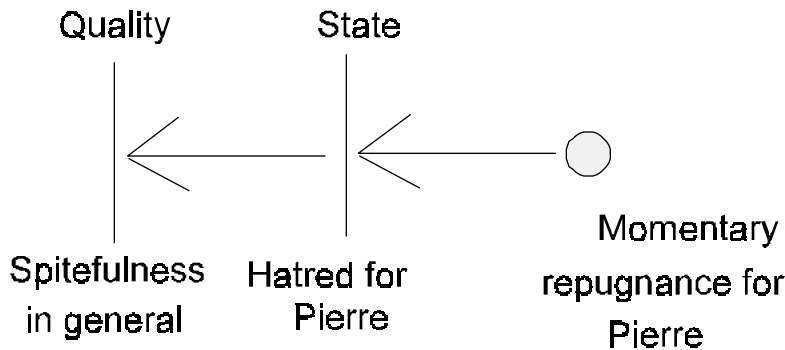
A momentary act is not an action, in this sense; an action is a longer-term project. Driving to Chicago is an action in a way that turning the ignition key is not. Playing the piano is an action in a way that playing middle C is not. Playing the piano is the sum total of a number of momentary finger and hand movements, which cannot be guaranteed in advance, just as hatred is the sum total of a number of momentary repugnances that may or may not actually occur.

We don't need to worry very much about actions for our present purposes. Just recall that, earlier (p. 60), Sartre had said that the I is the active part of the psyche, and is the unity of actions, whereas the me is the passive aspect, and is the unity of states and of qualities.

Let's set actions aside then, and concentrate on the others. We already have the notion of a state, but what about that third notion: qualities?

Well, just as we reflect on our momentary repugnance, and see implicitly behind it the state of hatred, so too the quality is something I see lurking behind the state. It's a kind of second-order unity, implied by the state that is in turn implied by the momentary repugnance for Pierre. (Thus, qualities are going to be indirect unities, whereas the states we said were direct unities.)

Thus we have:



(The things at the *head*-end of the arrows are supposed to be in some sense more *basic*, more *primordial*, than the things at the *tail*-ends.)

Just as the hatred is a summation of an in principle infinite series of momentary repugnances, both actual and implied, so too the quality is a summation of a possibly infinite series of states, a whole series of hatreds. I not only hate Pierre; I hate Jean-Louis too! In fact, I hate lots of people! I have a very spiteful or hateful personality. Spitefulness or hatefulness thus is a unity of hatreds, just as hatred is a unity of repugnances. Spitefulness or hatefulness is an example of what Sartre calls a quality.

Let's make two observations here:

- (1) Sartre says qualities are optional. What does this mean? Well, in the first instance all it means is that I can view my momentary repugnance for Pierre as a manifestation of my hatred for him without thereby also viewing that hatred as evidence for a deeper tendency for being spiteful in general. Perhaps I don't hate everyone; I'm generally pretty easy to get along with. But I do hate Pierre. In that sense, the quality is optional.

Now of course, in that same sense the state is optional too. I can reflect on my repugnance for Pierre without necessarily viewing it as part of a hatred; perhaps I view it

as just an odd momentary quirk, the result of a passing bad mood, but of no real long-term significance.

So I don't think Sartre means to be drawing a basic and fundamental distinction here between states and qualities by saying that the latter are optional. I think all he means to be saying is that even given that I view my repugnance for Pierre as a manifestation of my hatred for him, it is still optional whether I view that hatred in turn as a manifestation of a deeper character trait or quality.

- (2) Although Sartre mentions qualities only as optional transcendent unities of states, I don't see any reason why we couldn't also have qualities (or whatever we wanted to call them) as optional transcendent unities of actions. That is, just as my hatred for Pierre might possibly come on to me as an instance of my deep-seated quality of being spiteful or hateful in general, so too why shouldn't my action of, say, campaigning on behalf of the Flat Earth Society be a manifestation of a more general and deep-seated quality I have of engaging in eccentric and quixotic causes? I don't see that Sartre would object to this, but he doesn't mention the possibility — just as he doesn't say much of anything at all about actions.

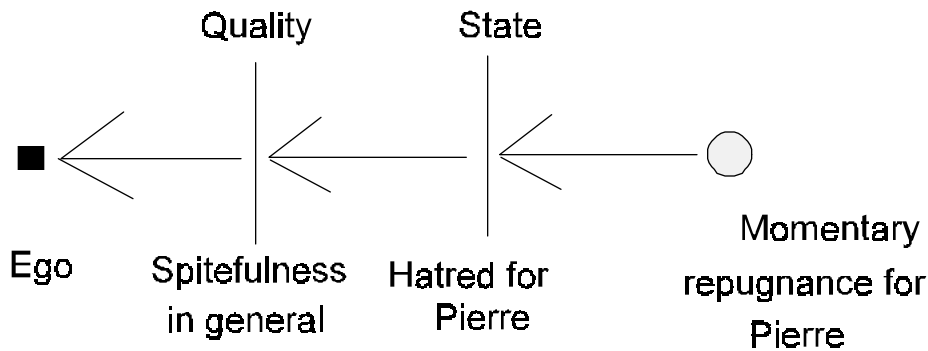
Now, it is only after all of this that we are in a position to talk about the Ego or Self, the personality.

Just as the state is a transcendent unity looming up behind the momentary repugnance for Pierre, and just as the quality is an additional, second-order transcendent unity that I may or may not see looming up behind the state, so too the Ego — the Self — is a further transcendent unity I see implied and manifested by the states and qualities.

Who am I? I am a person who hates Pierre, who is driving to Chicago, who also has a tendency to take on quixotic causes, who loves Griselde, and so on. You put it all together, and that's me. That's who I "really" am. That is the psychological Ego. As Sartre said (p. 60 again), the Ego is

... the ideal and indirect ... unity of the infinite series of our reflected consciousnesses. [Emphasis added.]

Thus, we have:



The Magical

But there is more to the story. How are these various levels and stages related to one another? That is, how do they appear to me on reflection as related to one another?

Let's ignore actions for the present, since they aren't going to add anything really new here. And let us include qualities. We have then four stages or levels:

- (1) The momentary repugnance for Pierre.
- (2) The state (hatred).
- (3) The quality (spitefulness).
- (4) The Ego.

And of course, in all of this, the only thing that in a sense is really given to me in reflection, the only thing that is directly given to me in the sense that I can be said to know it's there, is the momentary repugnance for Pierre. All the rest is implication, is promise; all the rest is what I see through my experience of that repugnance.

Let us start then with the momentary repugnance for Pierre. How does that repugnance come on to me in reflection? Well, it is active, it is spontaneous, it is a kind of process or event. In a sense, it is unpredictable. Even if I do view that repugnance as the result of a long-term hatred, still that hatred reveals itself in these moments of repugnance only sporadically, not all the time, and I never really know just when it is going to emerge.

On the other hand, the state of hatred comes on to me as being different. It is not spontaneous and unpredictable in that sense; it is not a process or event but a state, after all. It's much more permanent; it endures over time. It's still there even when I am not actively feeling any repugnance for Pierre. I hate Pierre while I'm cooking dinner. I hate Pierre while I'm driving to work. I hate Pierre while I'm asleep! I hate Pierre all the time, not just when my adrenaline gets going. (That's the whole point of distinguishing the

momentary act from the long-term hatred.) It's the sort of thing Aristotle called a "habit."

In short, the state of hatred appears to me as inert and stable, fixed, more or less like being-in-itself.

Nevertheless, the passive, inert state and the spontaneous, active repugnance for Pierre come on to me as related in a certain way. That is, they are not just two disconnected things. Although the feeling of repugnance comes on to me in reflection as spontaneous, nevertheless it is still somehow a manifestation of that hatred. It's because I hate Pierre that I feel this momentary repugnance. That spontaneous repugnance seems to arise out of the inert state.

And now we come to one of the central notions in Sartre, the notion of THE MAGICAL.

The connection between the state and the momentary repugnance is what Sartre calls a magical connection (p. 68):

We readily acknowledge that the relation of the hatred to the particular *Erlebnis* [= experience] of repugnance is not logical. It is a magical bond, assuredly.

Now the notion of the magical is a technical term for Sartre. Unfortunately, he doesn't define it for us in *Transcendence of the Ego*, at least not explicitly. But he does define it for us in *The Emotions* (pp. 88–91). There he is quoting someone else [= Alain], but he quotes the definition with approval. The magical, he says, is "the mind dragging among things."

Then he goes on to explain this in his own terms:

... that is, an irrational synthesis of spontaneity and passivity. It is an inert activity, a consciousness rendered passive.

But of course, passivity, inertness, are characteristics of being-in-itself, not of being-for-itself. Contrariwise, spontaneity and activity are features of being-for-itself, of consciousness, not of being-in-itself.

Hence the notion of "a consciousness rendered passive" is the notion of a being-for-itself that is also a being-in-itself. "An irrational synthesis of spontaneity and passivity" is, again, being-for-itself that is also a being-in-itself. So too with the notions of "an inert activity" and "the mind dragging among things."

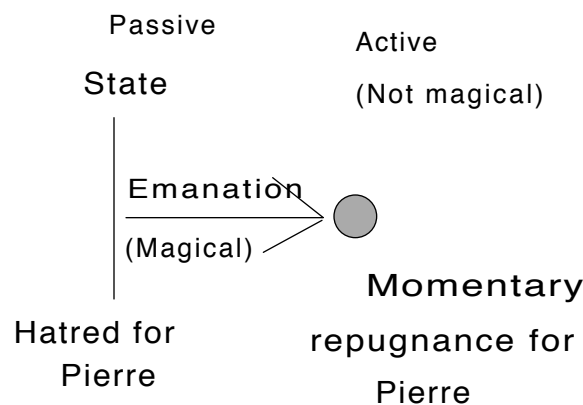
In short, the notion of the magical is the notion of that impossible combination of being-in-itself and being-for-itself that we have seen before. For Sartre, the Transcendental Ego is magical; the Freudian unconsciousness is magical; God is magical. Magical things don't really exist, of course; they are impossible. But we certainly encounter them in our theories about the world, in our ways of looking at things.

Why does Sartre call this “magical”? Well, it’s because it’s *spooky*. When inert *things* start behaving as though they were *conscious*, when they start acting *unpredictably*, that’s “magical.” When haunted houses begin to creak and groan of their own accord, when trees begin to speak, and heavy objects begin fly through the air as though they had a will of their own — that’s *magical*.

Now let’s apply this notion of the *magical* to our reflective enterprise. If we think about it, the connection between the *passive* state of hatred and the *active*, spontaneous, momentary repugnance for Pierre is a *magical* connection. It *combines* (links) the in-itself and the for-itself in the way we have just described. There is nothing magical about the *state* of hatred all by itself; it is inert and stable, a long-term matter. And there is nothing magical about the momentary repugnance for Pierre, all by itself; it is spontaneous and free. But the *linkage between the two*, the fact that the repugnance comes on to me *as though* it were somehow *produced* by the passive, inert state — *that* is a magical linkage.

Sartre has a special name for this magical linkage of the state with the spontaneous momentary act. He calls it *emanation* (p. 67, bottom). The state doesn’t exactly *cause* the spontaneous act, because then the act wouldn’t be spontaneous. Besides, the *state* is inert and doesn’t *do* anything. But nevertheless, the momentary act is given as *somehow* produced by that state. And that mysterious and ultimately *unintelligible*, “magical” connection is what he calls *emanation*.⁷ It is a kind of part-whole relation, like the relation of the three sides to the cube as a whole. Only in *this* case, that part-whole relation is “magical.” (It wasn’t magical in the example of the cube.)

⁷ Should we therefore *also* think of the “surging up” of the For-Itself from the In-Itself as a relation of *emanation* — a *magical* relation? The temptation is strong, but we should resist it. For, as we shall see, Sartre ultimately thinks that *magical* things are impossible; there really aren’t such things — even though there may well be real things that really *appear* to be magical. And surely we don’t want to have to say that the relation of the For-Itself to the In-Itself doesn’t really exist. That would be to deny that the For-Itself really *depends* on the In-Itself — that the For-Itself is *not* in-itself.



(Do not be confused. I am changing the function of the arrows now. Earlier, they went from the more *immediate* or *direct* objects of consciousness to the more *mediate* or *indirect* objects of consciousness. Now I am changing their function: the arrows go from *source* to *product* or *result* — in appropriately broad senses of those terms.)

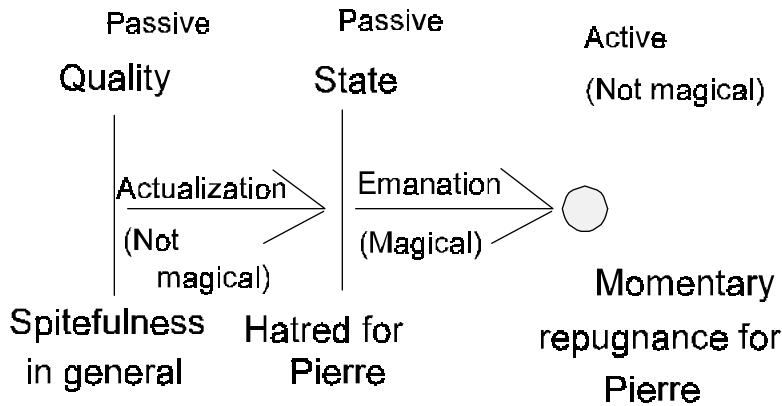
Let's look further. What about the linkage between the quality of spitefulness or hatefulness in general and the state of hatred for Pierre in particular. What kind of relation is that?

Well, what kind of thing is the quality? That is, how does it appear to me on reflection?

The quality of spitefulness is a kind of general notion. That is, whereas my hatred for Pierre is directed toward Pierre alone, my spitefulness is a tendency to hate just anyone. The relation between my spitefulness and my hatred for Pierre is a relation then of general to particular, of a tendency or potential for hatred in general to the actuality of my hatred for Pierre in particular.

There is nothing spontaneous or active about the quality of spitefulness; it is an inert, permanent potentiality, a kind of deep "character trait." Neither is there anything active about the state of hatred; it too is an inert, passive, stable thing, just as it was a moment ago. Hence the relation between them is not a relation that combines the passive with the active; it is not a magical relation.

On the contrary, my hatred for Pierre comes on to me as simply a kind of actualization in this particular case of my general potential for hating people. And so Sartre calls this relation: actualization. Note that it is not magical.



Finally, what kind of linkage is there between the Ego and the qualities, or — since qualities are optional — between the Ego and its states? How do these relations appear to me on reflection? How is my Self or personality connected with my spitefulness or hatefulness in general, or with my hatred for Pierre in particular?

Well, let's look at the two ends of the relation. Both the quality and the state are passive, inert, as we have seen. But what about the other pole, the Ego or Self?

As Sartre says (p. 77):

Everyone, by consulting the results of his intuition, can observe that the ego is given as producing its states. [Emphasis added.]

(The same thing goes, I suppose, for the optional qualities: the Ego is given as producing them too.)

That is, my hatred for Pierre is something I undertake, my spitefulness is something that is a result of who I am. So the Ego, the Self, the personality, is given as something active, that results in these features of me.

But it is also something passive. My hatred for Pierre is not something my personality generates and then lets fly away, as though it didn't matter. It's something that turns back on me, that compromises me, that affects me. When Raskalnikov, in Dostoevski's *Crime and Punishment*, kills the old pawn broker, it's not something he does and then that's over and done with. No — it leaves a mark on his soul, it changes him; it leaves what Sartre calls a "killing bruise."

Thus, although the Ego is active, it is in exactly the same respects passive as well. (The idea here is like a kind of principle of action and reaction.) Thus, the Ego itself is "an irrational synthesis of spontaneity and passivity." In short, the Ego itself, the Self, the personality — is already a magical object. As Sartre says (p. 82):

This is also why we are sorcerers for ourselves each time we view our *me*.

In other words, the *spookiness* we described in the notion of the “magical” is there too when we think about the Self, the personality. And in fact, this is true not just when I am reflecting on *my own* Self, but whenever I think about *other* people too. As Sartre says in *The Emotions* (p. 84):

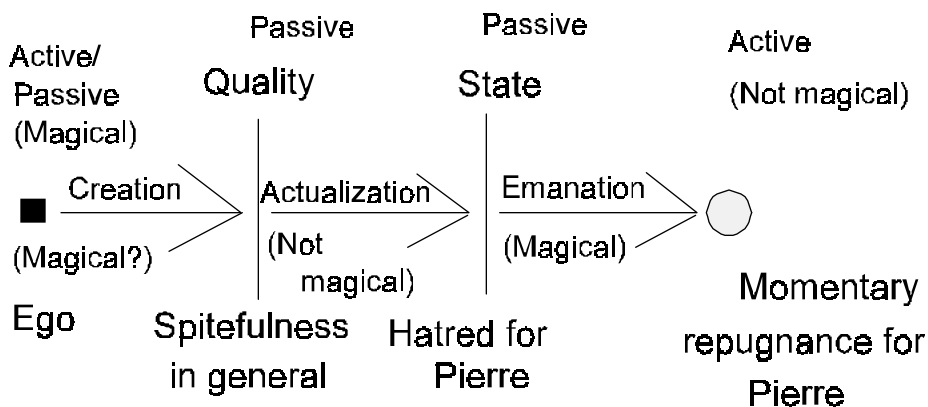
Thus, man is always a wizard to man, and the social world is at first magical.

So — unlike the relation between the *state* and the *act* of repugnance, a relation that is magical insofar as it links the passive state with the active repugnance, and unlike the relation between the *quality* and the state, a relation which is a *non*-magical linkage of *actualization*, joining two passivities — the linkage between the Ego or Self and its qualities and states is a linkage between something that is *already* magical (the Ego) and something that is passive. Is such a linkage *itself* magical or not?

Well, Sartre says, sometimes it is and sometimes it isn't. He says (pp. 78–79):

Most of the time the progression involved is magical. At other times, it may be rational

I'm not entirely sure I see which is which here. (Perhaps it depends on whether we are emphasizing the active or the passive side of the Ego.) But in any event, it is clear that we are dealing here with a relation unlike any of the others we have dealt with hitherto. Sartre calls it “*creation*.”



The term is appropriate, since the Ego is a *magical* notion, like the notion of God. And the way in which the Ego is given as producing its qualities and states is very much like the way in which God is thought of as producing the world. In both cases, a passive *product* is given as somehow emerging from a *magical* source. (But of course in traditional theology, God is not “compromised” by his products, as Sartre says the Ego is “compromised” by its qualities and states.)

So this is the story of the constitution of the Ego, of how the personality, the Self, is built up in consciousness, in reflection.

When I reflect on my momentary repugnance for Pierre, and see it as emerging — emanating — from my hatred for Pierre, which in turn is a special case of my spitefulness in general, which in turn is a product of my Self or personality, when I do all that, the picture I get is a picture according to which the direction of the development goes from the Ego to the momentary act of repugnance. The Ego is the ultimate source, and the momentary act of repugnance is its final result.

But of course that's all false! In fact, it's just the opposite of the truth. In truth, everything is just the other way around. What is directly given to me in reflection — and the only thing that is directly given to me in reflection — is the momentary act of repugnance. The rest is all inference; the rest is all “seeing as.” The rest is all the way I interpret it. In short, the rest is all a matter of how I constitute it.

This entire story about the Ego and its qualities, states and actions, is an elaborate lie. The truth is just the opposite. The Ego is not the producer and generator of my acts of consciousness (through the qualities and states); on the contrary, the Ego is produced or constituted by my reflecting on my momentary acts of consciousness.

In effect, the true story of consciousness is the one told in Part I of *Transcendence of the Ego*. The story in Part II is the distorted story consciousness tells itself whenever it reflects.

And now we come to something tremendously puzzling and tremendously important in understanding Sartre.

For Sartre, whenever we reflect on our acts of consciousness, we always and inevitably distort them in this way. As he says near the beginning of Part II of *Being and Nothingness* (p. 121):

It is often said that the act of reflection alters the fact of consciousness on which it is directed. Husserl himself admits that the fact “of being seen” involves a total modification for each *Erlebnis* [= experience].

(See also *Transcendence of the Ego*, p. 45.)

However this works for Husserl, it is easy to see why this should be so for Sartre. For we said long ago that every act of consciousness is consciousness of being-in-itself. (Every movie is a movie on the screen.)

Now, to see the point, let's take an analogy. Suppose that, in order to think about something, we had to make up a little wax model of it for ourselves. Then we would put the little wax model in front of our eyes, and then we could think about the object it's a model of. Just suppose, for the sake of the illustration.

Well then, as long as what we wanted to think about really was made up of wax, there would be no problem. If we wanted to think about candles, for instance, we just make up

some little wax candles and put them in front of ourselves, and then we could think about them. There would not be any distortion in this case.

But if what we wanted to think about were not really made up of wax, we would have a different situation. For example, if we wanted to think about a dog, we would have to make up a little wax model of a dog, and put the model in front of our eyes and then think about the dog. But of course, the wax model of the dog is not exactly like the dog itself. The dog is made of fur and flesh, while the model is only made of wax. When you hold a flame to the wax model, it melts; when you hold a flame to the actual dog, it scorches.

In other words, the wax model of the dog is fine as far as it goes. We can think about the dog that way, after all, and for certain purposes we can think about dogs quite adequately this way. But the model distorts the reality; it introduces features that are not there in the real dog, and leaves others out, or changes them.

Now SO TOO: Consciousness, we know, can only think about an object — any object — by looking at being-in-itself. It is as if, in order to think of something, we first had to make up a little model of it out of being-in-itself. Then we put the little model before our mind's eyes, and we can think about the object. This in effect is what the theories of *intentionality* and *constitution* come to for Sartre, as we have seen.

Now as long as the object we are thinking about really is made up of being-in-itself, everything is fine, and there is no distortion — just as with the candles when we were talking about the wax.

But if what we want to think about is not really made up of being-in-itself — and that means if what we want to think about is consciousness, since that is the only thing that isn't made up of being-in-itself — then we get an inevitable distortion, just as with the dog when we were talking about the wax. What we end up with is a little model of consciousness, a model made up of being-in-itself. The model is fine for certain purposes, and we can succeed in thinking about consciousness in this way. But it distorts. The model makes consciousness look like something that has some of the features of consciousness, to be sure, but that also has some of the features of being-in-itself — just as the wax model of the dog had some of the features of the dog and some of the features of wax.

In short, the model of consciousness, made up of being-in-itself, makes consciousness look like a combination of being-in-itself and being-for-itself; it makes consciousness look like God.

This always happens when we reflect, and we now see why. That's why the Freudian theory of the unconscious looks initially plausible. That's why the theory of a Transcendental Ego is attractive to some philosophers. And, in the end, that's why “man is always a wizard to man.”

Now, of course, there is an obvious problem — and this is what I said is absolutely crucial in our understanding of Sartre.

If reflection always distorts, then what about the theory in Part I of *Transcendence of the Ego*, which is very much a reflective theory, a theory of consciousness? Is that theory then itself a distortion? If so, is there any reason why we should take it seriously?

For that matter, if reflection always distorts, then it would appear that least one-half of *Being and Nothingness* is wrong!

The problem, therefore, is a very serious one for Sartre. How is he going to get started? He has got to find some way of getting around the fact that one of the consequences of his theory would seem to be that the theory itself has got to be a distortion and so wrong..

What about this, Sartre?

Well, Sartre is aware of the problem of course (see *Transcendence*, p. 46; also p. 41), and has at least a hint of a solution to it. The solution is going to be something he calls pure reflection.

Sartre refers to this several times in *Transcendence of the Ego*. For example, on p. 91, he says:

One can even suppose a consciousness performing a pure reflective act which delivers consciousness to itself as a non-personal spontaneity.

‘Non-personal’ here is the key term. That is, not endowed with a personality, an Ego. In short, not distorted in the way we have just described.

Again, on p. 101:

But it can happen that consciousness suddenly produces itself on the pure reflective level. [Emphasis added.]

Now what on earth is this pure reflection? It’s all very good to give a name to whatever it is that is supposed to solve our problem, but we need some further account of what this process is. And unfortunately, Sartre doesn’t say very much about it here.

It seems to me there are, at least initially, two things this might be:

- (a) A reflection in which everything proceeds just as we have seen it described in Part II of *Transcendence of the Ego*. The states, qualities, actions and the Ego itself all appear, just as Sartre described. But I am not fooled by any of this. I refuse to accept the promises of “more to come.” Perhaps all this would mean is a reflection in which I perform the epoché, the phenomenological reduction.

There is some textual evidence for this interpretation of what pure reflection would be. For example, on p. 101, we read (immediately following the passage just quoted):

Perhaps not without the ego, yet as escaping the ego on all sides
[Emphasis added.]

That is, the Ego does appear there, but we are not fooled by it, we “escape” it. Whether this is what Sartre means by these cryptic words here, I do not know.

Another thing “pure reflection” might mean is:

- (b) A reflection in which none of the things Sartre has discussed in Part II of *Transcendence of the Ego* even appears to me. This would be roughly like seeing the three sides of the cube simply as three surfaces, and not as three sides of anything.

(The difference between these two possibilities is that, on possibility (a), I see my momentary repugnance for Pierre as making further promises of “more to come,” but I do not commit myself to whether those promises are true or false. On possibility (b), by contrast, I do not see my momentary repugnance for Pierre as even making the promises in the first place.)

One difficulty with interpreting the notion of “pure reflection” this second way comes from a remark Sartre makes on p. 92:

A reflective apprehension of spontaneous consciousness as non-personal spontaneity [that is, as not endowed with any of the business discussed in Part II] would have to be accomplished without any antecedent motivation. This is always possible in principle, but remains very improbable, or, at least, extremely rare in our human condition.

Whatever this means, it would seem to make it unlikely that this is what Sartre means by the “pure reflection” that is going to make his whole enterprise possible.

It is important to realize that the notion of pure reflection is not just some technicality. It is presupposed by Sartre’s whole procedure. Without it, the whole project of *Being and Nothingness* is doomed from the outset.

I should also mention that Sartre returns to this topic in *Being and Nothingness* itself, although what he says there is still pretty obscure. The discussion occurs in Part Two, Chapter 2 (Temporality), section III: “Original Temporality and Psychic Temporality: Reflection” (beginning on p. 211).

There is a good Ph.D. dissertation on this topic, by one of our own recent Ph.D.’s, Christopher Vaughan, entitled *Pure Reflection: Self-Knowledge and Moral Understanding in the Philosophy of Jean-Paul Sartre* (Indiana University, 1993) — a dissertation I am proud to say I directed. I don’t think Vaughan has settled these questions once and for all (and neither does he), but he has certainly made a lot of progress. Highly recommended.

The Problem of Other Minds

We are nearly finished with our discussion of *Transcendence of the Ego*, and nearly ready to begin our discussion of *Being and Nothingness* in detail. But there is one last item I want to discuss from *Transcendence of the Ego*. And that is what Sartre has to say there about the “problem of other minds”: How can I ever be sure there are other consciousnesses like my own?

The “problem of other minds” is a kind of special case of the problem of *solipsism*, which we discussed in connection with Descartes a long time ago. Sartre discusses it briefly on pp. 103–104, where he *in a sense* gives us a “solution” to the problem of other minds.

The discussion is only two paragraphs long. It begins on p. 103 with the words:

This conception of the ego seems to us the only possible refutation of solipsism.

In effect, here is the essence of what he says:

What exactly *is* the problem of other minds? Isn't it just that I can be *quite sure* about *my* self in a way that I cannot be at all so sure about *yours*? I have a kind of “privileged access” to my own mind that gives me a kind of certainty when I am speaking about myself that I do not have when I speculate about what is going on in other people's minds — or even about whether they *have* minds.

Thus, the *problem* of other minds appears to arise because of a fundamental *disparity* between my Self and yours. I have a special access to my own Self in a way I do not have access to your Self or to anyone else's. And that “special access” gives me a kind of infallibility about my own case that I simply can't have when I am talking about you.

But on Sartre's view, this is just not so. I *don't* have any kind of special “privileged access” to my own Self that I lack to yours. Both *my* Self — my personality, the “real me” — and yours are objects *for* consciousness. They are both things we can study objectively. I am in no more privileged a position to pontificate about *my* Ego than I am about yours. I can be *mistaken* about *my* Self just as much as — perhaps even *more* than — I can about yours.

This is why other people frequently know us in certain respects better than we do ourselves! If I say “I hate Pierre!,” you may say “No — your relationship with Pierre is really much more complex than that.” And you may be right, whereas I am wrong! How would that be possible if I were somehow *infallible* about my own Self?

In short, for Sartre, *all* Ego's are *tentative* objects — just like the cube in perception. I can make mistakes about mine, just as I can make mistakes about any other. As long as we are talking about the *Self*, my Self is no more certain than yours.

And so *the problem of other minds is solved!* Now you may think this doesn't amount to a solution, but in an odd sense it does. The original *problem*, remember, arose because of

a disparity between my Ego and yours. That disparity has been disposed of in Sartre's theory.

In short, Sartre here tries to "solve" the problem of other minds by proceeding just the opposite of the way most people do. Most attempts to solve the problem of other minds try to find some way of allowing us in the end to be as sure about other minds as we are about our own. In other words, they try to find some way to raise the level of confidence we have in other minds until it is as high as the level of confidence we have in our own. And in that way, they remove the disparity that was the basis for the problem.

Sartre proceeds just the opposite. Instead of trying to raise our knowledge of other minds to the level of our knowledge of our own minds, he lowers the knowledge we can claim about our own minds until it is no better than the knowledge we can claim about other minds.

In both cases, the disparity is removed, and so, in a sense, the problem is solved.

You may not think this "solution" amounts to much. And Sartre, in the end, came to think that too. In *Being and Nothingness*, he finally decided that there was still a problem of other minds, a problem that arises in a form he had not yet addressed.

Listen to what he says in *Being and Nothingness*, p. 318 (in the discussion of "The Existence of Others"):

Formerly I believed that I could escape solipsism by refuting Husserl's concept of the existence of the Transcendental "Ego." At that time I thought that since I had emptied my consciousness of its subject, nothing remained there which was privileged as compared to the Other. But actually although I am still persuaded that the hypothesis of a transcendental subject is useless and disastrous, abandoning it does not help one bit to solve the question of the existence of Others.

And Sartre then goes on to give the problem another and quite a different treatment.

The Origin of Negation

After these lengthy preliminaries, we are now ready to get into the main text of *Being and Nothingness*.

After discussing the notions of being-in-itself and being-for-itself in a preliminary way in the "Introduction," Sartre now wants to talk about the relation between them. And this is what he starts with in Part I, Ch. 1: "The Origin of Negation."

Now, in a sense we already know what the origin of negation has to be — it's the for-itself. But I've told you that; Sartre hasn't told us that yet. He is now going to tell us, in this chapter.

We don't have to spend a lot of time on this chapter, but I don't want to just skip over it entirely. There are several very interesting things in it.

Sartre begins by giving some introductory remarks (pp. 33–34) to the effect that we shouldn't try to treat the two sides of the relation between being-in-itself and being-for-itself *in isolation*. This is a theme Sartre got from Heidegger.

The idea is that, if we start off by considering being-in-itself, matter, the world, *all by itself*, and *then* — only afterwards — worry about how we are going to bring *consciousness* into the picture, we are going to end up with a one-sided and hopelessly inadequate account. So too, if we start off, as Descartes did, with *consciousness* all by itself, and only *then* go on to consider how consciousness is related to the world, we will once again end up with a one-sided and inadequate account.

On the contrary, Sartre says, if we are ever going to get an adequate picture of the *relation* between being-in-itself and being-for-itself, we must start *from the outset* by considering both poles of the relation together

In a sense, this is related to (although it is not the same as) the point I made earlier, about how for Sartre you don't build up the individual as a product of the intersection of general principles. So too here, you don't build up a complex, concrete relation between two things by simply starting with the two things in isolation, and then trying to stick them together.

Now, one way in which consciousness is related to the world (the for-itself to the in-itself) is by *questioning* it — wondering about it, inquiring into it, just as we are now doing.

Sartre takes this concrete case as his starting point, his “case in point” — as he puts it, his “guiding thread” (p. 34).

So let's examine the peculiar *relation* by which consciousness stands in an *interrogative attitude* toward the world. WHAT IS REQUIRED FOR THIS RELATION TO BE POSSIBLE? (Note that what we have here is a kind of Kantian “transcendental argument.”)

Well, the first thing Sartre notices about the “interrogative attitude” is that *every question requires three kinds of non-being, three kinds of NOTHINGNESS* (pp. 34–36).

- (1) First, there is the *non-being* of knowledge in the questioner. In other words, in order genuinely to take an *interrogative attitude*, I can't already know the answer. If I do, my question is just a formality and not a real question. So the very fact of *asking* a question in the first place implies something *negative* — a *lack*, in this case — on the part of the one who asks the question.
- (2) Second, Sartre says, in every question there is always what he calls “the possibility of non-being of being in transcendent being” (p. 36). Basically, all this means is that, for every question, there is always at least the *possibility* that a *negative* reply is the correct

one. In other words, there is always the possibility that there is something about objective reality (“transcendent being”) that makes a negative reply appropriate — something lacking in objective being.

For example, “What’s wrong with the computer?” Perhaps the answer is: “It’s not plugged in.” Perhaps one of the memory chips has failed (that is, it’s not working any more). Perhaps there’s nothing wrong with the computer — it’s supposed to do that! All of these are negative replies that presuppose something negative about the computer itself.

Sartre goes to some length to try to make the case that every question can be construed in a way that leaves open the possibility of negative answers. I’m not sure he completely succeeds, and I’m not sure it really matters. If there are questions that cannot be cast in this form, let’s just set them aside and focus our attention on those that can. We are only taking this as an illustration, after all.

- (3) Third, Sartre says that each question presupposes a definite answer. That is, it presupposes that the correct answer is such and such, and (therefore) not something else. What time is it? It’s 12 noon (say), and not 12 midnight or 3:00 p.m.

Questions, therefore, implicitly presuppose that objective reality (the world) is differentiated, demarkated. The world comes divided up into parts such that one part is not another one. This differentiation, this distinction of one part from another, is yet another form of non-being. (Recall how for Parmenides, reality was not differentiated into parts, for exactly this reason.)

Note: Once again, we might ask whether this is really so for all questions? What about “yes/no” questions? For instance, is this class P535? But, in a sense, there is a kind of differentiation implied even here, insofar as the answer is, for instance, “Yes, it’s P535 and not P505.”

Now what has happened in all of this? In examining questions, we have encountered three kinds of non-being or “nothingness.” How can we account for that?

Well, as we know, in the end Sartre is going to have to say that consciousness is responsible for these three “nothingnesses” we have just encountered. But we still have to see just how he makes the point.

OK, we are now at the end of Section I of the Chapter.

At the beginning of Section II, entitled “Negations,” he observes that BEING-IN-ITSELF CANNOT ACCOUNT FOR THIS. We already know why not; being-in-itself is completely *positive*, completely *affirmative*.

In this Section II, he presents us with a theory that is *close* to what he regards as the correct one. But it is not quite right, and it will be instructive to see where he thinks it goes wrong.

He sets out the alternative theory on pp. 37–38. On pp. 38–42, he discusses certain *general* issues raised by the theory, and only comes back to give his verdict on the original theory at the end of the section, on pp. 42–44.

The theory Sartre has in mind is the theory held by Henri Bergson, a very important French philosopher in the early part of the century. He is sadly neglected in the American study of philosophy, but is well worth your taking a look at him.

(Bergson is one of those philosophers who has undeservedly “fallen through the cracks.” He was *enormously* important in the first half of this century, in a way people can hardly believe today, given the most of them have never heard of him. And furthermore, he is quite easy to read! His writing is not at all jargony and overly-academic. In fact, he won the Nobel Prize in literature — which is pretty amazing, given that (unlike Sartre, who also won the Nobel Prize in literature but had at least written a number of novels and other literary things) he *only* wrote in philosophy. Bergson’s influence on Sartre has never been systematically studied, but is definitely there. Again, see Christopher Vaughan’s dissertation.)

In any case, the theory Sartre is appealing to here is found in Bergson’s *Creative Evolution*, his most well known work. I have given you a passage illustrating the theory in the course packet. (It comes from Ch. 4 of the book.)

Here is what the theory says. (I’m following Sartre’s presentation here. You can verify for yourself whether it is a fair representation of what Bergson had in mind.)

Non-being, nothingnesses — the various lacks, absences, etc., that we encounter in the world — have no *objective* status out there in reality at all. They *couldn’t*. Being-in-itself, after all, (this isn’t Bergson’s term for it) is purely positive and affirmative. On the contrary, what we have is simply *negative judgments* about purely positive and affirmative being-in-itself.

The theory maintains that it is the negative *judgment* on our part that is responsible for the fact that we encounter negative features in our experience of the world.

It is easy to get confused in this chapter by Sartre’s terminology. In particular, there are at least *three* terms you may wonder about: *non-being*, *nothingness* (or the plural *nothingnesses*) and *negation*.

As I understand it, the terms ‘non-being’ and ‘nothingness’ are used more or less interchangeably. They both refer to things like *absences*, *lacks* — something *missing*, something *incomplete* or *defective* in some way. Later on in the Chapter (p. 55, in section

IV), Sartre introduces the term '*négalités*' for such things — as he describes them, these little “pools” of nothingness in the otherwise featureless desert of being. (The term '*négalité*' is a neologism on Sartre's part, and Hazel Barnes, our translator, just keeps the word intact in her English translation.)

By contrast, the term '*negation*' in this chapter refers to the *act* of forming a *negative judgment*. Roughly speaking, 'negation' here means '*negating*'.

Using this terminology, then, Bergson's theory amounts to saying that *negation* is what is responsible for *non-being* or *nothingness*.

Furthermore, according to the theory, these negative *judgments* do nothing but record *my comparison* between what is *actually* the case and what I *expected* or *imagined* or *wondered* or *feared* might be the case.

Sartre gives the example of *expecting* 1500 francs in his pocket, but when he looks, he finds he has *only* 1300 francs. The 'only' is crucial here. He finds that he has *only* 1300 francs — *and not 1500 francs after all*.

Now what's going on here? Well, basically, what we have are two facts that can be described in pretty much purely *affirmative* terms. (Actually, Sartre's more considered theory would have it that there is negativity involved even in these two putatively affirmative facts. But that deeper point only complicates the story unnecessarily here.) The two facts are:

There *are* 1300 francs in Sartre's pocket.

He *expected* 1500 francs.

So far, there's nothing *negative* about any of this. It is only when Sartre *compares* these two affirmative facts that he says “Oh, I have *only* 1300 francs — *not* 1500 francs after all.” And that process of *comparison* is a matter of forming a *negative judgment*.

Now of course, it doesn't always have to be a matter of *expecting* something other than what you find. Maybe I *wonder* whether there's any coffee left in the pot. Or perhaps I *fear* the bogeyman, and discover it's *only* the trees rustling. (I'm not actually *expecting* it to be the bogeyman.) Or perhaps I merely *imagine* the bogeyman (I'm not really afraid at all).

The general idea is that, in one way or another, I *am put in mind* of one state of affairs, and then *contrast* that with what I actually find. This *contrast*, the result of a comparative *judgment*, is what is responsible for my experiencing what I find as *only* what it is and *not* what I was put in mind of.

Now, Sartre asks, is this theory correct? Is this kind of negative *judgment* the only basis for our talking, by a kind of fiction, about *non-being*? Or is it the other way around? Must I encounter some kind of non-being out there before I can even formulate a negative judgment?

Sartre begins his answer by criticizing Bergson's theory on two grounds:

Copyright © 1996 by Paul Vincent Spade. All rights reserved. Permission is hereby granted to copy this document in whole or in part for any purpose whatever, provided only that acknowledgment of copyright is given.

- (1) First, he points out (p. 38), there are other attitudes besides judgment that are characterized by negation. There are what he calls “pre-judgmental attitudes” — that is, pre-verbal attitudes that already involve negativity.

For example, he says (p. 38), we question the carburetor. The car isn’t working, and so we get out, open the hood, and look in there quizzically. I adopt the interrogative attitude we talked about a little while ago. Thus, the three negative features that characterize all questions are present here as well. But I am not formulating a judgment. I am not yet at the level of putting anything into words (even silently).

On p. 39, Sartre gives a very interesting account of the notion of destruction. A hurricane comes along and utterly destroys a lot of property along the coastline. We don’t have to form judgments. All we have to do is open our eyes and watch it happen. To experience this event as destruction, rather than simply a rearrangement of matter, requires us to adopt a certain attitude toward it — an attitude that doesn’t presuppose any kind of judgment on our part.

In fact, as Sartre points out, the notion of destruction presupposes three kinds of non-being that parallel very closely what we found in our analysis of the question. (It is an interesting exercise to compare what Sartre says about destruction with what he says about the question, and to match up the various kinds of non-being presupposed. This actually works.)

Thus, as a result of this first line of criticism, we see that it couldn’t be just judgment that is responsible for our experiencing absences, lacks, and other forms of non-being in the world. So even at best, Bergson’s theory is not general enough to account for all the facts of our experience of negativity.

- (2) But suppose we confine ourselves to the case of judgment, by way of example. Nevertheless, Sartre says, the theory is still wrong for a second reason. If we examine a case of negative judgment closely, we’ll find that non-being or nothingness must precede my act of negating in a negative judgment.

On pp. 40–42, Sartre gives the example of the judgment ‘Pierre is not here’. Sartre has an appointment to meet Pierre in a café at such and such a time. But Sartre is delayed, and when he finally arrives, he wonders whether Pierre will still be there. As he enters the café and looks around, he comes to the conclusion “Pierre is not here.” Here there is a real judgment involved. But what does it presuppose?

According to Sartre, this judgment presupposes a twofold “nihilation.”

Digression on the word ‘nihilation’: This word can easily cause confusion. It is obviously reminiscent of the term ‘an-nihilation’, which means destruction, removing something. But that’s not what Sartre means here. In effect, what he means by ‘nihilation’ is “turning into nothing,” “turning into non-being.”

Now of course, if what we're talking about is a form of non-being to begin with, then "turning it into non-being" amounts to "making it," "producing it." So, when Sartre talks about "*nihilating a nothingness*," what he means is simply "making it into something negative" — that is, *producing* it.

In the case of Pierre's not being in the café, Sartre says there are *two* "nihilations" that must *precede* any *judgment* on my part that "Pierre is not here."

- (1) First of all, the whole scene when Sartre arrives in the café organizes itself in terms of "foreground" and "background," just as the Gestalt figure did that we discussed a long time ago.

That is, Sartre goes into the café *all set* to see Pierre. The *whole café* then serves as a kind of *background*, against which Pierre is supposed to appear. That is, the whole café is *downplayed*, reduced to the role of a mere *setting* for Pierre. The café, in other words, is "nihilated" — it's "made negative." Here is part of what he says (p. 41):

When I enter this café to search for Pierre, there is formed a synthetic organization of all the objects in the café, on the ground of which Pierre is given as about to appear.... Each element of the setting, a person, a table, a chair, attempts to isolate itself, to lift itself upon the ground constituted by the totality of the other objects, only to fall back once more into the undifferentiation of this ground; it melts into the ground. For the ground is that which is seen only in addition, that which is the object of a purely marginal attention. Thus the original nihilation of all the figures which appear and are swallowed up in the total neutrality of a *ground* is the necessary condition for the appearance of the principal figure, which is here the person of Pierre. This nihilation is given to my intuition; I am witness to the successive disappearances of all the objects which I look at — in particular of the faces, which detain me for an instant (Could this be Pierre?) and which as quickly decompose precisely because they "are not" the face of Pierre.

That's the first "nihilation." (It's "first" in the sense that it is *presupposed* by the other one; there's not necessarily any *temporal* priority here.)

- (2) The second "nihilation," the second negative feature of this situation, is of course the fact that Pierre *fails* to emerge against that backdrop; he is *not there*.

Of course, if Pierre *were* there, this second nihilation would not occur. But the first one would continue to hold.)

Now, the discussion on these pages is sensitive and very nicely done, but what exactly is the point of it all?

Well, the point is that the judgment ‘Pierre is not here’ that I make amounts to a discovery of a prior “nothingness” or lack in the café, not a producing of that “nothingness” or absence, as Bergson’s theory would have it.

In other words, these little “nothingnesses” in the world — lacks, absences, failures, destructions, etc. — all come on to me (that is, they appear to me phenomenologically) as being something I discover or learn. They appear as objective things, things I can be wrong about, and so about which there is a certain risk. (Perhaps Pierre is there in disguise.)

All of these features — the being able to learn, the risk, the objectivity — are features that characterize perception, as opposed to imagination or conception, as we saw in our earlier passage from *The Psychology of Imagination*.

But note that none of this means that Pierre’s absence from the café is a fact that is in any way independent of consciousness. It isn’t; Sartre agrees with Bergson on that. Consciousness constitutes Pierre’s absence from the café just as it constitutes all the other features of phenomena. All that follows from what we are now saying is that Pierre’s absence from the café is not something subjective, like imagination. It’s objective, like perception. And this is where Sartre disagrees with Bergson.

You may well wonder how that can be, since Sartre himself holds that consciousness constitutes Pierre’s absence from the café just as much as it constitutes any other phenomenon. If I’m the one who’s doing it all along, then how can I learn that Pierre is not in the café? If I’m the one who’s setting things up that way, then how can I be mistaken about it? How can I be surprised?

(We raised the same question in effect a long time ago, when we were talking about the “objectivity” of perception back in *The Psychology of Imagination*. We are now ready to start on an answer to it.)

But of course, once we think about it, there are lots of ways in which we are surprised by what we have done, are mistaken about it or learn from it.

Think of a novelist, for instance, who is writing a complicated novel with lots of richly developed characters in the story. (The model of writing a novel is, like the movie-theater metaphor, an excellent model for Sartre’s theory of consciousness.)

Novelists frequently report that they are surprised to find that their characters seem to take on a life of their own. They develop a kind of inertia, they come to take on personalities, characters of their own. So true is this that if the novelist tries to make a certain character behave in a certain way in his story, he finds that the character resists. It’s just not right!

Of course, in a perfectly obvious sense, the novelist is in complete control all along. If he wants to make the character behave in a certain way, all he has to do is write the words down, and it’s done! And there’s also a perfectly obvious sense in which nothing happens in the novel, nothing is true about the characters in the novel, except what the novelist makes true by his words.

Thus, the novelist (so to speak) constitutes his characters. He's the one who made them what they are. And yet he can be surprised by what he has written. He can be mistaken about the kinds of characters he has produced; he may think they have certain kinds of personalities, but come to discover that they are quite different.

None of this is to suggest that there is anything in the novel that was not put there by the novelist. And it does not mean the novelist put it there unconsciously, although some people would like to express it that way. The novelist wasn't unconsciously writing all those words down; he was awake the whole time and chose each word with the greatest care!

On the contrary, for Sartre all it means is that some of what the novelist put into his novel he put there unintentionally. And so he can be surprised by it, mistaken about it. But to say it was unintentional is not to say he didn't do it, and it isn't to say he did do it unconsciously. It's just to say that what he consciously did had an unexpected outcome. What's so difficult about that?

In the end, this is the main point of Sartre's discussion of Bergson's theory. He is rejecting a theory that would equate what is done consciously with what is done intentionally (= "on purpose," not "intentionally" in the sense of the theory of intentionality) or deliberately. Bergson's theory that non-being, nothingnesses, are the result of a negative judgment would have this effect: I can never be surprised by such judgments, or be wrong about them.

It is important to see the point here, because otherwise it is easy to get confused about the overall purpose of Sartre's Chapter 1. You might think that in Section 2 of the chapter, where he is discussing Bergson, he is arguing that non-being, nothingnesses, are not the products of consciousness, as Bergson thought. But then, in the last section of the Chapter, Sartre goes on to argue that the for-itself is the origin of nothingness — a conclusion I've already given you. And in that case, you might well wonder: Which is it?

But that is not what Sartre is doing at all. In Section 2, he is not arguing that non-being is not a product of consciousness. He is arguing only that it is not subjective in the way Bergson would have it.

Hegel and Heidegger

We can skip briefly over Sections 3–4 of the chapter.

Section 3 is a critique of the "dialectical" concept of nothingness, which is to say Hegel's notion of nothingness. I don't want to dwell on this, partly because it would take us too far astray, and partly because I don't know enough about Hegel to do justice to it.

But I do want to call your attention to a statement on p. 49, at the very end of the section:

Non-being exists only on the surface of being.

(That is, on the surface of *being-in-itself*.) In other words, non-being is something *imposed* on being-in-itself. *We* put it there.

Section 4 is a critique of the “phenomenological” concept of nothingness, which does not here refer to Husserl, as you might expect. It refers to Heidegger (who was really the first of the so called “phenomenological existentialists”).

Basically, this is the view that “nothingness” or “non-being” is to be viewed as something *outside* being, something *separated* from being. Heidegger likes to talk as if reality were, so to speak, a *little island of being in the middle of the great sea of nothingness*.

Well, that’s fine, Sartre says. But we don’t want to think that’s the *only* way we encounter nothingness or non-being — as *the great void beyond the edge of being*. No, we also encounter nothingnesses or non-beings right in the *middle* of being, *all through it*. These are the little *pools* of non-being that Sartre calls “*négatités*,” things like absences, lacks, failures, etc. Heidegger, Sartre charges, has taken no account of these.

In this connection, I should mention a hilarious, short article in *The Encyclopedia of Philosophy* by P. L. Heath (vol. 5, pp. 524–525) on “Nothing.” Heath distinguishes the Heideggerian notion of “nothing” from the Sartrean notion, which might be called the “Swiss cheese” view of nothingness. Here is what he says, in part:

The friends of nothing may be divided into two distinct though not exclusive classes: the know-nothings, who claim a phenomenological acquaintance with nothing in particular, and the fear-nothings, who, believing, with Macbeth, that “nothing is but what is not,” are thereby launched into dialectical encounter with nullity in general. [For Heidegger, the fear of death is a fear of total annihilation, a fear of nothingness.] For the first [for example, Sartre], nothing, so far from being a mere grammatical illusion, is a genuine, even positive, feature of experience. We are all familiar with, and have a vocabulary for, holes and gaps, lacks and losses, absences, silences, impalpabilities, insipidities, and the like. Voids and vacancies of one sort or another are sought after, dealt in and advertised in the newspapers. [By advertising “vacancies,” he means something like “Apartment for Rent.”]

He concludes the article:

If nothing whatsoever existed, there would be no problem and no answer, and the anxieties even of existential philosophers would be permanently laid to rest. Since they are not, there is evidently *nothing to worry about*. But that itself should be enough to keep an existentialist happy. Unless the solution be, as some have suspected, that is it not nothing that has been worrying them, but they who have been worrying it.

As an illustration of Sartre's view, consider the notion of distance. (Sartre himself uses this example at the end of § IV of the Chapter. What this notion of distance. It will play an important role as things develop in our account.)

Think of the road between Bloomington and Indianapolis. There are two ways we can think of this road:

- (1) We can think of it as the road, which is terminated at one end by Bloomington and at the other end by Indianapolis. If that is the way we are looking at it, then the road itself appears as positive, whereas the end-points are negative: they are where the road terminates.
- (2) Or we can think of the same configuration as consisting of Bloomington on the one hand, and of Indianapolis on the other, and the road is what separates them. If that is the way we are looking at it, then the two end-points appear as positive, and the road itself now comes on as negative.

Recall the Gestalt figure we discussed earlier, and how we can flip-flop from one way of viewing it to the other. Both in the case of the Gestalt figure and in the case of distance, we are the ones who make the overall phenomenon what it is for us.

Now the notion of distance (viewed in either way) is one of those phenomena Sartre calls "négativités." They are beings that appear to us riddled with non-being.

The Origin of Nothingness

We are now ready to look at § V of Chapter 1, "The Origin of Nothingness."

We already know roughly how this is going to go:

- (1) Nothingness cannot come from being-in-itself, as we've seen. (Being-in-itself is purely affirmative, and doesn't do anything.)

He goes on:

- (2) Neither can nothingness — lacks, absences, etc. — produce itself, or as Sartre says, "nihilate itself."

The second claim is part of what Sartre develops in the preceding section. Basically, it is a criticism of Heidegger. Heidegger had said, "Das Nichts selbst nichtet." That is (roughly), "Nothing itself noths." (Rudolf Carnap had a lot of cheap fun at the expense of this phrase.)

Of course, ‘to noth’ is not a normal verb in English any more than ‘nichten’ is in German. But the basic idea is that “noth-ing” is what Nothing does. (“What’s it doing? It’s nothing.”)

Apart from the verbal cuteness here, the basic idea for Heidegger is that nothingness is somehow self-producing, or as Sartre says, it “nihilates itself.” But Sartre will have none of that.

Basically, this much is just an elaboration of what Parmenides had already said a long time ago, and for pretty much the same reasons.

But instead of just rejecting non-being or nothingness as a dangerous and paradoxical illusion, as Parmenides did, Sartre wants to push further. Paradoxical or not, we encounter non-being, négatités — absences, lacks, etc. — and have to account for it. Even if these négatités were pure illusions, as Parmenides had said, we should still have to account for them somehow.

So, Sartre says — and now I’m going to try to give you a kind of explication of the almost unreadable passage on p. 57,

It follows therefore that there must exist a Being (this can not be the In-itself [for the Parmenidean reasons we’ve already seen], of which the property is to nihilate Nothingness [that is, to produce it, to turn it into nothingness], to support it in its being, to sustain it perpetually in its very existence, *a being by which nothingness comes to things*.

Furthermore — continuing our explication — he says that this special being must be one that is itself shot all through with nothingness — with absences, lacks, etc. If it weren’t, if it were purely positive, it would be just being-in-itself all over again. Hence it must be both. It cannot be mere nothingness; it has to be both a *being* and yet soaked all through with nothingness!

(This much is contrary to Bergson’s theory, which — as we saw — tried to get nothingness somehow out the juxtaposition of two purely positive facts.)

Thus, Sartre says (pp. 57–58):

The being by which Nothingness comes to the world must be its own Nothingness.

All this means two things for us:

- (1) Consciousness is going to have to be very paradoxical. In our phenomenological description of it, we are going to have to say things that seem incoherent and contradictory — and they are incoherent and contradictory. This much we have already seen (although here is where we see *Sartre* saying it for the first time).

Consciousness cannot be explained in the sense of giving a coherent account of it.

- (2) If we are going to be able to grasp adequately what is going on when we encounter négativités in our experience of the world, we are eventually going to have to turn to examine consciousness. That is, we are going to have to adopt a reflective attitude, to make consciousness our object and examine it. We see here a motivation for the predominately reflective tone of the rest of the book.

The examples we have considered so far all started off at least as non-reflective. When we talked about distance, about Pierre's absence from the café, etc., we were proceeding non-reflectively. We were not *especially* thinking about consciousness then (although we said some things incidentally about that too), but rather about distance, about Pierre's absence, etc.

Now — near the end of Ch. 1 — we are about to change that approach, and to adopt an explicitly reflective approach. We have got to the point of realizing that if we are ever going to get a grip on distance, on Pierre's absence, we are eventually going to have to adopt a reflective approach.

This is the main conclusion of § V of the chapter. But before we go on to look at some other things that go on there, let's pause to consider an obvious question:

What we've just looked at looks a lot like an argument: Consciousness has to have negativity running through it, because otherwise there would be no way for the appearances of negativity in the world to get there. But what is Sartre the phenomenologist doing arguing? I thought he was supposed to confine himself to pure description.

Well, I'm afraid we are going to see a lot of this in Sartre. And we can say one of two things about it. On the one hand, we can say that Sartre is just being pretty sloppy about his phenomenological method, and that he really is trying to be far more systematic and theoretical than strict phenomenology would allow.

On the other hand, we might also say that these apparent arguments are not meant to be the real bases on which Sartre's theory rests. Perhaps they are just meant as heuristic devices, as ways of getting you to see the point he is making. The point he is making — in this case, that consciousness is riddled with nothingness — is something that can be seen and described on its own, in the strictest phenomenological way. But first you have to see the point. And of course, it doesn't make any difference how he gets you to see the point, as long as you do see it. End of point.

There are several other things that go on in this last section of Ch. 1, and I want to look at some of them briefly.

There is, for instance, the very nice discussion of the notion of Anguish, which is something we will see a lot of later on. In this discussion, Sartre is concerned to contrast anguish from simple fear.

Anguish is fear of ourselves, fear of our own freedom.

In the discussion of this, Sartre gives us two important analyses that both illustrate the theme of anguish and also serve to lead us into Ch. 2. These are the discussions of vertigo and the case of the gambler.

In Ch. 2 (“Bad Faith”), Sartre is going to give us an absolutely brilliant discussion to try to show us something I told you a long time ago: that consciousness is not what it is, and is what it is not. The discussion there will attempt to show that this is so in a completely literal sense, with no trick whatever.

But here, at the end of Ch. 1, he gives us the two examples of vertigo and the gambler, which lead us right up to the same point. But here they do look like tricks, like merely verbal points that rely on playing fast and loose with the tenses of verbs. They’re not, but that’s the way they look at first. Sartre is in effect setting us up.

The Gambler

Let’s start with the discussion of the gambler (pp. 69–70). (Sartre in fact treats this one second in order.)

A certain man is a compulsive gambler. He spends all he has at the casino or at the racetrack. His habit is ruining his marriage, his children are starving, and things have really come to a crisis.

The man is no fool, and realizes the seriousness of his habit. And so he resolves to stop gambling, and his resolve is quite sincere. But the following day, he approaches “the gaming table,” and what happens? He is tempted.

He looks back into the past and sees himself yesterday. (Note: He is reflecting here.) Here’s what goes on in his mind:

That man back there in the past is me. It’s not someone else, after all; I recognize myself in that past man. And yet, in the sense that matters right now, that man is not me. That man has good resolutions that speak to him and are persuasive. But those resolutions do not affect me one bit, unless I make those resolutions anew — now. I do not find his resolutions affecting me.

So, here is a case in which I am that man, and yet am not that man. Thus, consciousness is what it is not.

Granted, the paradox looks merely verbal at this point. All that's really going on is that I'm not what I WAS, and that is hardly surprising. All it means is that I've changed. It's only by overlooking the obvious role of the passage of time here that we can make this situation look like a paradox.

Well, maybe. But let's look at what Sartre says about this here.

Consciousness in this instance is separated from itself (from its past self, to be sure). What is it that separates consciousness from itself here? Well, you say, it's time. Yes, but let's look at the question slightly differently.

What separates me from myself here? That is, what prevents me from being that man I see back there in the past — from being that man in such a strong sense that his resolutions are also my resolutions right now? What prevents me from adopting his resolutions as my own? Answer: NOTHING. Nothing whatever. I am perfectly free to make those resolutions anew if I choose to do so. Nothing is holding me back. Of course, by the same token, nothing is forcing me to renew those resolutions. All of which is just another way of saying I am free with respect to these resolutions. (We begin to see here the profound link between freedom and nothingness, a link that will be developed throughout the rest of the book.)

This freedom produces Anguish, a kind of profound panic at the thought that these matters really are up to us. Here is what Sartre says (pp. 69–70):

In reality — the letters of Dostoevsky bear witness to this — there is nothing in us which resembles an inner debate as if we had to weigh motives and incentives before deciding. The earlier resolution of “not playing anymore” is always there, and in the majority of cases the gambler when in the presence of the gaming table, turns toward it as if to ask it for help; for he does not wish to play, or rather having taken his resolution the day before, he thinks of himself still as not wishing to play anymore; he believes in the effectiveness of this resolution. But what he apprehends then in anguish is precisely the total inefficacy of the past resolution. It is there doubtless but fixed, ineffectual, surpassed by the very fact that I am conscious of it. The resolution is still me to the extent that I realize constantly my identify with myself across the temporal flux, but it is no longer me — due to the fact that it has become an object for my consciousness. I am not subject to it, it fails in the mission which I have given it. What the gambler apprehends at this instant is again the permanent rupture in determinism; it is nothingness which separates him from himself; I should have liked so much not to gamble anymore; yesterday I even had a synthetic apprehension of the situation (threatening ruin, disappointment of my relatives) as forbidding me to play. It seemed to me that I had established a real barrier between gambling and myself, and now I suddenly perceive that my former understanding of the situation is no more than a memory of an idea, a memory of a feeling. In order for it

to come to my aid once more, I must remake it *ex nihilo* [= out of nothing] and freely. The not-gambling is only one of my possibilities, as the fact of gambling is another of them, neither more nor less. *I must rediscover* the fear of financial ruin or of disappointing my family, *etc.*, I must re-create it as experienced fear. It stands behind me like a boneless phantom. It depends on me alone to lend it flesh....

So my own *freedom SEPARATES* me from myself, so to speak puts me at a *distance* from myself. (We have already seen how the notion of *distance* involves negativity.) And that separation, that *nothingness*, that *distance* is somehow a product of consciousness itself as part and parcel of its own freedom.

Of course, in this case we are talking about being separated from my *past* self, which I am *reflecting on* while I am being tempted to gamble again. That is, consciousness is separated from its *object*, which in this case happens to be its past self.

Now Sartre thinks this feature by which consciousness *separates* itself and *isolates* itself from its objects is *pervasive* of consciousness. It is a *characteristic* feature. (Recall that, for Sartre, an important feature of *intentionality* is that it is *irreflexive*.)

This is why I said a long time ago that, for Sartre, the *best model* for consciousness is the *stepping back* and *separating oneself* from an object, the *taking a point of view* on an object, the *putting oneself at a distance* from an object. (Recall the discussion of “distance” earlier; it’s no coincidence.)

Please keep all these threads in mind as we go on. Right now, they look like a hopeless — and unconvincing — tangle, but things will get better.

Vertigo

The example of the gambler involved the past. Sartre also gives a similar example that involves the future. This is the example of *vertigo* (beginning on p. 66.)

I stand at the edge of a precipice and look down. I begin to feel a little dizzy. What’s going on here?

It can hardly be that I am, in any objective sense, afraid of *falling* over the edge (at least not in most cases). Let’s suppose the ground is reasonably firm, the wind is not blowing so hard it’s going to puff me over the rim, there’s no real likelihood of an earthquake. None of that is what is really causing my dizziness.

No. For Sartre, what is bothering me is not the possibility that I might *fall*; it’s the possibility that I might *jump*. There is no other way to accommodate the facts.

I look, as it were, down there into the future and see myself tumbling head over heels over the edge to my death. Now, of course, in an obvious sense, I am *not* that man I see in the future. I’m up here on the top, reasonably intact; he’s down there on the bottom, all

smashed. But in another obvious sense, I am that man I see down there in the future. That is, I recognize myself in that moment. If I didn't somehow recognize myself in that future man, why would he bother me so much? The kind of vertigo I feel at the prospect of my tumbling over the side is quite different from whatever I might feel at the prospect that someone else might fall or leap over the ledge.

No, that's me. And yet, it's not me. I am what I am not, and I am not what I am. And, just as in the case of the gambler looking into the past, so too here in the case of the future, there is a way of putting this in terms of freedom:

What is it that prevents me from being that man in the future in so strong a sense that I too propel myself over the side? Answer: Nothing. What is it that compels me to do it? Nothing. In short, nothing SEPARATES me that prospect. And that nothingness is just another way of talking about freedom.

And in fact, the closer I get to the edge of the cliff, the more obvious it is that nothing prevents me from actually doing it. It's as if the literal distance between me and the edge is a kind of symbol of my own freedom. And that is what's so scary, what produces the dizziness or vertigo. This fear of my own freedom is what Sartre calls "anguish."

This notion of the fear of freedom is something we'll see much more of very soon. But, for the present, notice that, just as, in the case of the gambler, consciousness is separated from the past self it is reflecting on, so too here: consciousness is separated from the future self it is reflecting on.

Once again, then, we have the nothingness that separates consciousness from its objects. In these cases, of course, we are talking about reflection, and the object is my past or future self.

But the same point holds for pre-reflective consciousness. We "question" the carburetor, to use Sartre's own example. This requires us to draw back from the carburetor (perhaps even literally to draw back), to separate ourselves from it in order to consider it objectively, to put ourselves at a distance from it.

All this amounts to saying that consciousness, as it were, secretes a kind of nothingness (again, a kind of "distance") that isolates it from its objects.

Now — and this is why I have dwelt on this for so long — Sartre thinks the fact that consciousness can withdraw in this way, out of reach of its object — whether that object is something in the world, or whether it's me in reflection — IS PROOF OF ITS FREEDOM. Or perhaps, since a phenomenologist shouldn't be talking about proving things, I should say: THIS IS WHAT SARTRE MEANS BY FREEDOM.

Here is what he says on p. 60:

For man to put a particular existent out of circuit [that is, to put it out of reach, to separate himself from it] is to put himself out of circuit in relation to that existent. In this case he is not subject to it; he is out of reach; it can not act on him, for he has retired *beyond a nothingness*. Descartes

following the Stoics has given a name to this possibility which human reality has to secrete a nothingness which isolates it — it is *freedom*.

Of course, at this point that is just a bald claim. We still have work out all the details. In particular he have to ask whether *freedom* is simply freedom from being determined by my *objects*. But keep this passage in mind. It is one of the most explicit statements I know of in all of Sartre about how the notions of freedom and non-being are linked.

Now we have already seen that the awareness of our own freedom produces *anguish*. Moreover, since *every* act of consciousness is free for Sartre, and since there is nothing *unconscious* about consciousness, we ought to be *constantly aware* in whatever we do that we are acting freely, with nothing to compel us and nothing to prevent us from doing whatever we choose. It would seem to follow, therefore, that we are *constantly in a state of ANGUISH*.

And yet, there is a very interesting fact: We spontaneously and almost automatically act as though we were *not* free, as though we *were* compelled. We try to find *excuses*, *pass the blame*, *avoid our responsibility* — or, as Sartre says, to “*flee our anguish*.”

This is going to be a *very* interesting phenomenon. We are trying to *fool ourselves*, to distract ourselves from the fact that we are aware of our own freedom and responsibility. We are *pretending to ourselves* that we are *not* free, in the hope of convincing ourselves.

This is the behavior Sartre calls *Bad Faith*, and it is what is otherwise known as “*self-deception*.” And it is going to be Sartre’s main proof that consciousness is contradictory and paradoxical. It is going to be what finally *shows* us that the for-itself is what it is not and is not what it is — and that this is *literally* true, without any funny business.

The examples we have seen of this up to now — the examples of the gambler, and of vertigo — have all looked frankly like *tricks* involving some fast and loose playing with tenses. They have all involved the *separation* of consciousness from its intentional *objects* (and recall that intentionality is *irreflexive* for Sartre) — and in particular from its own past or future *reflected on* self.

But now, in the discussion of *bad faith*, we get a new kind of *separation* of consciousness. And this time it is *not* just a separation of consciousness from its *object* — whatever that object is. This time we are going to find that consciousness is separated from *itself* — not from its *past* or *future* self, but its own *present* self, and *not as an object of reflection*.

In short, we will find that negativity characterizes not just the relation between consciousness and its object, but is there in the very *act* of consciousness itself (the “being” of consciousness). So we are pushing our investigation deeper and deeper.

Bad Faith (Self-Deception)

Let's pause to see exactly what Sartre means by "bad faith." Basically, we said, it is self-deception. (Or perhaps it's best to say it is self-deception *about ourselves*. There is some unclarity in my mind about just how far Sartre is willing to extend the term 'self-deception'.) And there are lots of forms of it, some of them quite ordinary and commonplace, some of them pretty subtle.

Sartre gives lots of examples in the Chapter on Bad Faith, but it's not hard to come up with examples of your own.

Your girl-friend or boy-friend is cheating on you, and you know it. There's really no doubt about it. But what do you do? You don't want to believe it. And so you tell yourself various stories in an attempt to convince yourself that things are not the way you know good and well they are.

What is going on here?

It's fairly easy to get started.

Bad Faith or self-deception is a kind of lie — a lie we tell ourselves. Thus, among other things, it should exhibit all the structure of lies in general. So let's start by looking at The Lie (in general):

Any lie involves two sides: The Deceiver and The Deceived.

(We're talking here about a lie that works. Sometimes, of course, an attempt at lying doesn't succeed and no one is fooled. But let's set those cases aside.)

These two poles are related as follows:

- (1) The Deceiver knows the truth he is lying about. (If he doesn't, he's not really lying; he's just mistaken.)
- (2) The Deceived doesn't know the truth. (Remember, we're talking about a lie that works, where someone is fooled.)

(There are other factors involved too, but these are the main ones. For instance, I know my unlisted telephone number, and you don't. But that doesn't mean I'm lying to you about it. But this is enough for present purposes.)

In the case of self-deception, the lie to oneself, this simple and unproblematic structure becomes paradoxical. For in that case, The Deceiver = The Deceived.

Thus, one and the same person both knows the truth and doesn't know the truth, and that is a contradiction.

That's enough to get us started. Most of the rest of the chapter on "Bad Faith" is spent in illustrating the various ways in which this odd situation comes about, and in arguing that

there really is a contradiction here, that it is not merely a trick that can be avoided if we just make certain moves and distinctions.

For instance, Sartre says, the Freudian notion of an unconscious is frequently appealed to as a way of avoiding this paradox. But that won't work.

Here we get one of Sartre's most sustained attacks on the notion of the Freudian unconscious. There are other, less sustained and detailed discussions in his section on "Existential Psychoanalysis" in *Being and Nothingness*, in *Transcendence of the Ego*, and in his *The Emotions: Outline of a Theory*. The present discussion begins on p. 90:

To escape from these difficulties people gladly have recourse to the unconscious....

Now we know already, of course, that Sartre rejects the notion of a Freudian unconscious entirely. But that is not his point here. The argument here is not that there is no such thing as a Freudian unconscious (although Sartre believes that). He is here arguing instead that, even if there were, that wouldn't help one bit to get around the contradiction we are talking about.

Here is why people might think it would: You might think that the contradiction is only apparent. There is, you might say, *a sense* in which the same mind both knows and does not know the same truth at the same time. But this appears contradictory only because we have not cut our analysis finely enough. If we look more closely, you might say, we will see that there are various parts and subdivisions to the mind, and that one part knows the truth while another part does not. That's no more contradictory than the fact that my one hand might be in motion while the other one is not.

Let's do a short Freud lesson.

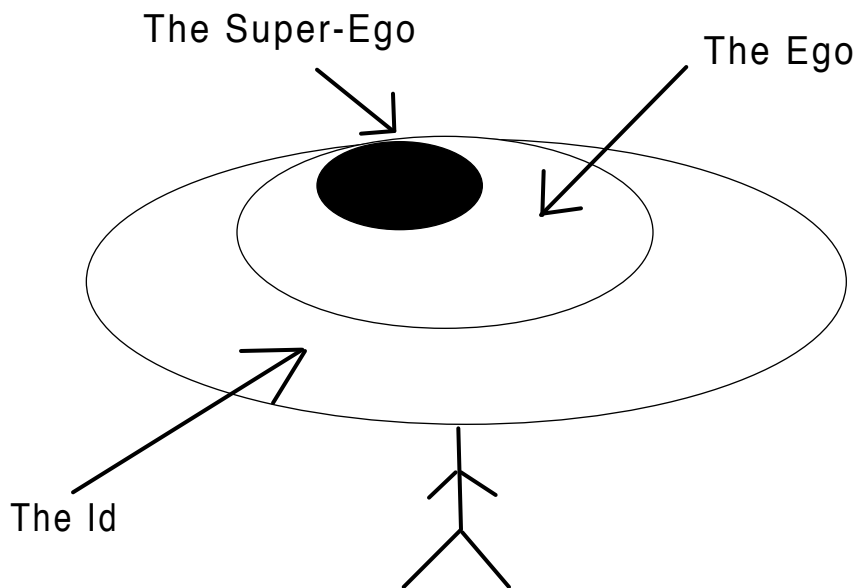
- (1) We start with the Id (= the "It"). This is an unorganized mass of drives and instincts — of libido urges. It is ruled by one great principle: The Pleasure Principle. All the drives and instincts in this psychic pool seek to be fulfilled, satisfied. (Compare Nietzsche's "Will to Power.")
- (2) As consciousness develops in the child, a part of the Id becomes organized, and the psyche begins to develop. This organized region of the Id is called "the Ego" (= the "I"). (Do not confuse this with the Husserlian "Transcendental Ego." The term is — partly — the same, and they serve many of the same functions, but the theories are really after quite different things. What we have here is closer to our earlier notion of the "psychological ego.")

The Ego is ruled by another great principle: The Reality Principle. The Reality Principle is what tells the psyche to wait to satisfy those Id drives until the appropriate time.

If the Pleasure Principle says “Yes! Now,” the Reality Principle in effect says “Wait!” Now, although it looks at first as though the Id and the Ego are in conflict, they really aren’t. The Ego is a subregion of the Id (the organized part), and is really at the service of the Id. It says “Wait!” only because that is in the long-term best interest of the Id’s own desires.

- (3) There is a third structure of the psyche that Freud recognized, called “the Superego.” It is specialized function or subpart of the Ego, and is what we normally call the “conscience” (not “consciousness”). Certain Id drives are so strong and so dangerous to the long-term health of the psyche that the Ego, in its special function as the Superego, says “No!” to them — not just “Wait!,” but “No!.”

Once again, the Superego is really at the service of the Id. In the end, it is the Id with its raw libido energy that is the motor driving the whole thing.



Note: It is only at the level of the Ego (including the Superego) that we have consciousness. This does not mean that all “Ego processes” are conscious ones for Freud, but they all can be; they can all be brought to consciousness. (The unorganized Id drives cannot be brought to consciousness in the same way, although we can become aware of them in the sense of inferring that they are there.)

So much for the various structures Freud distinguished in the psyche. Here’s how they apply to the notion of Bad Faith or self-deception:

Copyright © 1996 by Paul Vincent Spade. All rights reserved. Permission is hereby granted to copy this document in whole or in part for any purpose whatever, provided only that acknowledgment of copyright is given.

Freud noticed a very curious phenomenon in people, what he called “repression.” (This was a *general* term for him.) As we have seen in the case of the Superego, certain drives or instincts are so dangerous (or are perceived to be — whether they really are or not makes no difference) that the Ego, in its role as Superego, says “No!” And in some cases, the drives are so dangerous that not only should they not be satisfied, but they should not even be allowed into consciousness. They are just too dangerous!

But what happens to those drives then? They don’t just go away quietly and disappear. They are stronger and more insistent than that. And besides, the Ego and Superego, like everything in the psyche, are driven by those Id forces, and are in the long-term service of the Id.

What happens, Freud observed, is that these drives get redirected by the Ego, and satisfied in various safe but symbolically appropriate ways.

For example, I work in an office, let us say, and just hate my boss. I am bubbling over with seething resentment. But it’s dangerous for me to be explicitly conscious of this, since I might do or say something rash. (That’s the “Superego” speaking there.)

So my hatred is redirected and transformed, say, into a kind of nervous kicking of the office furniture — as though taking out my resentment on something that represents my boss and everything he stands for.

This “redirecting” (“repressing”) task, on Freud’s theory, is performed by the Ego (in some passages, Freud specifies more specifically the Superego). See the passages in the course packet on this.

Freud also observed that the same thing happens in dreams. There the process is called dream-censorship. And we have all heard about how that goes. (Sartre, oddly, uses the term ‘censor’ for any kind of repressing activity like this. Freud’s own terminology seems to restrict the term ‘censor’ to the dream-situation.)

And — an especially interesting case — Freud also observed that it happens all the time in the clinical situation of psychoanalysis. There he calls it “resistance.” As the patient and the analyst get closer and closer to the real source of whatever is bothering the patient, the level of tension rises and the patient begins to take various kinds of evasive behavior — to avoid letting this horrible truth rise to the level of consciousness. He begins to talk about irrelevant things, to lapse into complete silence, to try to change the topic, etc.

For Freud, the level of this resistance is a clue to the effectiveness of the analysis. A good analyst will keep the resistance as high as possible, since that means he is getting close to the main point. (Of course, the ultimate resistance is for the patient simply to stop coming. And the analyst must stop short of that.)

On all this, I have included some relevant passages from Freud in the course packet. Please go look at them now.

How does this apply to Sartre’s discussion of Bad Faith?

Well, plainly, the whole point of these evasive maneuvers is to keep consciousness — which is to say, the Ego (or Superego) — from becoming aware of what is really going on down there in the Id. The Ego is *deceived*. What we have here is *The Lie* — with all its duality: The Ego is deceived about the truth, which is conveniently kept down there in the Id (it is kept there by *repression*, in all its forms). The Id does know the truth.

So the Ego can be *separated* very nicely from the truth it is not supposed to know.

But, Sartre says, *it won't work!* What is it that does the *repressing*? What does the *dream-censoring*? What does the *resisting* in the clinical situation? *On Freud's own theory*, it is the *Ego* that does this (perhaps in its special role as Superego, but that makes no difference). (See the selection of passages from Freud in the course packet, to verify that this really *is* Freud's theory, and that Sartre is not just misinterpreting him here.)

The *Id* certainly isn't going to repress, censor, resist, its own drives. It doesn't want to hide anything. It wants these drives to be satisfied — and satisfied *now!* It *has* to be the Ego that does the resisting.

And in order to do this resisting effectively, the Ego must *know* exactly what is going on. It must *know* what it is that cannot be allowed into consciousness, in order to be able to take appropriate evasive action when necessary. The Ego, after all, is *very clever* at avoiding the real point in clinical analysis. It knows *exactly* what is going on!

But, on the other hand, the Ego — with its *conscious* Ego processes — is also supposed to be exactly what does *not* know what is going on. It is supposed to be *fooled*.

In short, *all the Freudian machinery of the mind has not succeeded in avoiding the paradox*. We are left now with the *Ego* that both knows and does not know the truth.

In short, the contradiction has *not* been avoided; it has only been *localized*.

Now you may think this is not very persuasive, that the appearance of contradiction persists only because we still have not cut our analysis finely enough. What we need to do is to add yet further refinements — so that, say, we have one *part* of the Ego being deceived by a second and *distinct* part.

Well, perhaps you can make some headway like that. But, in any event, Sartre has another piece of evidence that undercuts this whole approach. *There are certain cases that the Freudian approach simply cannot account for*.

Sartre cites some clinical reports by Wilhelm Stekel (a member of Freud's Wednesday night Vienna circle of psychoanalysts), concerning a frigid woman. According to her husband, she seemed to give all the objective signs of pleasure during sex. And yet she insisted she didn't. And there is no reason to think she was anything less than candid about this — she *believed* what she was saying.

Sometimes people object: Why believe the husband? Maybe he was a brute who was so involved in his own self-esteem that he didn't realize that this was torment for his wife. Well, yes, of course. Perhaps that's so. But it misses the point entirely. It doesn't really make any difference whether the case Sartre describes from Stekel actually occurred the

way Sartre — or Stekel — describes. The point is that, in that description, we can all recognize a particular kind of behavior, sexual or not, that we all engage in — and that the Freudian set-up doesn't account for. The behavior is a kind of attempt to distract ourselves from something we know good and well.

What is so important about this case?

What is important is: What is this woman deceived about? She's deceived about whether she feels pleasure or not. And we're not talking here about pleasure in the sense of "mental satisfaction"; we're talking about physical pleasure, which is, after all, a matter of nerve endings. Now there's nothing wrong, let us say, with the woman's physiology; all the nerves are intact and functioning. Of course she feels pleasure, as far as that goes.

Now of course, the reason she says she doesn't feel pleasure is no doubt because of some deep-seated Freudian complex, some "hang-up" that is getting in the way here, and will not allow her to admit what in the end is simply a matter of physiology.

All that may be correct, and there may be (let us say, for the sake of argument) such a deep "complex." But that's not what the woman is deceived about — or least that's not all she is deceived about.

The point is, it's not just the complex she won't admit to herself, it's also the pleasure. And the kind of pleasure we're talking about is not something hidden deep down there in the Id; it's not something down there in the inside; it comes from outside, and is something the woman can't help but be conscious of while it is occurring.

The Freudian mechanism of "repression," "censorship," "resistance," cannot be responsible for this, since its job is to filter out dangerous messages from the Id. But that's not what is going on here, at least not all that is going on here.

The Freudian theory mixes up the reason or motive for the deception (= the "complex") with the truth we are deceived about. They are not always the same. Freud can perhaps explain the former (although Sartre in the end thinks not), but he cannot account for the latter in this case.

On Sartre's analysis, what we have here is not a just matter of repressing something down there in the Id. What we have is a matter of the woman's also distracting herself from something she's very much conscious of. And of course that's the pattern we can all recognize in ourselves.

In short, Bad Faith — with all its contradiction — is back again, and cannot be avoided.

Sartre gives lots of other examples of the infinitely varied ways in which all this can be worked out in practice. One of them is the famous portrait of The Waiter.

The Waiter

Sartre's famous example of The Waiter begins on p. 101.

Copyright © 1996 by Paul Vincent Spade. All rights reserved. Permission is hereby granted to copy this document in whole or in part for any purpose whatever, provided only that acknowledgment of copyright is given.

Sartre is sitting in a café, and he has been sitting there a long time. He's been observing the waiter. And finally he realizes that there's something odd about this waiter (pp. 101–102):

His movement is quick and forward, a little too precise, a little too rapid. He comes toward the patrons with a step a little too quick. He bends forward a little too eagerly; his voice, his eyes express an interest a little too solicitous for the order of the customer. Finally there he returns, trying to imitate in his walk the inflexible stiffness of some kind of automaton while carrying his tray with the recklessness of a tight-rope-walker by putting it in a perpetually unstable, perpetually broken equilibrium which he perpetually reestablishes by a light movement of the arm and hand. All his behavior seems to us a game. He applies himself to chaining his movements as if they were mechanisms, the one regulating the other, his gestures and even his voice seem to be mechanisms, he gives himself the quickness and pitiless rapidity of things.

Finally Sartre says (p. 102):

He is playing, he is amusing himself. But what is he playing? We need not watch long before we can explain it; he is playing at being a waiter in a café.

Of course, he really is a waiter in a café; he's not a fake, he's not secretly a CIA spy. But nevertheless, he is playing a role. He is trying to fit into that “waiter”-role exactly.

You may well wonder, “What is there of Bad Faith in this?” Well, let's see if we can understand why the fellow is doing this.

Sartre analyzes the situation as follows: As a matter of fact — as a matter of facticity — the man is a waiter. He's not engaging in self-deception about that. (On facticity, see the discussion in my notes to “Existentialism Is A Humanism” in the course packet. That will do for starters. We will have much more to say about facticity as we go along.)

But of course being a waiter is not the end of the story about this man. It doesn't give him any kind of definition once and for all, as if he were a waiter and nothing else.

On the contrary, he is a waiter who is free. Every morning he freely gets up early and sweeps out the café, he freely starts the coffee. At any time, he could stop doing that. He could just decide to stay in bed some morning. He might get fired, of course, but he is free to get fired. He could quit, he could burn the café to the ground. He could run off and join the Foreign Legion.

In other words, being a waiter doesn't offer this fellow any kind of ultimate security. He still has to make ultimate decisions about what to do with his life. His being a waiter simply provides the context in which he exercises these free choices — they provide the starting point for his free decisions.

In short, the waiter is a waiter, to be sure, but that is not all that can be said about him. He is a waiter with choices. He is not defined by that. He goes beyond being “just” a waiter.

And in this sense of “going beyond,” Sartre says that the man “transcends” his being a waiter.

And so there are two sides to this waiter: his facticity and his transcendence. In other words, his context and his freedom that will be exercised in that context.

And of course, the same is true of all of us: We are all a kind of combination of our facticity and our transcendence. That is what it is to be a human being.

But of course freedom is a scary thing, something we don’t like to think about and take responsibility for. It would be much nicer if we didn’t have to make choices, if we had no risks, and just fit into our cozy little corner of life forevermore.

Now the waiter, by playing so hard at being a waiter, is in effect trying to deny his freedom, to play down his transcendence and to play up his facticity. He’s acting as if he were nothing but a waiter. He’s trying to be the perfect waiter. Or, to put it another way, he is trying to be a waiter in the sense that this would give him a definition, an essence.

He is trying to be what he is (a waiter) in the sense of being exactly that and no more (no possibilities, no further options).

In short, he’s trying to turn himself into a being-in-itself.

Why do that? Because of course, he wouldn’t have to face the risks and anguish of freedom that way. He would be secure. He would know exactly what was expected of him, what he could do and what he could not do. He could say, “Look at me. I’m just a waiter, and that’s that. That’s where I fit in.”

On the other hand, he doesn’t want to be a being-in-itself in the sense of being unconscious. He still wants to be conscious — if for no other reason than to enjoy the benefits and security of being “just” a waiter.

And so, the long and the short of it is that the man is trying to have it both ways. He wants to be both a being-in-itself and a being-for-itself at the same time.

And where have we seen this notion before, the notion of a combination of being-in-itself and being-for-itself? In the traditional notion of God.

In short, the waiter is trying to turn himself into God. Not the God of the Bible, necessarily, not the God who parted the Red Sea, but a very special and very personal kind of God. He is trying to become a Waiter God.

The waiter is just a type and symbol for all of us. We are all trying to be God — each in our own way. We all want the security of being-in-itself.

But of course we will never reach it, since God is impossible. We are doomed to frustration. (“Man is a futile passion,” as Sartre says at the end of the book.)

So although Sartre is an atheist, the figure of God is absolutely central to his philosophy. Our whole lives, everything we do, is so to speak aimed at this “absent God.”

This is the real and profound reason behind the traditional notion of God as the ultimate good.

Sartre goes on to discuss *other* ways in which we might try to deny the delicate balance between facticity and transcendence that is what we are. In the example of the waiter, the man was trying to emphasize his facticity at the expense of his transcendence (freedom). Why would he do that? Because freedom is scary stuff — it results in anguish.

And of course we not only like to think of ourselves this way, we also like to treat others the same way. You wheel your grocery cart up to the cash register, and as the cashier is ringing up your purchases, all of a sudden she stops, looks up, and asks you about your political preferences, or your views on abortion. What is your reaction? Well, unless you’re one of these know-it-alls who like to preach your views to everyone around, you’ll probably tense up. Whatever you say out loud in that case, you’ll probably be thinking, “That’s none of your business. Just shut and ring up the groceries.” In other words, just do your job, just be nothing but a cashier. Why? Because then I know how to deal with you, I know what to expect and what I am supposed to do in response. That’s secure and reassuring. But once the cashier begins to act in unpredictable and erratic ways, that cozy and familiar situation is shattered. And that’s hard to deal with.

All this involves denying freedom and emphasizing facticity. But sometimes we might do it the other way around. We might find that certain facts about ourselves (something we did in the past, for example) are so unpleasant and disagreeable that we are more comfortable facing the horrors of freedom than we are facing up to this horrible fact about ourselves in the past. In this case, we might try to deny our facticity and emphasize our transcendence. We might say, “Oh yes, I did that. But that was along time ago. I’m beyond that now.” (For example, criminals might say this.)

But of course, our facticity and our freedom are not things we can really succeed in fooling ourselves about. Every act of consciousness we make is a free act in a context. And since every act of consciousness is non-positionally aware of what it is doing, it follows that we cannot ultimately escape being conscious of our freedom and our facticity. Nevertheless, we can distract ourselves from these things (as the waiter is trying to do) — and that’s Bad Faith. We are fooling ourselves about something we know is otherwise.

Belief

Note: This section of the notes may be a little ragged, since I revised it in light of our discussion in class, but haven’t had a chance to go back and smooth it out yet.

We are now ready for § 3 of the Chapter, “The Faith of Bad Faith” (beginning on p. 112). And here things get subtle.

The odd thing about Bad Faith or self-deception is that it works — in a funny way. We can succeed in feeling more secure this way. We can end up reassuring ourselves in a way we are nevertheless fully consciousness is wrong! How is this possible?

OK, now pay attention. So far, we have been putting self-deception in terms of knowledge. And in fact that’s the way Sartre himself introduced the topic at the beginning of the Chapter: the Deceiver knows the truth, whereas the Deceived doesn’t know the truth. It’s all in terms of knowledge. And the idea, so far, is that in self-deception, where the Deceiver = the Deceived, this yields a contradiction.

BUT now consider Sartre’s discussion of Stekel’s case of the frigid woman.

Positionally, she doesn’t know she’s feeling pleasure. If you ask her, she’ll honestly (I don’t say “truthfully”) say no. So positionally, she’s deceived. So far, so good.

Nevertheless, we said, she does feel pleasure, and pleasure — like any act of consciousness — is always non-positionally aware of itself. Non-positionally, therefore, she is quite aware of the pleasure. So she deceiving activity (the “Deceiver”) must be located at the non-positional level.

Now, of course we know that positional consciousness and non-positional consciousness are not two things; they are simply two different facts about one thing. So we have the identity involved in self-deception.

Nevertheless — and this is a point that will become increasingly important as we go along — non-positional consciousness is not knowledge. (As we shall see, “knowledge” is confined to positional consciousness, to the relation of intentionality. Consciousness “knows” its objects. All this emphasis on non-positional consciousness is part of what Sartre meant in the Introduction when he said “We must abandon the primacy of knowledge.”)

Therefore, since non-positional consciousness is not knowledge, there doesn’t seem to be any sense in which the woman can be said to know she feels pleasure. (Or at least nothing we’ve said so far gives us any reason to think there is.)

But, if all this is so (and it is), then the contradiction seems to disappear after all. She does not both know and now know she feels pleasure. What we have instead is merely that she is aware she feels pleasure (non-positionally aware) but doesn’t know it.

Furthermore, apart from doing away with the contradiction, this situation doesn’t even seem to be what Sartre means by self-deception or bad faith. After all, no act of consciousness knows (is positionally aware) what it is aware of non-positionally — since intentionality is irreflexive, recall.

What has happened here?

As yet a further indication that something is wrong, consider the example of the waiter again.

The waiter is non-positionally aware that he is free. Nevertheless, Sartre says he is in bad faith insofar as he is trying to deny his freedom.

But what is the result of this bad faith? Does the waiter not only fail to know that he is free, but positively believe he is not free — as the woman positively believed she was not feeling pleasure? In other words, has the deception worked?

Surely not. Sartre doesn't discuss the point explicitly, but it can hardly be so. For if the waiter did believe he is not free, he would be reflecting. And that doesn't seem to be what is going on at all. The waiter isn't doing his job while reflecting all the while to himself, "I'm just a waiter." What he is doing is role-playing, which is not typically in this kind of case a reflective enterprise. (What he might be doing instead is saying to himself something like, "Waiters must do this, waiters cannot do that, there's no question of doing that other thing.")

What all this means is that we are still not clear exactly what bad faith is for Sartre.

In order to try to improve our understanding, let's turn to the discussion of belief — beginning on p. 112.

In this connection, Sartre discusses this notion of belief beginning on p. 114.

We must distinguish two senses of 'belief'. In one sense, belief is compatible with — and in fact implied by — knowledge. (I can't know something if I don't even believe it.) Knowledge implies belief (and more besides). This is not the sense Sartre is talking about here. ('Knowledge' in this context is not just knowledge in the sense of positional consciousness, as described above, but in the stronger epistemological sense in which we say, for instance, that knowledge is "justified true belief." This is positional consciousness, yes, but it narrower than that.)

What he is talking about might be called "mere" belief. (Does my friend Pierre like me? I don't know, I believe so.) In this sense, belief falls short of knowledge.

On p. 112, he defines belief in this sense:

But if we take belief as meaning the adherence of being to its object when the object is not given or is given indistinctly, then bad faith is belief.

Let's translate that. First of all, when he says "the adherence of being to its object," what kind of being does he have in mind? Well, what kind of being has objects? Not being-in-itself; it doesn't take an object, it just sits there inert. The only kind of being that can be said to have an object is being-for-itself. (And to say it has an object is just another way of affirming the notion of intentionality. So, once again, we are talking about positional consciousness so far.)

To "adhere" to the object in this case means something like committing yourself to it. So what the whole definition amounts to is: consciousness's committing itself to something

for which it has at best inadequate evidence. (If we had adequate evidence, if the “object” were “self-given,” as Husserl put it, then we wouldn’t be talking about mere belief any more, but about full-blown knowledge.)

The point here is about where the energy behind that “adherence” comes from. If what I am committing myself to is so perfectly obvious that in a sense I cannot help but assent (think of the “flash of insight” that sometimes comes when you’re trying to do a logic proof and suddenly “see how it goes”), then there’s nothing for me to do about it any more. It’s just a matter of opening my mental eyes and noticing the evidence.

But if what I’m committing myself to is not so obvious, and I nevertheless commit myself to it firmly anyway, then the “energy” does have to come from me. It involves an effort to believe like that.

Take another example. A few years ago we sometimes saw on television these pathetic parents of soldiers missing in action in Vietnam. Often they just knew their sons were alive and well somewhere over there in Indochina. How did they know that? Well, they just knew. They had faith, or however they put it. Did they have access to some secret State Department information that the general public did not have? Well, no, but they just knew anyway. (Note: If they really “just knew,” then wouldn’t be talking themselves into it; they wouldn’t have to.)

Now of course they didn’t know; they didn’t have any more information than the rest of us (in most cases). But they simply refused to acknowledge the possibility — even the probability — that their sons were dead. It’s easy to understand what was motivating such a denial. But it is harder to see how it would succeed. After all, when you tried to make them face the possibility that they could be wrong, they would just reaffirm their “knowledge” all the louder in a kind of stubborn incantation.

In other words, they went to a lot of trouble to hold this belief. And this lot of trouble was not something they were doing unconsciously. They were doing it out loud. It was so much trouble, in fact, that the average viewer’s reaction was often, “You’re protesting too much. Who are you trying to convince, me or yourself? If you really were as confident as you say you are, you wouldn’t have to insist so loudly.” This is exactly the kind of situation Sartre is talking about.

How would Freud handle this situation? He couldn’t — the situation is exactly the same as for the frigid woman. There may be a hidden motive for the parent’s denial, and perhaps Freud could explain that (although actually, the motive is not really very “hidden” at all). But that’s not all they’re deceived about, and the rest of the deception is something that’s totally beyond his theory.

Note that all self-deception involves belief in this sense. It always involves committing yourself to some view or claim. And since self-deception is not knowledge, the evidence for the claim has to be less than adequate.

Now you can't believe x, in this sense, without being aware (non-positionally) of believing x — because every act of consciousness is positionally conscious of an object and non- positionally conscious of itself.

Thus, you can't believe x, in this sense, without being aware that you are “adhering” to something for which you have at best inadequate evidence. You must be aware that you are making yourself believe it, despite the lack of evidence.

And of course, the more you are aware that you are making yourself believe, that you have to make yourself believe if you are going to believe at all, that your evidence is insufficient, the more you are aware that you could be wrong! The more you try to exclude that possibility, the harder you have to work, and the more you are aware of having to do that work — because of the inadequacy of the evidence. The more firmly you “hang on” (“adhere,” in Sartre's word) to whatever it is you're trying to convince yourself of, the more obvious it is that you're really “unstuck.”

In short, your belief is undercut. The more you believe, the less you believe. As he says (p. 115), “Every belief is a belief that falls short; one never wholly believes what one believes.” And (p. 114), “To believe is not-to-believe.”

And that's the contradiction in self-deception!

(The business at the beginning of the chapter about knowing and not-knowing the truth was then just a first approximation, a way into the topic.)

In short, this kind of belief Sartre thinks is self-deception. It is a delicate balancing-act that keeps threatening to disintegrate. Self-deception succeeds in a sense, but never completely succeeds at what it is trying to do.

It can succeed in making us feel better, in a way. But of course it is trying to believe without having to make itself believe. Self-deception, in other words, is never trying to be self-deception; it is trying to be knowledge. And it can never succeed at that.

It is what Sartre calls a *metastable* notion. It is unstable, constantly threatening to fly apart, and yet can be sustained — like bad faith — for long periods.

It is worth dwelling for a while on the kind of thing Sartre is talking about here. First of all, for those of you who know something about Søren Kierkegaard, you should recognize in Sartre's definition of belief what Kierkegaard calls truth as subjectivity. Kierkegaard defined it as “an objective uncertainty held fast in an appropriation process of the most passionate inwardness.” It's exactly the same notion. Kierkegaard thought it was something great. Sartre, by subsuming it under the notion of “bad faith,” regards it as a kind of existential vice.

The reason Kierkegaard thought it was something great is simply that when I believe (“hold fast”) in this sense, all the effort comes from me. If I were dealing with something that was so obvious and plain that there was simply no question about it, then I wouldn't have to bring myself to believe it. In that case, the blinding evidence overwhelms me and I cannot help but assent to it. It's the object and its evidence that prompts my assent. But

where what we are talking about is not overwhelmingly obvious, I have to work myself up into believing it. There I am the one doing it.

Kierkegaard thought this was something great because he was interested in the believer's own contribution and efforts. There's no glory in believing something, after all, unless the believer has some active role in acquiring that belief. Sartre is in full agreement about the mechanism here, even though his overall evaluation of what is going on is quite different.

For Sartre in the present Chapter, self-deception turns out to be inevitable — if not all kinds of self-deception, at least that kind of self-deception where we are deceived about ourselves, where we are reflecting, as in Sartre's example of the frigid woman. We cannot avoid this kind of self-deception. Look what happens if we try to avoid it — if we try to be sincere. (See the discussion on pp. 105ff.)

What are we trying to do when we try to be sincere — not so much to be sincere about whether we feel pleasure or not, or whether I am free to quit my job as a waiter, but sincere as a general policy? We are trying to see ourselves for what we really are, to see ourselves objectively and fearlessly. We are trying to face the fact that we are what we are, after all, and not to deny anything.

Note: We are trying to face the fact that we are what we are? That phrase should be a warning sign to us. For Sartre we are not what we are. The very attempt to look at ourselves “objectively” is false from the outset. There is nothing “objective” about us. It is just another attempt to pretend that we have a kind of “definition,” to get straight on “the real me,” so that it will turn out that there is some definite and settled truth about ourselves after all, and it is just up to us to face it fearlessly. But there isn't any such settled definition. The very attempt to be sincere in general is just another, particularly insidious way of being in Bad Faith all over again.

Thus, the goal of trying to be sincere is an impossible goal. It would require us to be an in-itself-for-itself. It is just another way of trying to be God.

(Note: This doesn't mean the attempt is impossible. It isn't; we attempt it all the time. But the attempt cannot succeed.)

In effect, this pessimistic conclusion comes from Sartre's analysis of the way consciousness works — an analysis that is in the broad phenomenological tradition stemming from Husserl. (Not that the details come from Husserl, of course.)

In this connection, it is very interesting to look at the footnote at the end of the chapter on “Bad Faith” (p. 116). Sartre has just been talking about how we cannot avoid Bad Faith or self-deception. But now he says in the note (emphasis added):

If it is indifferent whether one is in good faith or in bad faith, because bad faith reapprehends good faith and slides to the very origin of the project of good faith [translation: the attempt to avoid self-deception and be sincere is just another form of Bad Faith], that does not mean that we can not radically escape bad faith.

Oh, it doesn't, does it? I thought that was just what you were saying! How is this going to work? Well, here — at this crucial point — Sartre resorts to a metaphor. He goes on (emphasis added):

But that supposes a *self-recovery of being which was previously corrupted*. [Oh, that's what it is, is it?] This self-recovery we shall call authenticity, *the description of which has no place here*.

Or anywhere else in Sartre. He never really completely tells us how this “authenticity” is going to be possible, given the analysis he has just given. And the fact is, given that analysis, it doesn't look as if it is going to be possible — not if authenticity means avoiding bad faith.

Sartre describes authenticity in many places — in his plays and novels, for instance. (Perhaps most vividly in his play *The Flies*.) But he never really tells us how it is possible to be authentic, how it is possible to “get there from here.”

Sartre's philosophy shows the influence of two distinct traditions, one stemming from Kierkegaard and Nietzsche. This tradition puts a high value on the notion of authenticity. This tradition shows up in Sartre when he is talking about values and ethics. He wants the notion of authenticity in his philosophy. He is aiming to find room for it.

But he is also influenced by another tradition, the phenomenological tradition stemming from Husserl. This tradition shows up when Sartre is talking about metaphysics and epistemology.

The problem is: The two traditions don't mix very well. The result is a very volatile brew.

In this footnote, we see proof positive of the collision of these two traditions.

The footnote is very revealing, and I think it is right to view it as a kind of embarrassment for Sartre. Nevertheless, I don't mean to suggest there is ultimately no way out. Much of the recent secondary literature on Sartre's ethics is devoted to looking at how in fact Sartre (mostly in his later writings) does try to get himself out of this apparently insoluble knot. (See for example Anderson's and Detmer's books, in the list of *Books on Reserve*, in the course packet.) These recent studies, I think, are very promising, and help to correct a lot of misunderstandings about Sartre. They point out a lot of material, particularly in the later writings, where Sartre does talk about what it would take to be authentic. But Sartre never really says as much as we would like on this point. We will talk more about this later in the course.

But let me say *a little* more about even now. It seems to me that at this stage of his thinking, Sartre perhaps doesn't really know how he's going to be able to reconcile all the things he wants to say.

Certainly in this chapter, at any rate, he talks as if at least the kind of self-deception that involves beliefs about ourselves really is inescapable. (Perhaps it is possible, even in this

chapter, to avoid other kinds of self-deception — for instance about whether your son who is missing in action in Vietnam is still alive. That’s another question.)

And in a sense, it is not hard to see why self-deception about ourselves is going to look inevitable, at least, given that we reflect on ourselves at all. For, as we’ve already discussed in our treatment of *Transcendence of the Ego*, REFLECTION ALWAYS DISTORTS. At the time, I gave you the reason: remember the metaphor about making up little wax statues.

Now insofar as self-deception always involves believing something, self-deception about ourselves will always involve a kind of reflective belief, which then will always be wrong (since it’s distorted) and so deceptive (since we believe it).

How then can we ever avoid bad faith or self-deception about ourselves? How is authenticity possible? Well, you might suggest that one way is simply never to reflect. And perhaps that will work for this kind of bad faith, but even if it does, it certainly isn’t the kind of “escape” from bad faith that Sartre calls authenticity. The “authentic” individual is surely not someone who just never bothers to stop and reflect on what he or she is doing.

But there is another possibility too. When we were talking about the reasons why reflection always distorts, we said that this posed a serious methodological difficulty for Sartre’s whole enterprise, which is largely a reflective one. Sartre recognized this difficulty from the very beginning, we said, and had at least a name for what was going to allow him to get around it. It was called pure reflection. We speculated a little on what that might turn out to be, but we didn’t come to any definite conclusions.

Here we see then that the notion of pure reflection is tied up with the notion of authenticity. If pure reflection offers us a way of reflecting without distorting, then perhaps it will also offer a way to avoid at least one kind of bad faith. Pure reflection and authenticity are probably not exactly the same thing, just as it is no doubt too simple just to identity what Sartre calls “authenticity” with simply the avoiding of bad faith. But all these things are connected: escaping bad faith, pure reflection, authenticity.

If you are interested in pursuing this topic further, you should look carefully at Anderson’s book (who gives a kind of description of what authenticity really comes to and how it is possible), and Detmer’s book (who basically agrees with Anderson but provides some additional information). I think their picture of what authenticity is is not right, but is on the right track. And I also would urge you to look closely at Christopher Vaughan’s dissertation on the notion of pure reflection and how it is tied up with authenticity.

Let me add some last undigested thoughts on this topic of bad faith.

So far, it looks perhaps as if we’ve talked ourselves out of the contradiction the chapter was trying to convince us of. If the main purpose of the chapter was to convince us that consciousness is what it is not and is not what it is, that it is out and out contradictory, and

to do this by trotting out self-deception as the prime piece of evidence, the result is not very persuasive.

We've seen that the contradiction of knowing and not-knowing the same thing at the same time is not really involved here. Instead, the "contradiction" arose in the claim "to believe is not-to-believe." But, on closer examination, this really isn't contradictory either.

What we seem to have in the case of mere "belief" is not so much something that is contradictory as something that is self-defeating. And that's not the same thing at all. It's not really that "the more I believe, the less I believe." It's rather that the more I try to believe (to "adhere"), the less I succeed, the more I come "unstuck." That's not a contradiction; it's just failure.

In other words, what we have so far is a little like turning the steering wheel in a skid. If you turn it the wrong way in an effort to get out of the skid, you only skid all the more. And the more you turn it the wrong way, the more you just make matters worse. There's nothing contradictory about that.

If this were the end of the story, then I think we could all go home. Sartre would have failed to make his point. But it isn't the end of the story. There is one other factor that we haven't yet accommodated, and that isn't brought out in the example of skidding. And that is: self-deception works. It doesn't succeed at being knowledge, as we said. And it doesn't even succeed at being belief. But it does succeed at something! We can make ourselves feel better by engaging in bad faith. We can succeed in distracting ourselves from what we are fully aware of all the while.

And this is the feature that, I think, would give us the full-fledged contradiction after all, if we could only get a good grip on what is involved. But we haven't got it yet.

The Emotions

I want to look now at Sartre's discussion of the emotions, in his early book *The Emotions: Outline of a Theory*. (See the outline contained in the course packet.)

In connection with this, you should also know about an excellent book by Joseph Fell, *Emotion in the Thought of Sartre*. (See the list of books on reserve, the course packet.) This book is about Sartre's theory of the emotions, but it is also about lots of other things we've been talking about. It is an excellent book.

Sartre's *The Emotions* appeared in 1939, before *Being and Nothingness*, which appeared in 1943. In it, he discusses some of the psychological theories current in France during the early part of the century. Some of these theories, and some of the names, will very likely be unfamiliar to you. But that doesn't matter, since similar theories are still around today, and some of Sartre's comments apply just as well to present-day theories.

It is appropriate to look at the emotions, because they at first appear to present a difficult case for Sartre's theory that human beings are totally free and totally responsible for their actions.

It seems that, at least in some cases, we are not responsible for our emotions. Sometimes we are overwhelmed by our emotions, we lose control, we break down. Popular opinion (if not always the law) recognizes an important distinction between so called crimes of passion and cold-blooded, pre-meditated crimes, and regards the latter as involving somehow a greater guilt — implying that the former involves a lesser guilt, a lesser responsibility.

Even the term 'passion' here indicates a passivity, as though the emotion is not something we do, but something that happens to us.

Not so, for Sartre. For him, we adopt our emotions, we take them on. And thus we are completely and totally responsible for them. This is what he is going to try to show in this book.

He does this by first considering alternative theories. Let us look at what he says.

We begin with the "Introduction." Just as in *Being and Nothingness*, the "Introduction" is the hardest part of the text. In the present book, there are two main things you should get out of it:

- (1) The distinction between the two main kinds of theories of emotions.

There are two sides to emotions, Sartre observes: the conscious side (what it "feel like"), and the physiological side (sweaty palms, racing heart, rapid breathing, tears, etc.). The two kinds of theories differ on which of these two sides they take to be primary.

- (a) The so called intellectual theories hold that the inner state of consciousness determines the physiological disturbances. As Sartre pithily puts it: We weep because we are sad.
- (b) The so called periphereic theories. (The word may be found in the Oxford English *Dictionary*, but the discussion there isn't much help. It seems to have something to do with 'periphery'. And they put the accent on the penult: periPHERic. It should go on the antepenult: peRIpheric just as it does in 'periphery' and 'peripheral' — the second 'e' is short, and there are rules about these things in Greek, after all.) According to these theories, the physiological disturbances determine the state of

consciousness involved in an emotion. Thus: We are sad because we weep.

There are various kinds of peripheric views, some of them pretty crude and unconvincing. But note: Any kind of modern behavioristic view — whether crude or sophisticated — will fall into this category, any view that tries to reduce emotions (or any other psychological state, for that matter) to observable, laboratory testable phenomena — physiological, behavioral, etc.

Sartre's own theory of emotions will turn out to be of kind (a), an intellectual theory (although, despite the term, the intellect will not have any special role in Sartre's theory).

- (2) The second thing to get out of Sartre's "Introduction" is the notion of the signification or meaning of an emotion.

Sartre thinks emotions are like the three faces of the cube in perception (recall the passage from *The Psychology of Imagination*). They promise more than they directly show us. Emotions are not just brute facts of physiology or behavior. (For that matter, they're not just brute facts of consciousness either.) On the contrary, they have a human meaning that we can learn how to read if we are careful — just as the three surfaces we see in perception mean (= imply, promise) a whole cube.

This is the basic notion of signification or meaning, as it occurs throughout this book. We shall have to see how it works in detail.

After the "Introduction," Sartre looks at various alternative theories, and in discussing them, he gradually works up to his own. Throughout the discussion of these alternative theories, he adopts a kind of quasi-Hegelian dialectical procedure. That is, first he states a position. Then he raises objections to it. The reply to these objections leads to a new position, to which he raises new objections, and so on.

In Ch. 1 ("The Classical Theories"), we get a discussion of various peripheric views. In Ch. 2 ("The Psychoanalytic Theory"), we get — as the name implies — a discussion of the Freudian theory of emotions. Finally, in Ch. 3 ("A Sketch of a Phenomenological Theory") we get Sartre's own view.

That's the bare bones of the structure. Let's now sketch it in a little more detail, and then we'll look at it in much more detail.

I. Peripheric Theories.

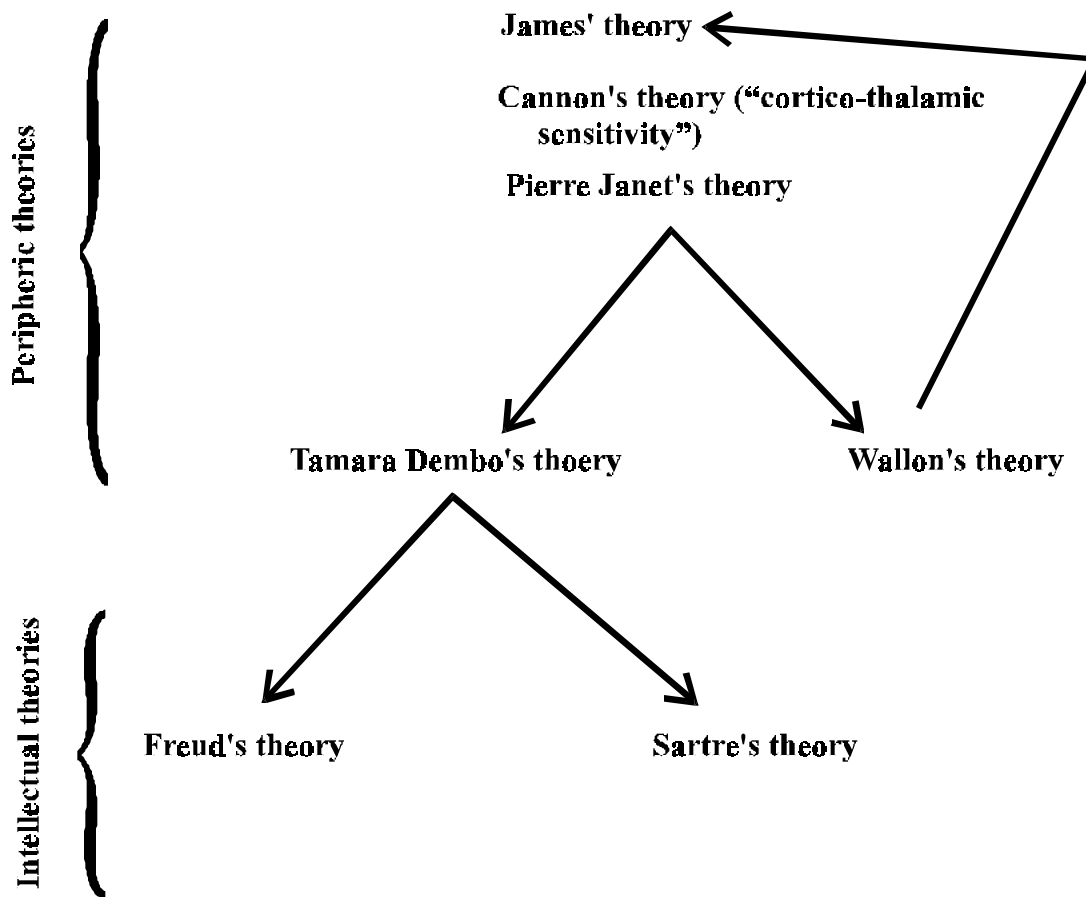
- A. First, Sartre considers peripheric theories in general, and raises some general objections against all of them as a group. For purposes of illustration, he takes William James' theory as a typical starting-point.

- B. Then Sartre considers a certain Walter B. Cannon's so called "cortico- thalamic sensitivity" theory as an attempt to answer one of the objections raised above. But other objections remain.
- C. In an attempt to answer these, we move on to Pierre Janet's theory of emotions. Sartre thinks this theory is in the end ambiguous. He thinks Janet was on to something important, but his behaviorist biases kept him from pushing it in the right direction.
 - 1. If you push the ambiguity in one direction (the behaviorist direction), you will end up back with James' theory. Sartre considers one attempt to prevent this (Wallon's theory). But that attempt fails.
 - 2. The failure of Wallon's theory indicates that we must go the other direction. Sartre considers an attempt to come to terms with what is involved. (This is Tamara Dembo's theory.) But this theory still hasn't quite got it. At this point, Sartre regards the peripheric alternative as exhausted. This leaves us only:

II. The Intellectual Theories.

- A. Among the intellectual theories, Sartre considers Freud's theory first, and argues against it.
- B. Finally, Sartre gives us his own theory, and declares it the winner.

Thus:



Now let's look at Sartre's discussion in some detail.

Sartre begins by giving some *basic* objections to the original peripheral theory (that "we are sad because we weep"). He lists *three* such objections at the outset (see the outline). But we only need to be concerned with two:

- (1) What about the "subtle" emotions? That is, what about the rather mild emotions that don't have any obvious physiological component? Emotion is not *always* a matter of racing pulse, flushed face, etc. What about things like "*passive sadness*"? (See the text on this.)

As it stands, this is a crude objection, and can be easily answered. ("You've just not looked hard enough for the physiological component! Nobody said the physiological component was a matter of *gross* physiology.") Sartre in fact is using this objection only to get things going.

- (2) The second objection is more serious: The physiological phenomena associated with *joy*, for example — rapid heart beat,

rapid movements — differ only in intensity (that is, quantitatively) from the physiological phenomena associated with anger.

The point is this: The physiological phenomena associated with various emotions tend to be pretty much alike, and differ only in degree. Yet anger is not just more intense (or less intense) joy!

It is perhaps hard at first to see the force of this objection. But in fact, it is a pretty good one. The basic complaint is that peripheric theories in general must try to explain qualitative differences among the emotions in terms of quantitative differences among the physiological phenomena associated with them. And, Sartre thinks, that just won't work.

This is one of the big themes of the book. Emotions are qualitatively different. They are different organizations, systems of behavior. And Sartre thinks that in the end this can only be explained in terms of the human meaning or significance of emotions; they cannot be reduced to simply a matter of chemistry or biology.

In order to make the case, let's look — as Sartre does — at some actual examples of peripheric theories. And, with Sartre, let's begin with William James' theory.

According to James' theory, the emotional state of consciousness is just consciousness of the correlated physiological disturbance.

For example, what is sadness? For James, it is just our consciousness of our weeping — or, in less extreme cases, of our depressed heart beat, lowered blood pressure, etc.

So too with anger: It is the consciousness of our clenched fist, grinding teeth, adrenaline, racing pulse, etc.

The two basic objections we listed above apply here: (1) What about the subtle emotions, where there don't seem to be any appropriate physiological phenomena to be conscious of? And (2), there is also the problem of reducing qualitative differences to quantitative differences. (Anger doesn't require that we clench our fists, after all. In many cases, the physiological side of anger looks just like the physiological side of sadness.)

In an attempt to answer the first of these objections, we turn next to the cortico-thalamic theory associated with Walter B. Cannon. This theory in effect says: "You're just looking in the wrong place for the physiological disturbances." In the case of the "subtle" emotions, there still are such physiological correlates. But they are not obvious ones, like clenched fists. Instead they are buried deep in the brain cortex.

Sartre's response to this theory is twofold:

- (1) First, the theory's claim is unverifiable. In effect, the theory is just saying that there must be such brain events in there, and they must be of the kind the theory requires, because otherwise the theory wouldn't work. For Sartre, that kind of pleading is a mark of desperation. And he's right. But you may well ask yourself whether

the theory really has to resort to such pleading. Is such brain activity really all that unverifiable?

- (2) But it doesn't matter, because even if you could verify such brain events, there is still the second line of objection. The theory is still committed to trying to explain qualitative by quantitative differences.

We turn next to Pierre Janet's theory. According to Janet, the problem with James' account is that he didn't take account of what Janet calls the "psychic" element in emotions. Sartre applauds this observation, and thinks Janet is basically right about that.

Janet goes on. He is still pretty much a behaviorist at heart, and wants to keep everything on the physical, empirical, measurable (and therefore quantitative) level. So what Janet does is to distinguish two levels within the physical, empirical realm:

- (1) The purely biological, physiological level — heart beat, adrenaline, etc. This seems to be what James had in mind. In effect, this level is reducible to a matter of biochemistry.
- (2) Organized activity, or what Janet calls "behavior."

The point of level (2) is that certain kinds of behavior are associated with each emotion. Each emotion presents an organized structure of its own. And, for Janet, this is how you can differentiate emotions qualitatively, in terms of their structure.

In effect, this is Janet's answer to the second line of criticism raised above against peripheric theories in general. Sartre thinks Janet is absolutely right so far. (In general, Sartre appears to have had a great respect for Janet.) But then, he thinks, Janet loses the thread.

Janet goes on: Emotional behavior is a kind of disadaptive behavior, resulting from a "setback."

The idea is this. You find yourself in a situation that calls for a certain kind of action, but that action is impossible, or at least very difficult. Thus, there is a tension in the situation. As a result of this tension, things just finally break down. The organized behavior that is well adapted to the demands of the situation just disintegrates, and is replaced by a disadaptive behavior, one that is not appropriate.

For example, certain patients came to Janet (who apparently had a clinical practice) for psychological help. As they discuss what it is that is bothering them, the level of tension in the situation rises. And as they get closer and closer to the real nub of the problem, the tension increases until finally they break down in sobs and can't continue. That is, whatever it is that is really bothering them is so painful to discuss that they cannot continue the "adaptive" behavior (which would involve getting to the bottom of it, talking it out, working it through) — and so it "breaks down." Even our colloquial speech captures what is going on here: a "breakdown." The sobs and tears are not at all well

adapted to the situation — in fact, they get in the way of working out the real problem. And in this sense, they constitute what Janet calls a “setback behavior,” a disadaptive way of dealing with the situation.

Sartre thinks Janet was definitely on the right track here, but on his own principles he has thrown the whole thing away. In the end, he is trying to explain all this on a purely physical level, thereby ignoring the excellent distinction he himself had made between the physical level and organized activity or behavior.

Here is Sartre’s critique of Janet: The “setback behavior,” according to Janet, is disadaptive, disorganized. It is the breakdown of the organized, adaptive behavior called for in the situation. But, Sartre observes, if what really happened is a breakdown of organization, then it was not a matter of “setback behavior,” but rather a lack of behavior — that is, a lack of organized activity.

Note: It was just the organization that Janet had said was what distinguished the purely physiological phenomena that James appealed to from real behavior. If the breakdown is the breakdown of organized behavior, then it appears we are in effect back with James’ theory. (You may want to ask yourself whether this really follows.)

What happened to the organized structure of the emotion itself, the organized structure that Janet had promised us? It is in this sense that Sartre thinks Janet has lost the thread.

At this point, Wallon comes to the rescue, touching up Janet’s doctrine. For Wallon, the breakdown does not result in activity that is totally lacking in organization. It’s just not organized in the same way.

For Wallon, the infant’s primitive nervous system has a built-in, inherited system of behavior. What Janet calls “setback” behavior is a case of abandoning the highly organized, appropriate behavior and reverting to the primitive behavior of the infant. We kick and scream — we quite literally become “infantile.” But the point is that this is still organized behavior, not just raw physiology, as on James’ theory.

Sartre thinks this just doesn’t help. First of all, Sartre remarks, James himself would be perfectly willing to accept this notion of the infant’s primitive level of behavior, if he saw any evidence for it. And so, once again, we are back with James’ theory.

I think this criticism is probably unfair. Sartre makes it sound as though Wallon’s theory reduces to the crude theory we found in James. But really what he has said is that James’ theory need not be so crude as we originally thought, and might in fact be as sophisticated as Wallon’s.

But, in any event, Sartre has a more serious line of objection. Why did Janet’s patients break down into sobs, instead of getting angry or going into a silent pout? The problem is that, on Wallon’s theory, there is only one kind of emotional behavior, “infantile” behavior. But in fact there are several different kinds of emotional behavior, with quite distinct structures. Wallon’s theory simply cannot account for this.

Sartre has another objection to Janet's setback theory: What about the very notion of a setback? Despite Janet's behaviorist convictions, this notion of a "setback" requires the notion of "finality" or "goal-directedness." And that is not something that can be explained on a purely physical level, but requires consciousness or some other goal-directed agency. So Janet's own theoretical principles are getting in his way.

Let's look at this notion of a "setback." What we have on Janet's theory is simply a case where one form of behavior can no longer be maintained, and so is replaced either by disorganized activity (Janet's own theory), or by another kind of behavior (Wallon's theory).

Far from being a "setback," this would seem quite the appropriate thing to do. The original behavior cannot continue to be maintained, after all. In order to think of this as a setback, we need to suppose some kind of agent that is trying to maintain the original form of behavior (a form of goal-directed behavior) but cannot do so. And this only makes sense if we are dealing with some kind of goal-directed agent.

So, there are really two serious criticisms here:

- (1) We need some kind of goal-directedness. But these theories studiously avoid appealing to anything like that.
- (2) We still have to explain the variety of organized emotions. That is, we still have to explain the qualitative differences we find among emotions.

Sartre goes on to consider the theory of Tamara Dembo, a theory that takes account at least of (2), although criticism (1) still remains.

For Dembo, an emotion involves "changing the form or structure of the problem." In her experiments, for example, a subject would be asked to perform a certain task, but then certain rules were imposed to make this difficult or even impossible. Thus, a subject might be asked to stand within a circle painted on the floor, and then reach out to grasp a certain object. But in fact things were set up in such a way that you *couldn't* grasp the object without stepping outside the circle. Here the original form of the situation involves:

- (a) a task to be performed;
- (b) certain rules to be observed in the performance of that task.

The frustration that builds up as a result of this impossible situation leads to the subject's transforming the situation:

- (1) He either gives up and just mopes — that is, regards the task as no longer having to be performed (thereby transforming part (a) of the situation). Or

- (2) He gets mad and breaks the rule. He just steps outside the circle and grabs the object, and that's that. In other words, he regards the rules as no longer having to be observed (thereby transforming part (b) of the situation).

In either case, what we have is a substitution of one form imposed on the situation by another form. There are a great number of forms that can be imposed on situations in this way. And so we have a way of accounting for the great variety of emotional structures, all qualitatively different.

This clever theory is very close to home for Sartre. But he thinks we still have not taken sufficient account of the notion of finality or goal-directedness that is required if this theory is to make sense.

There is another (in fact, a related) problem with Dembo's theory. The fact that the one form of the situation breaks down as a result of the tension is adequately accounted for on this theory. But how are we to explain the fact that a new form or structure is imposed in its place?

Consider an analogy. Recall the ambiguous Gestalt figure we have discussed many times, the figure with either two faces or a vase, depending on how you look at it. Suppose you start off by seeing the two faces in the figure. Now we might very well understand on purely physical and physiological grounds why that "form" might break down, why we can no longer see it in the figure. For instance, perhaps you turn the diagram on its side, or otherwise disturb the form. That much we can understand. But how do we explain the appearance of the other figure, the vase? How does the new form come to be imposed? Sartre thinks this is exactly where the notion of goal-directedness or purpose is required, and there is nothing in Dembo's theory — or any of the other theories we have considered so far — like a notion of goal-directedness or purpose.

At this point, Sartre thinks he has pushed the "classical" (peripheric) theories far enough. We need to find some way of getting goal-directedness or finality in our picture.

The Intellectual Theories

In Ch. 2 of *The Emotions*, Sartre turns to the so called "intellectual" theories. They take explicit account of the notion of goal-directedness or finality that has been bothering Sartre all along.

In Ch. 2, he says there are basically two types of intellectual theories. Either

- (a) Consciousness itself is the goal-directing agent. (This will be Sartre's own view.) Or else
- (b) An unconscious agent does it. (This is the Freudian view.)

(Of course, there are other theories of the unconscious besides Freud's, but Freud's can serve as a kind of paradigm for the others.)

Now we've already seen why Sartre rejects the Freudian notion of an unconscious region of the mind. So we don't need to go into this in great detail here. For present purposes, let's just look at one criticism Sartre raises.

On the Freudian theory, the conscious mental state that is one half of an emotion (the other half, recall was the "physiological disturbance") is significant. It means something, it fulfills a goal or purpose.

Quite so, Sartre agrees. But, on the Freudian view, this significance or meaning is something imposed on it from outside — just as the significance or meaning of a red traffic light at the corner is a meaning imposed on it by social conventions. (By itself, the red light doesn't mean anything at all.)

For Freud, what the conscious state of mind means is determined by the unconscious drive it fulfills, symbolically or otherwise.

On this theory, it is not consciousness that is goal-directed or purposeful in emotions, but the unconscious. It's the unconscious part of the psyche that is striving to satisfy its drive. (Recall our earlier discussion of Freud's "Pleasure Principle.")

The conscious state of mind, on this theory, takes on a certain meaning or signification only because it is caused to do so by the unconscious drives striving for satisfaction. Thus, consciousness is passive in this respect. It is a kind of thing, subject to the rules of cause and effect. (So too, the red traffic light is passive, subject to whatever conventions we care to invent about it.)

Sartre thinks this theory destroys the spontaneity of consciousness. It violates the spirit of the Cartesian *cogito*, which (he says) must be the starting-point.

Well, that last business sounds like sheer dogmatism. Freud disagrees with Descartes, and so we must reject Freud — because we just have to have our Descartes. But in fact, Sartre is on firmer ground than this. We have already seen the basic reason he rejects the Freudian notion of an unconscious: it is impossible, a combination of the in-itself and the for-itself.

In the end, Sartre thinks the only alternative left is to say that consciousness itself is what gives the meaning, the goal, the directedness to the conscious mental state in an emotion. Consciousness gives this meaning to itself. (It's as if the red light decided for itself what it is going to mean.)

What this implies, then, is that consciousness is the origin of our emotions. We consciously (and therefore freely) take on our emotions.

The threat that was the whole occasion for Sartre's writing this book — that the emotions appear to violate the radical kind of freedom Sartre thinks we all have — is thereby answered, and Sartre does not have to compromise his notion of freedom after all.

Before Sartre explains his theory in detail, he turns to two problems the Freudians would raise to this notion that consciousness is what is responsible in emotions:

- (1) If so, then why doesn't it seem so? It certainly looks as though we passively undergo our emotions. In fact, the very name 'passion' is derived from "passivity."
- (2) If so, then why do we so often consciously struggle against our emotions? After all, according to you, Sartre, it is consciousness that is doing it all to begin with? If it doesn't like what it's doing, why doesn't it just stop?

(Note: Sartre never explicitly replies to problem (2). And it's a good question. In several passages in his writings, Sartre compares this kind of situation to dreams, or to cases where we are fascinated — as it were, hypnotized — by certain objects or events. We are, so to speak, drawn into our dreams, and find it hard to wake up. We find it hard to break the spell of what fascinates us. Well, the comparisons are illuminating, but we still don't have a theory of how any of these things is possible if consciousness is doing it all along.)

Sartre's Own Theory

Sartre sets out his answers to these problems (although he doesn't say much about (2)), and explains his own theory, in the last chapter of the book, Ch. 3: "A Sketch of a Phenomenological Theory." Here he makes important use of some of the theory of mind we have developed so far. (Remember that *The Emotions* was written after *Transcendence of the Ego*, but before *Being and Nothingness*.)

The most important distinction at stake here is the distinction between the way in which we are aware of ourselves reflectively and the way in which we are aware of ourselves non-positionally.

Emotions, we know, are states of consciousness. Whatever else may be involved, such as physiological disturbances, etc., at least this much is involved. (In effect, this is what we learned in Ch. 1: The conscious side of emotions cannot be reduced to pure chemistry and physiology, as the peripheric theories try to do.)

Like all states of consciousness, therefore, an emotion is subject to the great law of consciousness: intentionality. Every consciousness is a consciousness of something. Now what are we conscious of in an emotion?

Frequently, it is thought that we are conscious of ourselves in an emotion. When I am angry, I am conscious of my own mental state of anger, etc. (So too, for James, in an emotion I am conscious of my own physiological state.)

Sartre thinks this is quite wrong, and the mistake is responsible for many of the problems in the usual theories of emotions.

What is the object of consciousness in a state of anger, for instance? Sartre recognizes, of course, that we can always make our own anger an object of consciousness — we can always say “I am angry.” And when we do that we are reflecting on our own anger, we are in a reflective mode of consciousness.

But while he says this is always possible, it is not what anger is at first. At first, what I am conscious of is what makes me mad! An emotion is, at least at first, a certain way of apprehending the world.

Of course, what confuses people is the fact that here, as in other cases, positional consciousness of an object is also non-positional self-consciousness. That is, when I apprehend what makes me mad, I am also aware of myself, of my anger. Sartre insists that what we have to remember here is that in such a case I am not aware of myself as of a new object. There is just the one object — namely, what makes me mad.

Recall that this same confusion was behind Husserl’s doctrine of the Transcendental Ego, behind the Self-Love Theory, and behind the Freudian theory of the unconscious.

On Sartre’s theory, an emotional act of consciousness involves an abrupt change of the form of the situation, just as on Dembo’s theory. For example, Pierre and I are engaging in a kind of playful game of “put-down,” in which we take turns insulting one another, each insult more outrageously witty than the preceding one. It’s all in good fun, and not to be taken seriously of course. But then all of a sudden it’s my turn, and I can’t think of an appropriate insult. So what happens? I get mad, and this time for real!

In this case, a situation on which I had formerly imposed the intentional structure “situation-calling-for-a-witty-put-down” is transformed, when I find I can no longer think of the proper reply, into a situation on which I impose the intentional structure “situation-calling-for-serious-abusiveness.” I adopt the emotional consciousness.

Recall that one of the objections the Freudians had raised (see above) was that if consciousness puts on its emotions freely (as Sartre says), and not as a result of causality from the unconscious (as the Freudians would have it), then why aren’t we aware of what we are doing? Sartre’s answer is now clear:

We ARE aware of what we are doing. But we are not aware of it as a kind of object for consciousness, at least not at first. We are aware of it in a non-positional way.

Now you might object here, as you might have objected on previous occasions when Sartre appeals to the notion of non-positional consciousness where others appeal to the notion of an unconsciousness: What has been gained here?

After all, why couldn’t a Freudian just reply that the disagreement is only terminological. You, Sartre, distinguish two sides to each act of consciousness. For you, each act of consciousness is both (a) positional consciousness of an object, and (b) non-positional self-consciousness.

But I (the Freudian) call (b) the “unconscious,” and restrict my usage of the term ‘conscious’ to (a). In other words, to put it in tabular form, we could translate between your theory and mine, as follows:

<u>Sartre</u>	<u>Freud</u>
positional consciousness	consciousness
non-positional consciousness	unconsciousness

In the end, don’t we really have the same theory? It’s just that you have changed the terminology.

Sartre would reply: No, we don’t have the same theory at all. In the first place, you Freudians do not think that everything in the unconscious part of the mind can be brought to consciousness. The mental events that can be brought to consciousness you call “Ego processes” (recall our Freud lesson in the discussion of Bad Faith). According to you Freudians, there are still the unorganized parts of the unconscious Id, and they forever remain hidden from consciousness. (That doesn’t mean we can’t know about them. We can know about them, by inferring what must be going on down there. But they are “hidden” in the sense that we can’t experience them directly.)

If I (Sartre) really had the same theory, as you are now suggesting, then I should have to maintain (according to the above translation-table) that not every act of consciousness can be made an object for itself. I should have to maintain that there are some acts of consciousness on which I could never — on principle — adopt a reflective attitude, that I could never view in an “objective” way. But in fact I do not hold that. On the contrary, I hold that every act of consciousness can be reflected on.

So that is one substantive difference between your Freudian theory and mine. Second (Sartre would go on), on your Freudian theory, the unconscious and the conscious are not just two sides of one and the same thing. For you there is a real distinction between the unconscious and the conscious, to the extent that the former can cause the latter. The unconscious and the conscious are as distinct as cause and effect. On my view, however, non-positional self-consciousness does not cause positional consciousness of the object; it strictly and literally is that positional consciousness of the object.

In other words, on my view (Sartre is still speaking), each act of consciousness is a unitary thing. The distinction between positional and non-positional is only a distinction between two aspects of this unitary thing. It’s like looking at it from the front and looking at it from the back — but it’s the same one thing we are looking at in each case.

But on your Freudian theory, each psychic event is not a unitary thing, but instead a pair of things, linked together by a bond of causality. The difference here is not just a difference of perspective, not just a difference of aspect, but a real difference. You Freudians introduce a split into the mind, and then try to repair the split by patching the two sides back together again with a causal link — a link that conjoins the active,

purposeful unconscious with the passive conscious mind in a way that can only be regarded as magical.

(Note: My impression is that Sartre regards the notion of “causality” *in general* as a magical notion, and so has to reject it. In the end, I take it, this means Sartre doesn’t really believe in the sciences — at least not in the sciences that appeal to causality.)

The upshot of all this is that the Freudian theory is not at all the same as the Sartrean theory, with only a difference of terminology. The differences are more substantive than that.

The Magical World

Sartre thus agrees with Dembo that emotion involves a “transformation of the world,” a change of the form of the problem. But there are many ways of changing the form of a problem, and not all of them are called “emotional.” So we have not yet got to the bottom of this.

We can always transform the world, for instance, by working on it, by acting and laboring in the world to bring about some goal. In this way, we can change a difficult problem into one more easily handled — or perhaps even solved! But there is nothing “emotional” about this. On the contrary, this is a quite reasonable, well-adapted way to approach a problem.

Emotion, on the other hand, involves changing the world by magic. We have already seen this notion of “magic” in Part II of *Transcendence of the Ego*. But Sartre didn’t discuss the notion very extensively there. We discussed it fairly extensively at that time, but we also referred ahead to Sartre’s *The Emotions*. Well, this is exactly the passage. Sartre gives a definition and discussion of this notion of the “magical” on pp. 83–91 (the definition is on p. 84). See our discussion of *Transcendence of the Ego* for the details.

Here is how this notion applies to emotions:

The idea is that there are two basic ways of “being-in-the-world.” That is, there are two basic ways in which the world can “come on” to us phenomenologically. The first way is the deterministic way. When we are in the world in this way, when the world appears to us in this way, the normal rules of means and ends apply. If you want to accomplish a certain goal, you must adopt certain means. If you want to go from here to there, you must pass over the intervening interval. This is a matter of changing the world by working on it, as we just mentioned.

(Note: Sartre’s use of the term ‘determinism’ here should not be taken to be a matter of causal determinism. We just saw him reject the notion of causality, in our discussion of the Freudian theory of emotions. Instead, the kind of “determinism” Sartre has in mind here is more a matter of means and ends.)

But there is a second way of “being-in-the-world” as well: the magical way. When we are in the world in this second way, when it appears to us in the magical mode, all these deterministic connections collapse. The normal rules and laws of nature are suspended. There is no predictability about the “magical” world. It has elements of the spontaneous — that is, of the free, of consciousness. But it is the world — a passive thing — that comes on to me this way. In short, what we have is “an irrational synthesis of spontaneity and passivity,” “a consciousness rendered passive.” These are exactly the definitions Sartre gives of the magical.

When I transform the world by emotion, when I adopt this kind of attitude, the world suddenly comes on to me in the magical mode.

If I cannot resolve the tension in the situation in a way that requires me to adopt the normal means to achieve the desired end, I just CHANGE THE WORLD. I make it one where those means are not required for that end.

For instance, if I am very afraid and faint, what is going on? The original situation was one that called for escape, for removing the fearful object out of reach. If for some reason I cannot do that, or if think I can’t, I (for example) faint. My fainting is a magical way of achieving my goal! I don’t have to adopt the normal means of removing the fearsome object. I adopt the much more radical technique of removing the object by making the whole world disappear! I BLACK OUT! In other words: THE MOVIE’S OVER!

This may not be a very well adapted way of avoiding the object (when I faint, after all, I am all the more vulnerable to the fearsome object), but it does remove it from consciousness. The criticism that it is not well adapted is a criticism someone else would make; it’s a criticism from the outside. From my own point of view, as the one who is experiencing the intolerable situation of being unable to escape from this terrifying threat — that is, as experienced from the inside — it’s quite effective: I am no longer experiencing that intolerable situation.

Thus, Sartre does not think emotions are things that come upon us from the outside. No — they turn out to be much too suspiciously convenient for that. In all cases, Sartre claims, the emotion is just way of achieving exactly the goal desired. This is what he meant all along by saying emotions are purposeful.

Take the case of Janet’s patients, who broke down in sobs. Sartre’s response is: How convenient! The patient just happens to end up not having to go through with the intolerable task of coming to terms with what is really bothering him. For Sartre, it’s not that the patient is unable to continue because he is all choked up and can’t talk. No — he breaks down precisely in order not to be able to go on. What we have here is a magical way of achieving the desired end (to relieve oneself of the intolerable demands of going on) by changing the structure of the situation. The patient chooses this magical approach, since there really is no other tolerable option.

What alternatives are there? He could always just get up and walk out, I suppose. That would achieve the end in a rational, “deterministic” way. But it’s not really a live option.

It would look blameworthy, cowardly. Stern old Janet would not be sympathetic. No, the patient is trying to avoid blame, trying to make it appear as though he weren't responsible for his failure to go on. In short, the patient's sobs are in bad faith.

This notion of the magical also allows Sartre to explain so called sudden emotions — that is, emotions not preceded by a state of tension or frustration. For example, he discusses the case of looking up and suddenly seeing a grinning face pressed against the window, and being immediately struck with terror, horror.

In a case like this, we first perceive the situation in the magical mode. It is not a matter of transforming an antecedently intolerable situation. (Thus Dembo's picture, although Sartre accepts a lot of it, does not apply to all emotions.) This passage (see the outline) repays careful reading.

This sharp distinction between the deterministic and the magical ways of “being-in-the-world” is just one example of a tendency we find a lot in Sartre, the tendency to set things up as stark contrasts, as absolute dichotomies. We have already seen one instance of this with his bifurcation of all reality into being-in-itself and being-for-itself. We also saw it in the passage from *The Psychology of Imagination*, where we had three and only three logically possible types of phenomena. (There, of course, it is a trichotomy rather than a dichotomy, but my basic point still stands.) And we shall see it again, for example, in his discussion of two (and only two) basic forms of interpersonal relations.

Every time we find Sartre doing this, there are always questions that arise about troublesome cases that appear to fall into none of the mutually exclusive alternatives Sartre allows. Are the higher animals — an alert dog, for example — being-in-itself or being-for-itself? If they're conscious, they have to be being-for-itself. But in that case, they are free and morally responsible for their actions (for biting the neighbor kid). On the other hand, if they are not conscious, they are just like rocks and can be kicked around at will. Neither alternative seems very attractive.

So too, what about emotions that don't seem to involve the kind of total transformation of things Sartre has in mind in his notion of the magical? After all, not all emotions are strong emotions. There doesn't seem to be any place in Sartre's scheme for these cases. Again, when we come to Sartre's account of interpersonal relations, there may well be cases that don't obviously fit into either of the two great patterns he allows.

In all these instances, the troublesome cases are exactly the ones to look at to get deeper into Sartre.

False Emotions and the Physiology of The Emotions

Sartre also discusses what he calls “false emotions” (see the outline), as when I pretend to be overjoyed by a gift that doesn't really interest me at all.

Here I impose the form “calling for joy” on the situation, and I behave accordingly. But it’s all a lie. I don’t believe in what I’m doing; I don’t believe the situation really “calls for joy.”

Thus the difference between real emotions and fake ones is exactly this element of belief.

And this allows us, finally, to account for the physiological facts of emotions — the racing heart, the sweaty palms, etc. If I believe in my joy and am not just pretending, I transform the world magically into a joyful world and then LIVE in it.

The fact that I have a body means that I have to live in the world I am conscious of, and not just be a neutral witness to it. Thus, if I magically transform the world in an emotional moment, I must live in that magical world I’ve made. And if I really believe what I’ve made for myself, this living in the magical world will take the form of the physiological phenomena of emotions.

These physiological facts, then, are the phenomena of belief. If I’m convinced by my fear, then the mark of that conviction as lived in the world is my sweaty palms, my racing heart, etc.

Sartre’s theory of emotions, then, is one of the so called “intellectual” theories. (As noted earlier, the term is a bit of a misnomer, since the intellect plays no special role in the theory.) Consciousness is responsible for the physiological disturbance in an emotion, not the other way around. We weep because we are sad, not conversely.

Questions: In looking through Sartre’s text, you may want to ask yourself the following:

- (1) Is Sartre’s account of joy adequate? He’s trying to fit it into the general framework of what Janet had called “setback behavior.” But it sounds strained to me.
- (2) What would Sartre say about drug-induced emotions? For example, what about the general torpor and lassitude produced by tranquilizers? Or what about emotions prompted by electrical stimulation of the brain? Do these things fit Sartre’s analysis?

Part II: Being-For-Itself

We turn now to Part II of *Being and Nothingness*, entitled “Being-For-Itself.”

At this point, I’m going to have to step back from the text a bit, and not follow it so closely as I have been doing. Instead, I am going to try to give you a general picture of the main things that are going on, and the main considerations that are motivating Sartre.

First, let’s do a capsule review of how we got here.

In the “Introduction” to *Being and Nothingness*, we got a few preliminary statements about Being-for-itself. We learned, for instance, about the distinction between positional and non-positional consciousness, and about reflective vs. non-reflective consciousness — distinctions we had already seen in *Transcendence of the Ego*.

But the main emphasis of the “Introduction” was on Being-in-itself. At the end of the “Introduction,” you will recall, Sartre gave us three preliminary characterizations of being-in-itself:

- (1) It is in-itself. (That is, it is metaphysically uncaused, independent.)
- (2) It is. (That is, there is no sufficient reason for it.)
- (3) It is what it is. (That is, it is purely affirmative or positive.)

Sartre did not in the “Introduction” give us a set of similar characteristic features of the for-itself. I gave them to you, but Sartre didn’t. The features I listed were:

- (4) The for-itself is not in-itself. (That is, it is metaphysically dependent.)
- (5) It is. (There is no sufficient reason for it either.)
- (6) It is what it is not and is not what it is. (That is, it has negativity all through it.)

Sartre reveals these characteristics (and others) to us only gradually, throughout the whole book. We began in Part I by taking negativity as our starting point, and our ability to question things as our “guiding thread.”

We saw in Ch. 1 (“The Origin of Negation”) that the presence of objective negativity in the world (translation: the appearance of *négalités* on our phenomenological screen) required that consciousness be its source.

We also saw that in the case of reflection, at any rate, consciousness is separated by a kind of distance (a negative notion) from itself, from its own past, reflected-on self (in the case of the gambler), and its own future, reflected-on self (in the case of vertigo). So, in that sense and in at least those cases, consciousness is what it is not (any longer), and is not (yet) what it is (in the future).

In Chapter 2 (“Bad Faith”), we pushed further. All the cases we treated earlier involved a distance between consciousness and its objects:

- (1) In non-reflective consciousness, the object is the world, or part of it. Consciousness stands at a distance from it. Recall the business about “questioning the carburetor.” (This was a proof of its freedom, he said.)

- (2) In reflective consciousness, the object is ourselves — our reflected-on self.

In bad faith (Ch. 2), however, we found that the same kind of distance, the same kind of nothingness, was involved in consciousness itself, not just between consciousness and its objects.

By studying bad faith, then, we learned that consciousness is separated from itself, at a distance from itself, and is therefore not itself.

This is not to say that one part is separated from another part, as we saw in the discussion of Freud. No, for Sartre the whole of consciousness is separated from the whole of consciousness.

Consciousness then is mysterious, paradoxical, contradictory — but undeniable.

But so far (at the end of Part I of the book) all we have is the mere fact of this separation. We have not yet explored all its implications, all the ways this fissure in consciousness manifests itself. We begin to do that now, in Part II.

Part II, Ch. 1 is called “The Immediate Structures of the For-Itself.” The first thing we need to ask is what the word ‘immediate’ is doing here.

In order to bring this out, we need the notion of an ekstasis (plural = ekstases.) Note the spelling of this word; it is a transliteration of the Greek. Etymologically, the word just means “standing outside” — just as in the English term ‘ecstasy’ (with the different spelling), when we are “beside ourselves” with joy or rapture.

Sartre uses the term ‘ekstasis’ in a technical sense he got from Heidegger. For Sartre, there are three fundamental ekstases of consciousness, three basic ways of getting outside the confines of the momentary, instantaneous cogito.

The term ‘ekstasis’, then, basically means a way of getting outside those confines. Recall, this was exactly the task Husserl had taken on in The Idea of Phenomenology: how to get around Descartes’ problem. In effect, what we are seeing here is Sartre’s answer to the same question.

Some of that answer will look a lot like what we have already seen in Husserl. But other parts will be new to us. (That’s not to say there aren’t similar things in Husserl, but we haven’t looked at them in Husserl, and we are not going to now.)

These three fundamental ekstases are the topics of subsequent chapters in *Being and Nothingness*. Here they are (we’ll look at them in more detail later):

- (1) Temporality — time. My momentary act of consciousness, right now, also has a past and a future. Sartre discusses time in the next chapter: Part II, Ch. 2.
- (2) Transcendence. This term is used in lots of different senses by Sartre. Basically, it means a “going beyond.” In this general sense, temporality is a kind of “transcendence” too; consciousness

transcends itself back into the past and forward into the future. But as he uses the term here (for the second ekstasis — discussed in Part II, Ch. 3), it is more restricted. In this context, it just means the relation of positional consciousness to its objects, the “reaching out” toward objects. In this usage, then, ‘transcendence’ is just another term for intentionality.

Let me say a little more about this now. In this restricted sense of the term, transcendence is sometimes called knowledge. This is important. Positional consciousness knows its objects. (This knowledge need not be verbalized or articulated. That is a later development)

Thus, when I perceive the cube, Sartre will say I know the cube. I know it is a perceived cube. That is not to say I know it really is a cube (there is always the risk, remember). But I know it as a cube. And, once again, this need not be a matter of verbalized or articulated knowledge. I don’t have to know that what I’m perceiving this thing as is what is called a “cube.” In short, I don’t have to know any geometry to know that, yes, this is the way I am perceiving this object.

The same thing applies to imagination and conception. They are other ways of knowing the cube — or knowing the object as a cube.

Get used to this terminology. For Sartre, knowledge is confined to this relation of positional consciousness to its objects.

This is not just a matter of terminology; it isn’t sheer arbitrariness. Sartre adopts this terminological convention because he thinks it accurately reflects our normal notion of knowledge. And to some extent it does. “Knowledge” here is what we ordinarily mean by “acquaintance,” which surely counts as one familiar kind of knowledge.

When, for example, consciousness believes something in bad faith, its belief is a failure, as we saw in the Chapter on Bad Faith. Consciousness does not succeed in being fully belief. Furthermore, consciousness is non-positionally aware of the failure of its belief. It is aware that it is making itself believe, despite inadequate evidence, so that it might be wrong. It is aware that it does not fully believe.

But this does not mean consciousness knows it does not fully believe. That would require a positional consciousness of the failure of belief. We know it — now, after the fact, by a kind of reflection in which we become positionally aware of what we were doing. But at the time we did not know it.

We were conscious of it at the time; aware of it. But this awareness was not a kind of knowledge; it was a kind of being, the particular kind of being characteristic of the for-itself. In short, at the time I did not know the failure of my belief; I was not thinking about it. But I was certainly experiencing it, undergoing it. What I was doing at the time was exactly failing to believe fully. And since for consciousness there is no distinction between the action and an agent (consciousness is a process, an event, recall), this means that what I was at the time was a failing to believe fully.

Copyright © 1996 by Paul Vincent Spade. All rights reserved. Permission is hereby granted to copy this document in whole or in part for any purpose whatever, provided only that acknowledgment of copyright is given.

This illustrates an important point Sartre made very strongly in the “Introduction” to *Being and Nothingness*, a point we are only now in a position to begin to see the importance of:

BEING IS NOT TO BE MEASURED BY KNOWLEDGE.

As Sartre sometimes puts it, *knowledge is not primary*, or *we must abandon the primacy of knowledge*. (See, for example, § 3 of the “Introduction,” around p. 11.)

The solution to a number of *important* philosophical problems will rest on this.

For example, to say that *being is not measured by knowledge* is in effect to say that the way things *appear* to *positional* consciousness does not necessarily *exhaust* the way things are.

To think it *does* is a kind of *idealism* — the view that reality is composed of, constructed out of, *ideas* (in this context, read: *phenomenal objects*.)

Sartre rejects this. Recall how in the “Introduction” (§ 2) he insisted that *the being of the phenomenon is NOT the phenomenon of being*. That is, the reality of what’s out there (the *being* of the phenomenon) is not exhausted by what I’m thinking about when I focus on the *fact* that the thing exists (the *phenomenon* of being). That’s not to say that the reality of what’s out there is *hidden* from me, like some kind of Kantian thing-in-itself, but only that there’s more there than what I’m positionally conscious of. In the “Introduction,” Sartre used the term ‘*transphenomenality*’ to make this point. (Go back and review the outline of the “Introduction.”)

We have in effect seen the same point in other ways. *Being is not to be measured by knowledge*. It was, in effect, confusion on this point that was behind the *Self-Love Theory* we saw in *Transcendence of the Ego*. The people who held this theory thought that the way my *motives* appear to me on *reflection* is the way they originally are; I go to help Pierre *in order* to reduce my discomfort at his need.

Sartre thinks Husserl too was guilty of the same mistake. The fact that consciousness appears *on reflection* to be endowed with an Ego led Husserl to suppose that *even before reflection*, consciousness was presided over by an Ego.

For Sartre, then, it is Husserl’s later *idealism*, his *measuring being by knowledge*, that is the radical origin of his doctrine of the Transcendental Ego.

So much for now about *transcendence*, the second of Sartre’s three *ekstases*. The third one is:

- (3) *Being-for-others*. That is, the presence of “other minds” — other “for itselfs.” This is the topic of Part III of *Being and Nothingness*, and some of the implications are worked out further in Part IV.

These then are the three Sartrean *ekstases*. Now I introduced these as part of my explanation of what Sartre meant by the title of Part II, Ch. 1: “The Immediate Structures of the For-Itself.” We are now in a position to see what he meant:

To talk about the *immediate* structure of the for-itself is to talk about the structure of the for-itself *without* bringing these three *ekstases explicitly* into consideration. They are of course *implicit* from the beginning, but they can be brought out only gradually, and we’re not yet ready to do that.

Presence to Self

Let’s begin with Part II, Ch. 1, § 1: “Presence to Self.”

On p. 121, Sartre makes a point we have already seen in our discussion of *Transcendence of the Ego*:

THE ACT OF REFLECTION ALTERS THE FACT OF CONSCIOUSNESS ON WHICH IT IS DIRECTED.

In general, of course, it is *not* true that positional consciousness *distorts* the object it is conscious of. But in the special case of positional consciousness of *consciousness itself* — that is, in the special case of *reflective* consciousness — there *is* a distortion for the reasons we saw earlier. (Recall my example of the little wax statues.) And this distortion may mislead us *if* we assume that the consciousness that appears to us in *reflection* is structurally the same as it was *before* we reflected on it.

This was the mistake of the Self-Love theorists, and of Husserl. They overlooked the fact that reflection *distorts* its objects.

(Sartre grants that Husserl himself had made the same point, that he recognized that reflection distorts, but seems to think he then lost track of it and did not realize its implications.)

Now, while *non*-positional self-consciousness is not *reflective* in the sense we have talked about all throughout this course, it is nevertheless LIKE (as Sartre says, “*homologous*” to) reflective consciousness in the sense that it too *alters* what it is consciousness of: namely, *itself*.

Non-positional self-consciousness thus *alters itself, changes itself*. And, of course, non-positional self-consciousness is not a *distinct* consciousness from positional consciousness, but only a distinct *aspect* or *side* of one unified consciousness. Thus, consciousness itself *alters itself, changes itself*.

That is to say, it is constantly *in flux*, changing. In fact, it *is* a flux, a change. The *being* of the for-itself is not like the stable, inert being of *things*, of *substances* — that kind of being is appropriate to the *in-itself*. Rather, the being of consciousness is like the being of an *event*, a *process*. We should not think of consciousness in terms of *nouns*, but of *verbs*.

Strictly speaking, I suppose, we shouldn't even say that the for-itself exists, but rather that it happens. From this point of view, to say that there can be no combination in-itself-for-itself is just to say that there can be nothing that is simultaneously a thing and a process (event).

There is something like reflection at the non-positional level. Indeed, Sartre uses similar terms here, and it's easy to get confused. He calls consciousness a dyad: the dyad of reflection-reflecting. (Don't be confused; this has nothing to do with the problematic notion of "pure reflection.")

But be careful. The English term 'reflection' here is misleading. This is not "reflection" in the sense we are familiar with, in the sense of reflective consciousness. In fact, in the French, Sartre uses a different term altogether here, one that doesn't really have a smooth English equivalent. What Sartre is talking about here is not "reflection" in the sense of reflective consciousness, but in the sense of an image in a mirror. We use the term 'reflection' in English for this too.

On the other hand, the other term in the "dyad," the "reflecting," means what does the reflecting — the mirror itself.

Thus, Sartre's talk about consciousness as a dyad: reflection-reflecting really means that consciousness is a lot like a mirror that reflects itself.

Don't think of the familiar department store arrangement in which you have two mirrors reflecting one another. Consciousness is not two things (that is what Freud mistakenly thought), but one. Instead, think of a single mirror shaped into a hollow sphere, with the reflecting surface on the inside. That's a good image of what consciousness is like, of the way in which it is non-positionally aware of itself. (Don't worry about how you get any light in there; it's just an analogy.)

This "dyad" of reflection-reflecting, which is like but not the same as "reflection" in our earlier and more familiar sense, is what Sartre calls "Presence to Self."

Now of course the fact that consciousness exists like an event, a process, means that it cannot be pinned down. It is not stable. It is never entirely what it is; it is already — without any passage of time as in the cases of the gambler and of vertigo — something else.

Facticity

Let's turn now to § 2 of the chapter.

Consciousness is. Like the in-itself, it exists (although in a quite different way). It is a brute fact, a violation of the Principle of Sufficient Reason. There is no good reason why consciousness should exist.

Furthermore, Sartre goes on, to say that it exists is not just to make a general, empty, abstract claim about consciousness. It always exists in a very particular way. It exists as a Philosophy professor at Indiana University, or as a student who finds himself or herself in very particular circumstances. In effect, all Sartre is doing here is making a point we have seen a long time ago: you don't get the individual, the particular, as a product of generalities.

Now, given that there is no good reason why I should exist at all, why in addition should I exist in the particular, unique, individual way I do? For Sartre, there is no good reason for that either.

As he puts it, there is no foundation for it. But we have to be careful. There is a sense in which being-in-itself is a kind of foundation for my existence — in the sense that I cannot exist as consciousness without being positionally conscious of being-in-itself. My being rests on the being-in-itself of the screen. That is in effect the point of Sartre's so called "ontological argument" in § 5 of the "Introduction." And, as he makes clear there, it is just another way of putting the theory of intentionality.

But that kind of foundation is not enough. Being-in-itself, although it is a necessary foundation for my being-for-itself, is not a sufficient one. There is no sufficient one.

In particular, I am certainly not a sufficient foundation for my own particular being. I am not ontologically or epistemologically "self-justifying." (Sartre does say that I am in a sense a sufficient foundation for my own nothingness, but not of my being.)

That is, there is no good reason why I should exist at all, and exist as the particular, unique individual I am in the unique circumstances in which I find myself. This point is what Sartre calls my Facticity. It is a fixed and perfectly definite fact I have to come to terms with. (We'll have to be a little more careful later on about the sense in which it is fixed and perfectly definite.) It is immovable, solid. And since it is fixed and definite, immovable and solid, it smacks of the in-itself, whose characteristic it is to be immovable and solid.

But we have to be careful. I am certainly not a being-in-itself; I am a being-for-itself. And there is no being that can be both a for-itself and an in-itself. That would be magical; it would be God.

So when Sartre says, as we have already seen him say in the Chapter on Bad Faith, that we are all combinations of "facticity" and "transcendence," we should not think of this as in any way a combination of the in-itself and the for-itself. Sartre won't allow any of that.

But that doesn't mean that the in-itself and the for-itself are cut off from one another, unrelated, without communication. The two are intimately connected, at least in one direction. Consciousness always refers to the in-itself. (By contrast, the in-itself does not refer to consciousness, or to anything else. It just sits there.)

How does it do this? Well, we have already seen one way in which consciousness refers to the in-itself: Sartre's so called "ontological argument" in § 5 of the "Introduction."

Every act of consciousness is positional consciousness OF being-in-itself. (That's just the theory of intentionality.)

Thus, we are already familiar with the way in which the in-itself, so to speak, haunts the positional side of consciousness. But now we see — in the present section on “Facticity” — another way in which consciousness “refers to,” is “haunted” by, the in-itself. The in-itself haunts the non-positional side too.

The fact that consciousness exists at all, and that it exists as the particular, unique being it is — this is the in-itself haunting the non-positional side of consciousness.

To put it in the terms that we've learned from our preliminary discussion of the second ekstasis (above), this means that consciousness is related to the in-itself not just with respect to what consciousness knows, but also with respect to what it is.

How then can I be said to be “responsible” for my facticity?

I have nothing to say about my facticity. I can't control it, I can't prevent it. It's not my fault. I didn't ask to exist, after all. And yet, in an important sense, Sartre says, I am responsible for my facticity.

I am responsible for it in the sense that it's up to me what to do with it, what I make of it. I am completely free in that respect — the possibilities are endless. It's up to me how I go beyond my circumstances, how I transcend them (in the broad sense, not the narrow sense of Part II, Ch. 2, where transcendence = intentionality = knowledge).

If we think of the familiar model of choice as “the fork in the road,” we have lots of alternative forks to choose from. The possibilities are endless. I can go any direction I want from here. But — and here is where facticity comes in — I start here! That I have no control over.

I am a combination of facticity and transcendence. Recall the example of the Waiter in the Chapter on Bad Faith. I am a facticity transcended, the transcending of a facticity. I am an event that takes place in a certain definite context. I have no control at all over what the context is; but I do decide what happens in that context, and in that sense am responsible for it.

I am not responsible for my facticity — the context — in the sense that I can bring it about or can prevent it. But I am responsible for it in the sense that it is IN MY CARE. It is entrusted to me.

Thus, while I am not the foundation of my own being (and nothing else is the foundation for it either), I am the foundation of my own transcendence, Sartre says. That is, of my own nothingness, of my own “going beyond,” my own putting myself at a distance from my facticity.

Lack

In § 3 of the Chapter, Sartre discusses the notion of “lack.”

Since consciousness is always non-positionally aware of itself — that is, of what it is, not of what it knows (the latter it is aware of positionally) — and since, as we have just seen, I am not the foundation of my own being, it follows that I am always non-positionally aware that I am not the foundation of my being, that I am a gratuitous fact, a logical outrage. In short, I am always non-positionally aware that I am imperfect. Something is lacking to me. (I am, so to speak, “not all there.”)

Recall the objective “lacks” in the world we discussed in Part I, Ch. 1, “The Origin of Negation.” Those lacks were négatités, negativities in the world. But now we are talking about “lack” in consciousness itself.

Descartes built a proof for the existence of God out of this deep awareness of our own imperfection or lack. He argued in *Meditation* III that, since I am imperfect, there is no way to account for the idea I have of a perfect being, unless such a being really exists. I certainly could not have come up with such an idea from my own resources.

Sartre rejects the argument, of course. But it is not a silly argument, and it is based on the profound realization of our own imperfection that Sartre now wants to focus on.

One perfectly familiar way in which this lack or imperfection shows itself is in the fact that human beings have desires.

Hegel was one of the first philosophers to point out the metaphysical significance of the fact that human beings have desires. Sartre is developing that point here. (Plato had talked about it earlier — in the *Symposium* — but in very metaphorical terms.) For both Hegel and Sartre, desire is only possible for a being that is missing something, that is “not all there.”

Sartre analyzes the notion of lack into three components:

- (1) The Lacking. That is, what is missing, the part that is not there.
- (2) The Existing. That is, what misses it, the part that is there.
- (3) The Lacked. That is, the whole, what you would get if the Existing ever got what was Lacking to it.

The formula, then, is:

$$\text{The } \underline{\text{Lacked}} = \text{The } \underline{\text{Existing}} + \text{The } \underline{\text{Lacking}}.$$

Sartre gives an example of a crescent moon. The crescent moon is (2), the Existing. What it is (1) lacking is the rest of the moon. The result of adding the crescent moon to the rest of the moon is (3) the Lacked — that is, the full moon.

Now, let's apply this to consciousness. In the case of consciousness, what is the existing? Well, it's just consciousness itself. What is lacking to consciousness? What is it missing? Well, what it is missing is exactly itself. Consciousness, remember, is not fully itself. It is not what it is. There is a fissure in consciousness.

The waiter in the Chapter on Bad Faith plays so hard at being a waiter exactly because he is not fully a waiter. (If he were, he wouldn't have to work so hard at it.) If only he could be fully a waiter, nothing but a waiter, he would know exactly where he stood. And of course, that is just what he wants; that is his desire. Perhaps by playing the role of a waiter so conscientiously, he can somehow become a waiter in the full sense, perhaps he can finally be a waiter, he can "catch up" with himself.

That is exactly what is lacking. He is not fully a waiter.

What then is the lacked? What would be the result if the waiter ever did "catch up" with himself, "recovered himself," as Sartre says? What would be the result if he succeeded?

Well, again the answer is: himself. He would finally be one with himself; the wound in his being would be healed; the fissure filled in. In short, he would be a being-in-itself. But he would still be a waiter — that is, a being-for-itself. He would be both.

But of course that is impossible. "Man is a useless passion," as Sartre says at the end of *Being and Nothingness*. He is in bad faith. He is trying to become the sufficient foundation of his own being, just as he is indeed the sufficient foundation for his own nothingness. He is trying to be God, who is the sufficient foundation of his own being. (This is what is behind the traditional notion of God as in some sense "self-caused." the existence of God is not supposed to be a gratuitous fact.)

So the for-itself is simultaneously the existing, the lacking and the lacked.

Perhaps what Sartre is talking about here will be a little clearer if we think of the various popular "self-help" books that talk about the "real you." The "real you" is supposed to be some kind of ideal self, buried down deep inside you, hidden by the everyday you. And the point of these "self-help" books is to get you to let your "real you" out, to let it shine forth in all its glory. (Because, of course, the "real you" is *very, very good!*)

Notice the implications of this kind of talk. That "real you" is supposed to be who you really are. In a deep and profound sense, that is your identity. And there is some truth to this, after all. No matter how vague and unarticulated that notion of the "real you" is, nevertheless we recognize ourselves there. It is a notion that speaks to us. But of course if that really is who you "really are," then there shouldn't be anything left for you to do, and the whole point of the self-help book would be lost. The notion that the book's advice and encouragement is going to do us any good presupposes that we are not who we "really are," but still have to work at it.

This is the kind of thing Sartre is talking about. He does not think the popular psychological talk about a "real you" is silly. On the contrary, he thinks there is a profound truth behind it. But he also thinks the "real you" (= the lacked) is impossible to achieve. It is a goal out of reach. So all such self-help books are doomed to failure.

Furthermore, he thinks the “real you” is not something given in advance, so that our task would be simply to find out what that real self is and then try to identify ourselves with it. No — for Sartre, we choose that goal ourselves.

Thus, human beings are lacks. But this abstract structure is never found in practice just in the abstract. It is always very particular, unique. Our attempt to recover ourselves, to “catch up” with ourselves, is always colored by the particularities of our facticity. The waiter, after all, is trying to be a waiter, not a diplomat or a carpenter. The abstract is not primary, the particular is. Husserl’s eidetic abstraction, which gives us universals, is very useful for Sartre. But it will not give us the full particulars.

Note: This feature of consciousness by which we aim at ourselves on the other side of our facticity is what Sartre calls “The Circuit of Selfness.” We start from ourselves (the existing), look at the world and our situation in it, and aim at ourselves (the lacked) on the other side of the world. I must say, I find Sartre’s discussion of “The Circuit of Selfness” just plain obscure, and more than a little confusing. Don’t worry over it.

Value

Let’s now talk about Sartre’s notion of value. (We are still in § 3 of the chapter.)

From what we have just discussed, it follows that every individual person is a particular project, a particular attempt to transcend a particular facticity and become one with himself or herself, to become the foundation of this particular being. (Sartre doesn’t say that human beings have projects; they are projects.)

The particularity of this project is what gives rise to different values. For the waiter, things take on value precisely in the light of his project to be a particular kind of waiter, to be a particular kind of waiter-God. Everything is evaluated in terms of that goal, the impossible, ideal goal. Everything is evaluated in terms of, measured against, the particular kind of God the waiter is trying to be.

This is the profound meaning behind the traditional doctrine that God is good, and indeed is the standard of goodness, and that we ought all to try to be like God. You should also think here of the famous line at the beginning of Aristotle’s *Nichomachean Ethics*: “The good is that at which all things aim.” And in fact, Aristotle there goes on to give an excellent discussion of “means and ends,” of how we evaluate lesser, subordinate goods in terms of their utility in achieving higher goods, so that the highest good of all would have to be totally useless! This is very much the kind of picture Sartre has too, except of course that Sartre doesn’t think there is anything fixed and absolute about these goods.

We frequently tend to think that values are found ready-made in the world, that ethical and moral standards are absolutes, that certain things just are good and others just are bad, just in themselves. Certain standards or codes of ethics just are authoritative, and it is up to us to find out which ones they are and to conform to them.

Notice: On Sartre's doctrine, this would be true if all our projects were necessarily the same — and the same in detail. In that case, the same standards of right and wrong would apply to everyone, and would not be subject to negotiation. But in fact, of course, we do not all have the same projects in mind — much less necessarily having the same projects. (If we did all have the same projects, then since a human being just is such a project, there would be only one human being!)

This notion that values are ready-made absolutes found in the world is what Sartre calls the "Spirit of Seriousness." He accuses the bourgeois of having this attitude. They regard their norms and values as eternal truths. One does this; one does not do that — and that's the end of the matter. Everything in its place — a place that is carved out by the very nature of things.

In his novel *Nausea*, Sartre at one point describes going into the local art gallery in a provincial town, and seeing there the heavily framed portraits of all the town's founding fathers and civic pillars. They are all hanging there, very serious and very stone-faced. They are frozen "absolutes." They are the very embodiments of that town's values, solid and fixed — and above all, objective.

A world like this, with objective values, is comfortable. It is a world in which it is possible to fit in, to belong.

Sartre thinks this "Spirit of Seriousness" is in bad faith. To be sure, values do appear in the world. In a famous phrase, he says values "spring up like partridges." But that doesn't mean they are there ready-made.

No — they are projected there. We find them there because we read the world in terms of our own project to be God in our own unique way. In the end, consciousness is the source of value.

The situation here is the same as what we have already seen with emotions. My anger first appears in the world — as a particularly outrageous deed, for example. But of course I am the one who made it outrageous.

As far as ready-made, absolute values are concerned, there are none. The Ten Commandments have no ultimate authority. If I accept them, if I feel their weight, it is because I choose to. In the end, I am the final authority.

If we ask about absolute values, ready-made values carved in stone — well, from that point of view, Sartre says (p. 797),

... it amounts to the same thing whether one gets drunk alone or is a leader of nations.

This obviously makes the notion of an "existentialist ethics" very hard for Sartre. At the very end of the conclusion of *Being and Nothingness*, he promises that he is going to write another book, on ethics, spinning out in detail the ethical theory implicit in *Being and Nothingness*.

But the book never appeared, and it is more than a little hard to see what it could possibly say. It looks as though his ethics could amount to no more than “Anything goes.”

And yet it is clear that Sartre doesn’t want that. That is one of the charges against existentialism he considers in “Existentialism Is A Humanism,” and it is clear in that essay that he thinks the accusation simply doesn’t apply to his doctrine.

Note: Although the promised book on ethics was never published, Sartre did do a lot of work on it. His notes and rough drafts of portions of it were published posthumously in 1983 under the title *Cahiers pour une morale* (Paris: Gallimard, 1983). It has been translated into English under the title *Notebooks for an Ethics*, David Pellauer, trans., (Chicago: University of Chicago Press, 1992). And it must be said that the book is something of a disappointment. What the book contains is not Sartre’s notes for the ethics implied by *Being and Nothingness*, but for something rather different — the sort of ethics that fits in better with his later, more Marxist views.

It is fairly clear what Sartre wants. He wants there to be one absolute value after all, one ultimate existentialist virtue: namely, authenticity. (Recall the cryptic note at the very end of the chapter on “Bad Faith.”) When we are authentic, we reject the “Spirit of Seriousness,” we recognize that we determine our own values, and that we are utterly without guidelines in picking our values. If we pick a certain set of guidelines — for example, the Ten Commandments — why did we pick those instead of others? Isn’t it because we find them especially worthy — that is, haven’t we already chosen our values before we can see the Ten Commandments as good ones, as ones that deserve to be followed? There is no avoiding it: we are ethically on our own.

To be authentic means to realize this fact — and then to act. It makes little difference in the end how we act. Just do! Take a stand! Create your values. And be fully aware of what you are doing. Realize your freedom in this respect, and act anyway. Act in anguish. (Recall how anguish results from the awareness of our freedom.)

The point is, you are going to “create” your own values anyway, whether you admit it or not. What Sartre wants us to do is to realize this, to know what is going on.

But how can he say this? What is so great about authenticity? Is it any more an absolute value than all the rest? What is good about authenticity? What is bad about bad faith?

Is Sartre simply practicing what he preaches here, taking a stand in favor of authenticity, with no guidelines? But then why should we accept it? Sartre might reply that we don’t have to; we must make up our own values, and all he is doing is recommending authenticity to us.

But if that is all Sartre can say, then it follows that there is no special ethical attitude that follows from the ontology of *Being and Nothingness*. It allows all ethical codes, and requires none of them — not even authenticity. But that is just to say that there is no such thing as an “existentialist ethics” — that is, an ethics that follows naturally from the ontology of *Being and Nothingness*.

There are at least five books you should know about in this connection, books that try to sketch an existentialist ethics despite the difficulties I've raised. They are all classics, and can all be recommended. (For full references, see the list of *Books on Reserve*, in the course packet.)

- (1) Simone de Beauvoir, *The Ethics of Ambiguity*.
- (2) Hazel E. Barnes, *An Existentialist Ethics*.
- (3) Thomas C. Anderson, *The Foundation and Structure of Sartrean Ethics*. (This is an excellent work.)
- (4) Francis Jeanson, *Sartre and the Problem of Morality*.
- (5) David Detmer, *Freedom As A Value: A Critique of the Ethical Theory of Jean-Paul Sartre*.

All but the Jeanson book are available in paperback.

These books are not confined to the limits of *Being and Nothingness*. That is, they look at Sartre's thinking as a whole.

From my own personal point of view, I found Anderson's and Detmer's books to be extraordinarily rich. They go a surprisingly long way toward establishing that Sartre could and did have a legitimate ethical theory that grows naturally out of *Being and Nothingness*. Nevertheless, I think both books ultimately get that theory wrong. We will have a little more to say about this at the end of the course.

As for Jeanson's book, Sartre himself wrote a letter that is published in the same volume as a kind of Foreword. Here is what he says, in part (p. xxxix):

... you did not hesitate to take existentialist morality as your guiding theme, which was all the more worthy since that part of the doctrine has not yet been really treated — at least not in its totality — and since most of the critics, choosing to refute theses which I have not yet advanced and which they do not know, have introduced the deepest confusion into this matter. ... you have so perfectly followed the development of my thought that you have come to pass beyond the position I had taken in my books at the moment I was passing beyond it myself and to raise with regard to the relations between morality and history, the universal and the concrete transcendence, the very questions I was asking myself at that same time.

Jeanson's book first appeared in French in 1947, just four years after *Being and Nothingness* was written and only a year after "Existentialism Is A Humanism" was published. Sartre's letter was included in the first edition.

Notice what he is saying in the passage I just quoted you. In effect, he is saying, "This is the book on morality I would have written myself, but never did." You could hardly have a stronger recommendation than that!

Copyright © 1996 by Paul Vincent Spade. All rights reserved. Permission is hereby granted to copy this document in whole or in part for any purpose whatever, provided only that acknowledgment of copyright is given.

But notice also that Sartre talks about “passing beyond” certain positions. That is, Sartre’s ethics is not just a matter of turning the crank of *Being and Nothingness* and generating some consequences for morality. It’s a matter of thinking hard about the topic, and being prepared to revise one’s former opinions if that is necessary.

Possibility

In § 4 of the Chapter, Sartre discusses the notion of possibility. There is an initial problem with the notion of possibility:

- (1) On the one hand, we say that possibility is prior to being. That is, possibility is the weaker notion. Something can be merely possible and yet not exist. For example, it is possible that you will all fail this course. But that possibility is not a reality, and (let us hope) never will be.
- (2) On the other hand, possibilities have some kind of being. They are in some sense real. We speak of real possibilities. It is a real possibility that you will all fail this course. And yet that possibility is not a reality, and (we hope) never will be. At the same time, other things are not real possibilities. It is not really possible, for instance, to find a largest prime number.

Mere possibilities are therefore somehow real, and yet not real. Thus they are what they are not, and are not what they are (namely, real). This of course should tell us right away where Sartre is going to locate the origin of possibility.

Sartre insists that possibility is grounded in the actual. Things have possibilities. We look at the sky, for instance, and say it is possible that it will rain. For Sartre, this possibility appears to us as a property of the sky.

It is perhaps not immediately obvious what Sartre’s point is here. But it fits into a quite traditional way of thinking about the relation between possibility and actuality. One way of thinking about this is to start with a large supply of possibilities, among which certain possibilities are somehow singled out as the actual ones. Possibility in this sense is in some way a matter of logical consistency. Leibniz had a view like this.

Sartre rejects this theory. It would ground the actual in the possible rather than, as Sartre holds, the other way around.

For Sartre, the notion of possibility is based on the powers of things. To say it is possible that it will rain is not to say merely that it is logically consistent that it should rain, but rather that the present state of the weather might come to produce rain. This is not some abstract logical point but rather a matter of the natures of things, their causal powers.

This notion of possibility is also behind the Aristotelian notion of the *potentialities* of things. The acorn is *potentially* the oak tree, meaning that it is within the *power* of the acorn to develop into an oak tree.

Sartre's account of possibility in effect sides with Aristotle rather than Leibniz. (And, it is perhaps important to note, until very recently most of the rest of the history of philosophy sided with Aristotle too.)

So the *possibility of rain* is a kind of *property* of the sky. But it is a property that in a way *goes beyond* what is merely *given* — much as, in perception, the three sides of the cube *promised* three more sides around in back.

There is a difference, of course. The sky doesn't yet *promise* rain; the connection is *weaker* than that. It is only *possible* that it will rain. But the reference to *more* is there.

(The reference is not *guaranteed*, of course. It may *not* rain in the end. This means that, although there is an important difference between the mere *possibility* of rain and the perception of the cube — the possibility of rain is not a full-fledged *promise* — nevertheless both belong under the broad heading *perception* as described in *The Psychology of Imagination*.)

In possibility, as in perceiving the cube, there is a *reference* beyond, a *transcendence*. And, Sartre says, in possibility just as in perceiving the cube, this reference beyond is something *we* impose on things. We *project* possibilities onto things. (That doesn't make it *subjective*, for the reasons we have discussed before in connection with Bergson's theory of negativity.)

We can do this, we can project possibilities onto things, because *we ourselves* can go beyond the *given* too. We *transcend our facticity*. The possibility that is what it is not and is not what it is (namely, a *reality*) must ultimately — like negativity in general — come from a *being* that is what it is not and is not what it is, from the *for-itself*.

When we considered nothingness, we found not only that nothingness in the world came from consciousness, but also that *consciousness itself IS* a nothingness. So too, *possibility* not only comes from consciousness, consciousness *is its own possibilities*. Its *freedom*, its ability to transcend its facticity, *that* is what marks out the for-itself's possibilities.

Here is an example of a point we will see over and over again in Sartre. There are many philosophical problems that seem insoluble; they have been with us since the beginnings of philosophy, with no apparent progress. Certain philosophical puzzles appear to admit of *no* resolution. One perhaps suspects they are just plain *contradictory*, although logically minded philosophers don't want to admit that, and fondly *hope* such problems can in principle be resolved.

- (1) The nature of consciousness is one of these traditional philosophical puzzles, with consciousness's peculiar ability to be *self-consciousness*, and with its peculiar ability to *deceive* itself.

- (2) The notion of non-being or nothingness is another such problem. There is an obvious sense in which non-beings are real; Pierre really is absent from the café. And yet how can non-being be real? (Recall Parmenides.)
- (3) The notion of value is another traditional sticking point. The things most really valuable (namely, ideals) are the least real things of all.
- (4) We now see that possibility is another such apparently contradictory notion. There are real possibilities that will never be “real.”
- (5) Time, as we shall see, will be another such traditional philosophical puzzle. (See Part II, Ch. 2. We will be talking about this in just a moment.)

Now, one of the beauties of Sartre’s philosophy is that, although it has the perhaps unwelcome feature of openly embracing contradiction (at least that’s unwelcome to some people), nevertheless it also has the considerable theoretical economy of reducing all these contradictions to one big one: to consciousness. In the present instance, as we have seen, it is possibility that is rooted in consciousness.

Time

So too, time is grounded in consciousness. (And now we turn to Part II, Ch. 2.)

For Sartre, there are lots of wrong ways of thinking about time. For example, there is:

- (1) The picture of time as a great container, in which the events of the world all take place in sequence. This is the notion of what Sartre calls “the time of the world,” and discusses in Part II, Ch. 3, § 4. For Sartre, this notion of time is a secondary, derivative notion.
- (2) The picture of time as the summation of “times.” That is, the picture of Time as a whole as what you get if you add a whole bunch of instants or little intervals together. There are “times” — yesterday, today, tomorrow, for example — and Time at large is just all of them put together.

For Sartre, the “summation” picture (2) puts the emphasis in the wrong place. And now I’m going to give you an elaborate parallel:

JUST AS for Existentialism in general it is fruitless to try to start with the two isolated and independent notions “man” and “world” and then try to explain the complex relation “man-in-the-world” in terms of those two isolated poles — rather we must start with the

concrete totality “man-in-the-world” and only in terms of that can we understand the notions “man” and “world” (recall the passage on p. 33, at the very beginning of Part I, Ch. 1: “The Origin of Negation”):

It is not profitable first to separate the two terms of a relation in order to try to join them together again later)

— SO TOO, Sartre thinks it is fruitless to start with the notion of isolated, individual times, and then construct time as a whole out of them. Instead we must start with the notion of time as a whole and let that notion shed light on the notion of “times.” (But be careful: We don’t want to begin the notion of time as a whole in the sense of the great container theory. That theory is too abstract too.)

For Sartre, the trouble with both the above pictures of time — the great container theory and the summation theory — is that they make most of time not exist!

The past, of course, doesn’t exist; it did exist but not any longer. The future doesn’t exist either; it will exist, but doesn’t yet. Sartre warns us, for example, to beware of thinking of the past as though it really does still exist, but is somehow in a kind of state of retirement, just isn’t active any more.

Thus, only the present exists. But the present is just an infinitesimal moment, a kind of limiting point, the last of the past and the first of the future — a moment so short that it has no duration at all. It does not last. It endures for no time at all.

So both the above theories are inadequate. Time is not just a great container in which the events of the world take place in a certain order. It couldn’t be that, since on such a theory most of the container would not exist.

Neither can time be just the summation of a series of times, of instants or intervals. Most of these times — in fact, all but one, the present — don’t exist, and so can hardly be called parts of a whole.

So the problem about time is that it is both

- (1) Quite real — it is real enough to “heal all wounds,” for instance. The simple passage of time seems to have very real effects of its own. And
- (2) Yet most of time doesn’t exist, and so is not real.

The problem with both the great container theory and the summation theory, then, is that they do not do justice to (1). They do not explain the reality of time.

Here is another of those paradoxical situations we have already seen with nothingness, with value, with possibility.

Instead of starting with more or less scientific and mathematicized theories of time, which try to be consistent and are then embarrassed to find that they lead right into this paradox

(or can avoid it only by denying the obvious reality of time), Sartre thinks we should recognize the paradox of time at the outset and then construct our theory in accordance with that paradox — not even try to avoid it.

The trouble with time, like the other puzzling things we have investigated, is that it is both real and unreal; it is both being and nothingness. And so we should expect time to be based on consciousness, which is ultimately the only kind of being that is also a nothingness.

In the Chapter on time (Part II, Ch. 2), Sartre discusses it under two main headings:

- (1) The phenomenology of time, how it appears to us (§ 1 of the Chapter).
- (2) The ontology of time, what it really is (§ 2 of the Chapter).

These two discussions are closely related, and ultimately in agreement with one another. But they are not the same thing. To say they are is to measure being by appearing, to measure being by knowledge. And we have already seen what Sartre thinks about that.

Nevertheless, for present purposes, we don't have to worry too much about the distinction between these two parts of Sartre's discussion. The points we need to note don't require that level of sophisticated refinement.

First of all, time comes in past, present and future. But we shouldn't think of these as absolutes, as general things, independent of the events that occur in them. (That would be the great container view again.)

On the contrary, things HAVE a past. Things HAVE a future. Things HAVE a present. (Recall how the similar point worked when we were discussing possibility.)

The notion of past, present and future is thus always a relative notion — relative to the being of which it is the past, present or future.

For instance, I have my past; you have your past. And they are not the same. I did graduate work in Canada; you did not.

Similarly, I have my future, although I don't yet know what it will be. And you have yours. They will not be the same.

(The present is an odd case, as we shall see.)

So too, the whole world has a past (namely, world history) and a future too. (But only in a special and derivative sense, as we shall see.)

Thus, to be past, present, or future is to be SOME X'S past, present, or future.

Now a being's past has a peculiar feature. It is settled, over and done with. It cannot be changed — although perhaps it can be added to as time goes on and the being gets older, and although perhaps there once was a time when what is now in the being's past could have been prevented or done otherwise.

There is nothing mysterious about what Sartre has in mind here. All he is saying is that what's done cannot be undone. Not even God, according to the traditional notion, can now bring it about that Rome was not founded. Before the fact, he could have prevented Rome's founding, and after the fact he can destroy Rome. But that is not the same thing. Once Rome was founded, the deed was done. It was fixed for all eternity and cannot be undone. Forever after, it is (and will continue to be) true to say "Rome was founded."

Thus, a being's past is fixed, immutable. But these are infallible characteristics of the in-itself. A being's past, therefore, smacks of the in-itself. It is what we have called facticity.

On the other hand, the future is wide open. There are infinite possibilities there, as we saw in the analysis of possibility in the previous chapter.

It is in a being's future that it "transcends" its past; it is not confined to the past. It is not limited or "defined" by its past. A being with a future is a being that cannot be captured like that.

Thus, a being with a past and a future is a being that involves both a facticity and a transcendence. That is, it can only be a being-for-itself.

And the converse holds too. A being-for-itself is a being that is haunted by facticity and yet transcends it. Such a being will have a past and a future.

Put these two results together, and we can see that consciousness is the type of being — and the only type of being — that exists in a temporal fashion. The existence of time follows from the nature of consciousness as we examined it in the last chapter — as a combination of facticity and transcendence.

Long ago I told you that, for Sartre, it was probably best to think of consciousness as like an event. The being of consciousness was more like the being of a process, an event, than like the being of a thing, of a substance. Consciousness, we said, was always in flux, it was constantly changing, it flowed.

Well, now we see that this is how we account for the notion of the flow of time!

When we try to visualize time, we usually picture it in spatial terms — for instance, as a line, running from earlier to later. But that is to miss the whole essence of time; it makes time static. As in the diagram below, point A is earlier than point B, and that's all that can be said about it.



Time, on the other hand, flows; the future becomes the present, and then the past. Time involves change. But this is exactly the kind of being consciousness has, and nothing else does.

This notion of the “flow” of time cannot be handled very well mathematically. This is a point that has been raised repeatedly in the literature, and sometimes people have trouble seeing it. The point is made several times by Arthur Prior, for instance, in his writings on tense logic. And well before that, it was made by the British philosopher John McTaggart Ellis McTaggart (1866–1925). McTaggart distinguished two ways of “ordering” times, or events *in* time, according to what he called the “B”-series and the “A”-series.

According to the “B”-series, events are ordered as either “earlier than,” “simultaneous with” or “later than.” Mathematics has no problem dealing with that. But according to what he called the “A”-series, events are ordered as “past,” “present” or “future” — and *those* relations keep changing. It is perhaps an open question whether the “A”-series relations can be accounted for in terms of “B”-series relations (although the answer to the “open” question is generally suspected to be no.)

Bergson makes exactly this point in his *Introduction to Metaphysics*. Sartre’s view of time owes a lot to Bergson’s influential writings. (See Christopher Vaughan’s dissertation for some details.)

Sartre is here talking about McTaggart’s “A”-series, and claiming that it is inevitably grounded in a *conscious* being, a for-itself.

Two questions arise at this point:

- (1) Past and future have been accounted for in terms of facticity and transcendence. What about the present?
- (2) If only consciousness can have a past and a future, then how can we speak of the past of the world (= World History), or the future of the world?

With respect to the first question, Sartre treats the present a little differently. There is an ambiguity in the term ‘present’, and Sartre thinks this ambiguity is not just a coincidence.

On the one hand, ‘present’ can mean present as opposed to past or future (recall McTaggart’s “A”-series.). On the other hand, it can also mean present as opposed to absent.

There is a connection between these two meanings. A being that has a present is a being that is present to something, in the presence of something.

Now consciousness is exactly the kind of being that is present to something. We have already seen, in the preceding chapter, the notion of presence to self. But more important here, consciousness is always present to the world. Consciousness is always a kind of witness to the world. (Note: Sartre always puts it in terms of consciousness’s being present to the world, not the world’s being present to consciousness.)

This notion of being a witness is what Sartre pushes here. A being that has a present is exactly a being that is present to an object, that is a witness of something.

This “presence” to an object is just what we already know as intentionality. Consciousness, insofar as it is intentional — which is to say, positional — is present to the world. That is, it has a present.

So here is the final structure. There are two sides to consciousness: (1) positional consciousness of an object, and (2) non-positional self-consciousness.

The former is the intentional aspect of consciousness, and pertains to knowledge, to what I know. (We have seen this connection before, when I discussed the second of the three Sartrean ekstases, namely transcendence.)

The latter, the non-positional side, pertains to the being of consciousness, what consciousness is. (We saw this connection too in our discussion of transcendence.)

This is what Sartre means in the slogan he repeats often: “The being of consciousness is SELF-consciousness.”

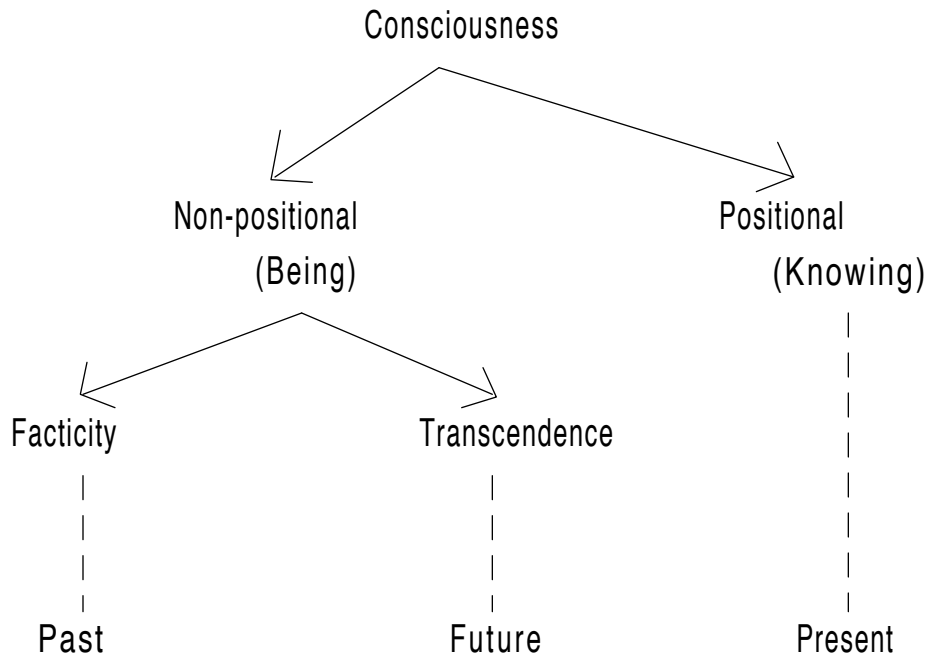
For example, look at the “Introduction,” p. 13 at the bottom:

Every conscious existence exists as consciousness of existing.

It is a matter of existing, not of knowing.

Now the positional side of consciousness is where we get the present. The non-positional side pertains to the being of consciousness. This being, we have seen in the previous chapter, involves both facticity and transcendence, and so both the past and the future.

Thus, we have:



Now what about our second question, the time of the world?

For Sartre, this a derivative notion. The world has a future, in a secondary sense, only because we have a future in the primary sense. So too for the past.

Similarly, recall how in the analysis of possibility, we said that the sky had the possibility of rain only because we who looked at it had our own possibilities.

We read the former in terms of the latter. The sky's possibilities are simply a disguised projection of our own possibilities.

What we have just talked about is basically what goes on in § 1 of Part II, Ch. 2, the “phenomenology” of time. I am going to skip over § 2 of the chapter, “The Ontology of Temporality.” Read it on your own.

(Note: There is a confusing misprint that occurs twice on p. 193. The volume has something that looks like a division sign: \div . It should be a non-identity sign (an “equals”-sign with a line through it): \neq .)

I do want to mention one odd doctrine Sartre discusses in this context. That is the theory that the past does not exist in any sense except with respect to people's minds. Thus, a person who dies and is completely forgotten is not just no longer existing; he never

existed. And that doesn't just mean that he never existed "as far as we can tell"; it means he never existed at all! (Assignment for homework: Figure out how this can be reconciled with the view that the past is fixed and unalterable.)

This is a peculiar view, but it is easy to see why Sartre holds it, given his account of how time in any event originates in consciousness. This theory is very nicely illustrated in his play *No Exit* (from 1944 — the year after *Being and Nothingness*). The three characters in that play wake up dead — literally. They are in hell, and the play is about their interaction with one another. But it is striking in the play how the characters are terribly worried about being forgotten by their survivors back here in the world. And the source of their worry is just that, once they are completely forgotten, they will not only be dead; they will never have existed — they will not even leave a legacy!

Thus, consciousness is the fundamental source of time. The paradoxes and puzzling features of time are just the paradoxes and puzzling features of consciousness all over again.

This then is the first of Sartre's three ekstases. It unfolds from the very being of consciousness.

The picture we end up with is a picture of the for-itself lurching toward its future (that is its transcendence), dragging its past behind it (that's its facticity.) I cannot do anything about my past, and yet I am responsible for in the sense that it is in my care. It is up to me what to do with it.

The past is in a sense growing ever larger, as more of my future slips into my past; the burden of my facticity grows ever heavier. But never mind; I keep transcending it. I am never confined to my facticity. I cannot ever really be in the manner of that being-in-itself that is my facticity.

Until the moment of death. At that precise moment, I almost succeed in "catching up" with myself. At that precise moment, my transcendence is almost over; time's up, there is nothing more to do. I am almost completely identified with my facticity, just as the waiter (in the Chapter on "Bad Faith") was trying to do. But, of course, then it's too late!

It's almost as if reality were designed as some kind of cruel practical joke, guaranteed to frustrate us most on the very verge of our success.

As we conclude our discussion of temporality, it is important to notice what has happened. When we first began this course, we talked about consciousness as though it were a matter of instantaneous flashes of consciousness. We spoke, for example, of acts of consciousness. Of course, if we do that, we are going to have the problem of tying these isolated acts together. What makes for the unity of consciousness over time?

Now we are in a position to see that our earlier way of viewing the matter is not right. Consciousness is not a string of instantaneous cogitos. It's a flow. The way to think of consciousness is not as a momentary flash, connected by only the most dubious of links to an earlier and a later flash. No, the way to think of consciousness is as an overall process. It takes time. In fact, insofar as the for-itself is the whole human being (and not just the

mind), we are now in a position to see that the proper way to think about the for-itself is as a whole human life.

This sort of thing often happens in Sartre. We revise our earlier ways of thinking about things in the light of later considerations. Our understanding deepens in this way.

Pure and Impure Reflection

You can pretty much skip over § 3 of the chapter on temporality. But there is one discussion there that I want to focus on briefly. It is the very important section on the difference between pure and impure reflection (pp. 211–237).

Indeed, this passage is the only head-on discussion I know of this important notion.

The topic is absolutely crucial for Sartre. As we have seen, reflection distorts its object. We talked about that as long ago as *Transcendence of the Ego*. And we saw Sartre say it explicitly on p. 121 of *Being and Nothingness*, in the Chapter on “The Immediate Structures of the For-Itself”:

It is often said that the act of reflection alters the fact of consciousness on which it is directed.

On the other hand, a large portion of the enterprise we have been engaging in in this course is a reflective one.

So our theory has reached the point where it is in danger of undercutting itself.

Sartre recognized this problem quite early — as early as *Transcendence of the Ego* — and there drew the distinction between a pure kind of reflection and an impure kind. The latter, is the one that distorts; the former, pure kind does not.

So everything now rests on this distinction. What exactly is the difference between pure and impure reflection? What can it be, given what he have done so far?

Let’s begin by asking what it is about impure reflection — the distorting kind — that makes it distort?

We discussed this briefly before. Basically, you will recall, it is because every act of consciousness makes up its own object by imposing form and structure on being-in-itself. As long as we are conscious of trees and tables and the like, this is fine. They after all just are formed or structured being-in-itself, no more and no less, so that there need be no distortion here.

On the other hand, if what I am conscious of is consciousness itself (as happens in reflection), then distortion inevitably occurs. What I am trying to think about is not in fact formed or structured being-in-itself. But that is the only way I can think of it. Hence, the distortion.

We have seen all this before, I have gone through it again because it now leads to a crucial point:

If distortion occurs whenever we try to think of consciousness as though it were structured being-in-itself, then the only way pure reflection can avoid distortion is by not thinking of consciousness in that way.

But Sartre's version of the theory of intentionality (see § 5 of the Introduction: "The Ontological Proof") holds that every act of consciousness structures and forms being-in-itself, and thinks of its objects in those terms exclusively.

Hence, and here is the astonishing conclusion: It appears that Pure reflection must VIOLATE the doctrine of intentionality!

I put the point this way because, if we now look at the discussion of reflection on pp. 211–237, we seem to be leading toward the same result. What seems to be happening, then, is that Sartre is revising and adjusting his views as he goes deeper and deeper into his theory.

Recall how when we discussed Part II of *Transcendence of the Ego* (on "The Constitution of the Ego"), we saw that the actual form the distortion takes there consists of, for instance, my momentary repugnance for Pierre's coming on to me as simply one manifestation of, one profile of, a longer-term state of hatred for Pierre, which in turn came on to me perhaps as an actualization of a more general disposition or quality of being hateful. And beyond all this, there appeared the Ego, which somehow generated everything else.

Now notice, all of this was put in terms of profiles, of promising more around in back, etc. If we recall the passage we discussed from *The Psychology of Imagination*, it is plain that the reflection going on in Part II of *Transcendence of the Ego* is much closer to perception and imagination (although imagination less so) than it is to conception. Such promises and profiles were not involved in conception.

When we were discussing this earlier, I said that if pure reflection, whatever it is, is going to avoid this kind of distortion, then that can happen in one of two ways:

- (1) Either the promises are still there, the object still comes on as referring to "more to come." But I am not fooled. I refuse to accept the promises. This sounds as if perhaps it is only a matter of performing an epoché, adopting the phenomenological reduction.

When I discussed that alternative, I quoted you a passage in which it seemed perhaps to be implied (*Transcendence*, p. 101, describing pure reflection — emphasis added):

Perhaps not without the ego, but as escaping the ego on all sides.

That's one possibility:

- (2) Or, alternatively, the promises are not still there. It's not just that I am not fooled by the distortion; it does not even occur. This would be like seeing the three faces of the cube as simply three surfaces, without viewing them as parts of any larger whole.

At the time, I said it wasn't clear which of these two alternatives Sartre had in mind. But now, by the time we get to the present section of *Being and Nothingness*, it is quite clear that it is alternative (2) he intends. Consider the following passages (pp. 218–219):

Actually the consciousness reflected-on [in pure reflection] is not presented yet as something *outside* reflection — that is, as a being on which one can “take a point of view,” in relation to which one can realize a withdrawal, increase or diminish the distance which separates one from it.

... It does not then detach itself completely from the reflected-on, and it cannot grasp the reflected-on “from a point of view.” Its knowledge is a totality; it is the lightning intuition without relief, without point of departure, and without point of arrival. Everything is given at once in a sort of absolute proximity. What we ordinarily call *knowing* [recall, knowledge = transcendence (the second ekstasis) = intentionality] supposes reliefs, levels, an order, a hierarchy. Even mathematical essences are revealed to us with an *orientation in relation to other truths*, to certain consequences.... At the same time, it [that is, pure reflection] is never surprised by itself; it does not *teach* us anything but only *posits*.... Reflection [again, pure reflection only] is recognition rather than knowledge [note again: knowledge = intentionality].

Now what Sartre is trying to do here is to do justice to the recognition that occurs in reflection, to the fact that there is big difference between thinking about myself and thinking about someone else, between thinking about my acts of consciousness and thinking about yours. There is an obvious sense in which I recognize my own thoughts as mine, and do not recognize yours in the same way.

But while that is what Sartre is doing, notice how he is doing it. As you listened to the passage, you should definitely have been reminded of that earlier passage we discussed from *The Psychology of Imagination*.

Pure reflection, he says, does not present its object “in profile,” or from a “point of view.” It is given “all at once” (that is, none of it remains that is only “promised”). It can teach us nothing. It cannot surprise us.

In short, pure reflection is here begin described exactly as conception was described in the passage from *The Psychology of Imagination*. When we talked about reflection just a moment ago, we said it was more like perception. So something new is happening in the present discussion.

At the same time, recall what Sartre's example was of conception in the passage from *The Psychology of Imagination*. It was the example of conceiving a cube — that is, the way a geometer conceives of a certain mathematical structure. But now he says in the passage just quoted:

Even mathematical essences are revealed to us with an orientation in relation to other truths, to certain consequences.

In short, Sartre seems to be doing some serious revising here. He is in effect saying that his earlier example about conceiving a cube was perhaps a bad example, and that perhaps even putting his point then in terms of conceiving was a bad choice of terminology.

The so called conceived cube in that earlier passage is not presented without “profiles” after all, it is not presented “all at once.” Some of the parts that were not presented “all at once” in the case of perception and imagination — the sides around in back, the corners that are not directly facing me — those, to be sure, are now presented all at once. But there is more involved than that. The notion of a “profile” need not be confined to this narrow visual analogy.

For example, as a geometer, I can conceive of a cube in such a way that I am thinking of all the sides and angles all at once. Nevertheless, I can be thinking of it as related in certain mathematical ways to certain other geometrical figures. I can think of it as the kind of figure for which certain theorems hold. I can think of it as the kind of figure for which such and such formulas apply to calculate the area and the volume. And, of course, I can be wrong about all of these. Perhaps when I work them out, those are not the formulas that apply, etc. In other words, even in the case of what Sartre had earlier called “conceiving the cube,” the object is still presented from a profile — still makes promises, and in fact still makes promises in a way that does not guarantee the promises will come true.

In short, it now appears as if Sartre's example of “conceiving” (and perhaps even the term ‘conception’) in *The Psychology of Imagination* is a bad one. It does not illustrate what Sartre is trying to illustrate there.

He now seems to be saying that if we want a real example of what he was talking about earlier — an example of a phenomenon presented with no profiles, all at once — we should take pure reflection as our example. And I suspect that is the only good example there is of this.

But there is more. On p. 218 again [emphasis added]:

[In pure reflection] the reflected-on is not wholly an object but a quasi-object for reflection.

Notice what he is saying here: In pure reflection, consciousness has no object, but only something he calls a “quasi-object,” whatever that is.

But I thought every act of consciousness had an object. That was the fundamental claim of the theory of intentionality.

So once again, it looks as if the notion of pure reflection violates the doctrine of intentionality.

But what about this “quasi-object”? What is that?

The context makes it clear that the problem with the so called “object” in the case of reflection is that it is identical with the act of consciousness that is reflecting on it. That, after all, is why reflection is a recognition of myself. There has got to be some sense in which the act that is doing the reflecting is identical with the act it is reflecting on, or else there would be no basis for the sudden sense of recognition we get.

But of course, the theory of intentionality, at least according to everything we have said so far, requires that the intentional object be quite distinct from the act of consciousness that is aware of it. This is true for any act of consciousness, reflective or not. The relation of intentionality is irreflexive, we said. It was this feature, after all, that was why people thought of the theory of intentionality as allowing us to break out of the confines of the Cartesian cogito.

Nevertheless, as we have just seen, in pure reflection the so called “object” of consciousness is identical with the act that is conscious of it. That is why Sartre says it is not really an object of consciousness, but only an odd quasi-object.

There is no escaping it. No matter how we look at it, it is clear that something funny is happening here with the notion of intentionality here.

We can either regard this as a hopeless muddle on Sartre’s part, a weakness in the theory, and perhaps a symptom that everything has fallen apart. Or we can regard it as a renegotiating of the doctrine and the notions involved.

We have already seen this kind of thing before. We can view the notion of facticity either as a violation of Sartre’s theory of radical human freedom, or else as a enriching and filling out of exactly what that human freedom is. So too, we can view the notion of the temporality of the for-itself as just an abandoning of the Cartesian cogito, which thought in terms of instantaneous, momentary acts, or else we can view it as a revision of that earlier picture and an enrichment of our way of thinking about consciousness.

Here are a few additional, undigested thoughts on this topic:

Sartre now seems to think pure reflection is the only mode of consciousness that presents its object (or “quasi-object,” as he now says) without profile, all at once. In particular, conceiving a cube does not do this.

Still, what bothers us about this notion of pure reflection? Why do we have the feeling something serious and unsettling is happening to our theory here, when we did not get that same feeling with the example of consciousness’s conceiving a cube? The case of conceiving a cube may not really be an example that does what Sartre originally said it

did. But we thought it did, and had no special problems dealing with this at the time. So why are we now bothered when Sartre says the same things about pure reflection?

On the one hand, Sartre says, pure reflection has no real “object.” Why does he say this? Well, in part (look again at the passages quoted above) it is because what pure reflection is aware of is presented without profiles, without making further promises. There is no “object” here in the sense that there is no objective claim being made. There is nothing that can be tested.

But if that is all that were involved, there would perhaps be no special problem — just as we had no special problems with the example of the cube.

On the other hand, another reason Sartre says pure reflection has no genuine “object” is that the only thing there that could serve as an “object” is not distinct from the act that is aware of it. Does this in any sense follow from the previous point, from the lack of profile?

Sartre talks as if it does follow. He speaks of “absolute proximity” in the case of pure reflection, where consciousness “cannot withdraw at a distance” from its object or quasi-object. This proximity, this inability to stand back, amounts to consciousness’s being identified with the object or quasi-object.

But does it really follow? In the case of conceiving the cube, we had no temptation to think the object was identical with the act just because (we thought then) it was presented without profiles, all at once. Is it that we were not tempted then because the example of the cube is just a bad example or because the identity of subject and object in pure reflection comes from another source altogether?

One might say that the identity of subject and object Sartre finds in pure reflection is his attempt to account for the fact that we recognize ourselves in reflection. And surely this is part of what Sartre has in mind.

On the other hand, the element of recognition is part of any reflection, pure or impure. The fact that I recognize the so called object as myself in pure reflection is because it is reflection, not because it is pure.

So if the element of recognition is why Sartre says subject and object are identical in pure reflection, then it is not after all just a matter of lack of profiles, a matter of “absolute proximity.” For in impure reflection that proximity is lacking, the “profiles” are very much present, as we saw in Part II of *Transcendence of the Ego* (“The Constitution of the Ego”).

Still, our dominant model of consciousness throughout this course (and Sartre’s dominant model too) is one of “taking a point of view” on something. That model does imply that where there is no point of view, no perspective, where the consciousness is that close to its object, it is identical with it.

Finally, note that Sartre regards pure reflection as a kind of ideal. In practice, reflection starts off as impure (p. 218 top):

Pure reflection, the simple presence of the reflective for-itself to the for-itself reflected-on, is at once the original form [not in the temporal sense but in some sort of structural sense] of reflection and its ideal form; it is that on whose foundation impure reflection appears, it is that also which is never first *given*; and it is that which must be won by a sort of katharsis.

Note that if pure reflection is an ideal, then it would seem we can never achieve it. Recall our discussion of ideal *values* in the chapter on “The Immediate Structures of the For-Itself.” But then, if we never achieve pure reflection, then all the real reflection we ever actually engage in is the impure kind, and *distorts* its objects. We are back where we started with this problem.

What *motivates* pure reflection? Perhaps, if it is an *ideal*, there need not be any special problem here. Perhaps we just automatically aim at that ideal, just as we automatically aim at being God. So too, just as we never succeed in being God, so too we will never succeed at a really *pure* reflection. But that doesn’t mean we can’t realize where we are going wrong.

It seems that what we have here is a little like Husserl’s recurring problem of how to be sure he is applying the phenomenological method correctly. The correct application of that method seems to be a kind of *ideal* goal Husserl is aiming at. He is never quite satisfied with what he has achieved, which is why he was constantly *starting over*, writing one “Introduction” to phenomenology after another. Is that more or less the same situation Sartre is in? If so, it means the project of *Being and Nothingness* could never on principle be completed. Sartre *must* keep revising and adjusting the theory, starting over again and again, reconsidering his most basic starting points.

The Existence of Others

For our purposes, you can skip Part II, Ch. 3: “Transcendence.” This, you recall, was the second of the three *ekstases*.

There is interesting stuff in that chapter. He talks there about knowledge. He talks a lot about *how* consciousness constitutes its objects. He does an analysis of *space* in § 2 of the chapter.

All of this is very interesting, but we are going to move along to Part III of the book, where we finally face squarely the problem of the existence of *other* people, the classical “problem of other minds.”

We have seen this problem several times in Sartre. In “Existentialism Is A Humanism,” you will recall, this problem was in effect the basis for the objection Sartre attributed to the Marxists, who charged that existentialism, since it started with the Cartesian *cogito*, was committed to being through and through *subjective*. It treated people in total isolation from one another, and in fact couldn’t account for the existence of other people at all.

Sartre responded that this accusation did not apply to his version of existentialism, although it might apply to others. But it was unclear in “Existentialism Is A Humanism” just how Sartre himself avoided the problem. For that we have to look back to his earlier *Being and Nothingness*.

We saw a second treatment of the problem — actually, chronologically his first treatment of it — in *Transcendence of the Ego*. There Sartre argued that I have just exactly the same grounds for saying you exist as I do for saying I exist. For me, you are objects, with all the tentativeness and risk entailed by objectivity. (Recall the sense in which the perceived cube involved an “objective” risk.) But for me, I too am just an object. My Ego is a transcendent object in the “world” just as much as yours are. That is why sometimes other people know me better than I know myself.

There is something quite unsatisfactory about this response, as we recognized at the time. Sartre recognizes it too. And now, near the beginning of Part III, he says it explicitly (p. 318):

Formerly I believed that I could escape solipsism by refuting Husserl’s concept of the existence of the Transcendental “Ego.” At that time I thought that since I had emptied my consciousness of its subject, nothing remained there which was privileged as compared to the Other. But actually although I am still persuaded that the hypothesis of a transcendental subject is useless and disastrous, abandoning it does not help one bit to solve the question of the existence of Others.

Let’s look more carefully at the nature of the problem. It is going to be a problem for anyone who stands in the Cartesian tradition and who tries to suspend judgment about everything that can possibly be doubted, anyone who tried to confine himself to what is directly given to consciousness.

Can I doubt that your minds are really there, hidden in your bodies? Of course I can. Are your minds directly given to my consciousness. No, only your bodies are — and even there, only one profile, one Abschattung is directly given to me, so that the whole business is very tentative and subject to error.

(Yes, I know I told you earlier that the for-itself is its body; there is no dualism in the for-itself between mind and body. But that is something that will come out only in the next chapter of the book. We won’t be saying much more about it.)

Let’s look at the problem the way Sartre does. He considers two kinds of classical, traditional views first, and then gives us his own. The two classical views are presented in Part III, Ch. 1, § 2: “The Reef of Solipsism.” They are: realism and idealism.

By idealism in this context, he means a theory very like George Berkeley’s (although he is clearly thinking more of Kant), a theory according to which what we think of as objects are mental constructs out of ideas, not things-in-themselves behind the ideas. (This was Kant’s view, remember. Kant also believed in a thing-in-itself, but that was not what

Kant regarded as the object of consciousness.) Husserl too belongs under this heading, although Sartre will give Husserl's theory a separate treatment later on in the chapter. Husserl, you will remember, regarded the object as the infinite series of appearances, and the essence of the object was the principle of that series. (Recall the "Introduction" to *Being and Nothingness*. For Sartre, of course, the object involves not just an infinite series of phenomena, but also being-in-itself, the being of the phenomena.)

By realism, Sartre clearly has in mind here Descartes' doctrine. Descartes, for all his methodological doubts and for all his adopting of what is in effect a phenomenological reduction, in the end had a representational theory of knowledge. For him, our ideas and thoughts were representations of external objects. And the big question for him, you will recall, was whether the ideas he had in his mind corresponded to things in reality, were accurate representations of them. In short, Descartes has a correspondence theory of truth and knowledge, not a coherence theory as we saw with Husserl and Sartre.

Descartes argued that, although he started by doubting the existence of those real "things- in-themselves" (as Kant would call them) outside his mind, he could in the end — as the result of a curious argument that need not detain us here — be sure of certain things about those external objects. He thought he could show too in the end that our ideas of external objects, or at least some of those ideas, were caused in us by the action of the external objects of which they were the ideas.

This is the "realistic" side of the doctrine. (The ideas come from something "real" out there.)

All right, given that, how does Descartes think we come to know, or even come to think, there are other minds? Well, listen to what he says in his *Meditations*. He speaks of (emphasis added):

... human beings passing on in the street below, as observed from a window. In this case I do not fail to say that I see the men themselves ... and yet what do I see from the window beyond hats and cloaks that might cover artificial machines, whose motions might be determined by springs? [Notice: He is in effect saying here that I see only a "profile."] But I JUDGE that there are human beings FROM THESE APPEARANCES, and thus I comprehend, by the faculty of judgment alone which is in the mind, what I believed I saw with my eyes.

In effect, what Descartes is saying is this: What I literally see are bodies, or more strictly only clothes that seem to cover bodies. But even supposing that the bodies are really there, I haven't yet got to the minds. I come to the conclusion that there really are minds in those bodies by a kind of analogy. In effect, I reason like this:

- (1) I observe that my body (presumably we have already solved Descartes' problem of determining that I really have one) behaves in certain ways that I soon come to correlate with certain states of

my consciousness. Wincing is associated with pain, tears with sadness, laughter with joy, etc.

- (2) I observe other bodies behaving in the same, or at least similar, ways.
- (3) And so I infer by a kind of analogy that those other bodies are endowed with a mind just as mine is. And, of course, I know and am quite infallibly certain about my mind.

Sartre's objection to this view is that it is not what it purports to be. It is not a realism, after all, but rather a version of idealism.

Thus Cartesian realism fails; it becomes idealism when it faces the question of the existence of others. Realism, if you push it, is forced to concede that our so called "knowledge" of others is only an ideal mental construct we fabricate as an "hypothesis" to explain the behavior of bodies.

Now, you may say, wait just a minute! This isn't "idealism" as we just described it earlier, or as the term is used elsewhere in *Being and Nothingness*.

Idealism does not just hold that our knowledge and our ideas of the object are constructs out of phenomena, or out of other ideas. It holds that the reality of the object is a construct out of phenomena or ideas. Idealism is not just a claim about how we come to know about other things, but also a claim about what kind of reality they have.

Now Descartes' argument by analogy is certainly not committed to holding that the reality of other minds is reducible to some construction I make up out of my own ideas or phenomena. On the contrary, the realist view holds that other minds are not just theoretical constructs, hypothetical entities I postulate to account for the observable data (as I hypothesize quarks, for instance, or gravitational fields). No, other minds are really out there (Descartes has a correspondence theory, after all), although our knowledge of them is gained only by a kind of hypothetical argument.

So Sartre's charge that Descartes' realism becomes an idealism when faced with the problem of other minds is entirely unfounded. He seems to have confused a theory about how we come to know something with a theory about what kind of thing that something is.

Of course, idealism "measures being by knowledge," for Sartre, so that if you are an idealist, you can make the transition quite easily from a theory of how we know something to a theory about what kind of reality it has.

But the question of course is: Is Descartes' theory an idealism on this point? It's not fair to say yes on the grounds that if he is an idealist already, then Descartes' argument will be equivalent to an idealist argument. That obviously begs the question.

Nevertheless, even though the charge that Descartes' realist view turns to idealism on this point may break down, still it is clear that Descartes' argument is not very persuasive.

First of all, arguments by analogy are notoriously weak and inconclusive.

But that is not the main problem here. Quite apart from the certainty we might or might not derive from our argument by analogy, the fact is that Descartes' view seems to distort the facts of the situation.

The view seems to say that first I have a clear idea of myself, and then by analogy I construct a notion of other minds. That is, that my knowledge of myself is somehow prior to my knowledge of others. But that's not right. (This much of what was said back in *Transcendence of the Ego* is right.) Sartre will argue — although this is getting ahead of the story a bit — that, on the contrary, the awareness I have of myself is already — at the outset — an awareness colored by an awareness of other people too.

So in the end, Descartes' theory does not seem adequate. Let us then turn to the other main classical attempt to explain other minds: idealism. And here we are basically talking about Kant.

The problem for Kant is not the correspondence of our ideas, including our ideas of other minds, with noumenal “things-in-themselves” out there, but rather the coherence of our ideas with one another. Kant, like Husserl and Sartre, has a coherence theory of truth.

The difficulty for Kant arises from the fact that our idea of the Other is a very odd idea indeed.

Consider once again the case of perceiving the cube. The three sides of the cube facing me refer to, promise, three more sides that I could see if I just turned the cube around. (Of course, if what I am perceiving is not really a cube, then that promise is a false one.) For Kant, as for Husserl (although not for Sartre, who does not “measure being by knowledge”), the cube just is the sum total of all its perspectives, all its profiles — both those it does and those it doesn't but could display to me.

Thus, what is promised in the case of perception is additional appearances that I can on principle get myself into a position to observe. In short, I can actually test the promises.

Now, when I experience another person — that is, when I view another body (rightly or wrongly, it makes no difference) as endowed with a consciousness — what I am experiencing (my phenomenon) refers once again to other phenomena. But this time — and this is the crucial point — the additional phenomena include ones I on principle can never get myself in a position to observe. They are the other person's private mental experiences.

Thus the experience of the Other is like perception and unlike imagination and conception insofar as it makes promises that are not guaranteed to be true. There is the risk of error. I can be wrong about what is going on in your minds, if anything. I can be wrong about whether you even have any. And in that sense, what we are dealing with here is like an ordinary case of perception.

But it is unlike perception — or at least unlike the ordinary cases of perception (perceiving the cube) — insofar as those risky promises can never on principle be tested by me. I can never be sure what you are thinking. And I can never be refuted.

But this means idealism is in trouble. Idealism says that I construct the world and everything in it, I constitute it — not just my knowledge of it but the very reality of it — out of my actual and possible phenomena, the ones I actually observe and the others that are referred to or promised. But the point of idealism is that, in every case, they are my phenomena; they all are phenomena from my point of view.

Indeed, how could consciousness constitute a world out of phenomena except from its own point of view? How could a movie be projected on a screen except from the “point of the view” of the camera?

And yet, in the experience of the Other, there are phenomena promised — and so they are part of the whole texture of my world — that are nevertheless not my phenomena but someone else’s.

It is perhaps hard to make this point clearly in terms of our movie theater analogy. But put it like this. Think of the movie theater, and the movie that always shows things from a certain angle, the “eye of the camera.”

The camera then always “sees” everything from its own point of view. There are promises, references to “more to come,” to be sure. In fact, a skillful director can lead the viewer to expect certain things to happen next — to expect the killer to be just around the corner in the hall. And such expectations may or may not be correct ones. But what the director cannot do is to lead the viewer to expect another movie entirely (as opposed to just a sequel to the present movie).

Idealism, faced with this problem, can make one of two moves, Sartre says:

- (1) Solipsism. But this alternative is unpalatable. As Sartre says, “it is opposed to our deepest inclinations.” (This is badly stated. There’s more to it than just saying that solipsism is so horrible we’d better not think about it. In fact, solipsism simply denies facts of our experience, Sartre thinks. We do experience the world as containing Other consciousnesses. Rightly or wrongly, we experience it that way. The main problem then for idealism is not that it cannot account for the truth of those experiences, but rather that it cannot account for the fact that he have them at all.)
- (2) Or idealism can say that the Other is not something we constitute after all. The Other, and our experience and knowledge of the Other is not something that comes from inside us but from outside. But that, of course, is just to return to realism all over again.

Hence, Sartre concludes, unless you are willing to accept solipsism, realism will lead you to idealism (although we've seen that part of what Sartre says doesn't seem to hold) and idealism will lead you to realism. We are obviously caught up in a hopeless vicious circle.

It is time to step back and take an overview of the situation. What is it that got us into this mess to begin with? Sartre considers this at the end of § 2 of the Chapter (pp. 312 ff.). There he argues that there is a fundamental assumption underlying both the realist and the idealist attempts at a solution here. The assumption is that my consciousness and the other consciousness I experience are related by an external negation.

What does this mean?

Well, first of all, how are they related by a negation at all? And, then, what is external about it?

They are related by a negation in the sense that I am not you. To say that A is not B is to relate them in a negative way. I and the Other are related in this way; we may be related in other ways too.

When we say that A is not B, we have an external negation provided that the negation does not originate in either the A or the B, and provided the negation does not in any way profoundly affect the A or the B. That's still too vague. Let me explain.

First of all, external negation is what we normally have in mind when we say that A is not B. Thus, the table is not the door. The table is just what it is, and so is the door. The table is just a table, and that's the end of that. Its not being the door is not a constitutive ingredient of the table. The negation here does not arise from the table. And so too for the door. Its not being the table is not a constitutive ingredient of it.

Likewise, the table's not being the door is not something that profoundly affects the table. Even if the door never existed, the table would remain exactly what it is.

Of course, it is certainly true that the table is not the door. And it is also true that this fact could not be changed — so that the table would be the door - without doing something pretty serious to the one or the other or both. But that's not what we're talking about here. The fact is, tables and doors just go their respective ways with complete disregard for one another. Their nonidentity is definitely a fact about them, but it is a fact that almost accidental to them. It doesn't get at their real core.

Sartre discusses the notion of external negation (and its correlative, which we shall see is called "internal negation") back in Part II, Ch. 3: "Transcendence." We skipped over this chapter by and large, but here in part is what he says on this topic (p. 243):

Actually we should distinguish two types of negation: external negation and internal negation. The first appears as a purely external bond established between two beings by a witness. When I say, for example, "A cup is not an inkwell," it is very evident that the foundation of this negation is neither in the cup nor in the inkwell. Both of these objects are what they are, and that is all. The negation stands as a categorical and

ideal connection which I establish between them without modifying them in any way whatsoever, without enriching them or impoverishing them with the slightest quality; they are not even ever so slightly grazed by this negative synthesis. As it serves neither to enrich them nor to constitute them, it remains strictly external.

On the other hand, consciousness is not what it is. This is a negation that does originate in consciousness itself, and profoundly affects it, as we have seen in the Chapters on “Bad Faith” and “The Immediate Structures of the For Itself.” Likewise, consciousness is not its objects; it always stands at a distance from its objects. But this is not some extraneous, accidental feature of consciousness. No, this not-being its objects is the deepest stuff of consciousness; it is what consciousness fundamentally is. In the end, the most profound way to say what an act of consciousness is is to say that it is not-being its object.

A negation like this is called an internal negation. To say that A is not B is to describe an internal negation if the negation arises from one or both of the terms, and profoundly affects their being. On p. 243, again, he says:

By an internal negation we understand such a relation between two things that the one which is denied to the other qualifies the other at the heart of its essence — by absence. The negation becomes then a bond of essential being since at least one of the beings on which it depends is such that it points toward the other, that it carries the other in its heart as an absence.

The notion of internal and external negation can be generalized, of course, to the notions of internal and external relations at large. You find this kind of talk of internal and external relations in Hegel, Marx, F. H. Bradley, etc.

Now, to get back to the point, the difficulty with both realism and idealism when it comes to the problem of other minds is that both theories assume that my consciousness and yours are related only externally. In particular, the fact that I am not you is only an external negation, according to these theories.

What’s wrong with that? Well, you will recall from Part I, Ch. 1: “The Origin of Negation,” that consciousness can be the only source of negativity, of nothingness. (Being in itself certainly can’t give rise to it; being-in-itself is purely positive.) The same point is elaborated in a manner more relevant to our present topic in the preceding chapter, Part II, Ch. 3: “Transcendence.” If the table is not the door, in the purely external way we have described, then it is consciousness that constitutes that negative fact. We have already seen Sartre say as much, in the passage I quoted you just a moment ago on external negation (p. 243 again). It is:

a purely external bond established between two beings by a witness.

He makes this same point on several other occasions too. Here he is again on pp. 255–256 (in the same chapter — underlined emphases added):

It remains to determine what type of being the external negation possesses since this comes to the world by the For-itself. We know that it does not belong to the *this*. This newspaper does not deny concerning itself that it is the table on which it is lying; for in that case the newspaper would be ecstatically outside itself and in the table which it denies, and its relation to the table would be an internal negation; it would thereby cease even to be in-itself and would become for-itself.

In other words, external negation always requires a witness.

But if this is so, then both realism and idealism are in trouble. They assume that my consciousness and yours are related by an purely external negation. But external negation requires a witness. So, whether they like it or not, there must be some third consciousness looking on, constituting you and me as externally distinct from one another.

Some philosophies will appeal to the notion of God here. It is God who constitutes you and me as what we are, and as externally distinct from one another. But this won't work, because God is himself another consciousness, another mind, so that we have gained nothing. Here is why we have gained nothing:

If God is related by an external negation to your consciousness and to mine, we'll need yet a fourth mind as yet a new witness, and so on. Either we will go on like this to infinity, in which case we never do ultimately account for all these external negations but just keep passing the buck, delaying the question yet another time. Or else we stop at some point and admit that some consciousness is related to the others by an internal negation, not by an external negation. But if we are driven to that conclusion in the end, why not just accept it at the outset: Two consciousnesses are related, not by an external negation, but rather by an internal one. There is no theoretical advantage to be had by delaying this inevitable conclusion.

So the result of this discussion in Part III, Ch. 1, § 2: "The Reef of Solipsism," is that we learn that the notion of "otherness" in the problem of "other" minds must be treated as an internal negation, not a merely an external one. Otherwise the problem will be insoluble, as our discussion of Descartes and Kant has shown.

But of course recognizing that we are here dealing with an internal negation is not by itself enough to solve the problem. We have uncovered a necessary ingredient of a solution, but not the whole story.

As a final remark on this § 2 of the Chapter, let's just note once again that the critique of idealism is much stronger than the critique of Cartesian realism. This is not surprising, since it is clear that Sartre thought much harder about the idealist approach.

After all, Sartre himself has a doctrine of constitution and something like a coherence theory of truth. His whole ontology and epistemology, although it is not an idealism (since he maintains that the being of the phenomenon, the underlying reality of it, is not consciousness but being-in-itself), is much closer to idealism than to Cartesian realism, which he never took very seriously.

Husserl

Let us see what else is required for a solution to the problem of other minds. We now turn to § 3 of the Chapter: “Husserl, Hegel, Heidegger,” beginning on p. 315.

These three philosophers all realized the conclusion we have just come to. But for Sartre, they still have not gone to the root of the problem. All three of these authors accept the conclusion that the relation between myself and the Other must be an internal negation. But they continue to think that my basic connection with the Other is one of knowledge. Recall from the “Introduction” that Sartre insists that “knowledge is not primary.” Sartre is going to invoke this here. He thinks the primary relation between myself and the Other is not one of knowing but one of being.

Let us look more closely. Consider Descartes again. He put the whole problem in terms of knowledge — which is to say, in terms of positional consciousness. When I look out the window, I see the hats and coats moving back and forth in the street below. Even if I assume that the hats and coats are not illusory, I still have to infer that there are bodies underneath them, not machines or hot air. And I have to infer further that the motions of those bodies are controlled by minds in a manner similar to the way my mind controls the motions of my body.

Notice how all this is put at the level of knowledge, of positional consciousness.

So too, “Kant” (that is, really, the “idealist” position) was concerned with consciousness’s constitution of the Other as part and parcel of its constitution of the world. The world is an object of positional consciousness, and we are related to it by knowledge.

Now, while Husserl, Hegel and Heidegger have made an advance over Descartes and Kant by realizing that you need an internal negation, they nevertheless retain this fundamental orientation in terms of knowledge.

Sartre considers Husserl first — out of chronological order, since Hegel antedated Husserl. But Sartre seems to think Hegel’s theory was more advanced than Husserl’s on this particular point. So he begins with the most primitive and unsatisfactory of the three theories.

Husserl discussed the problem of other minds in *Cartesian Meditations*, Meditation V, and lots of other places — many of which have only recently been published, and no doubt some of which haven’t been published at all. It is not clear to me just what exactly Sartre is thinking of in Husserl here. (The *Cartesian Meditations* was a series of lectures given in Paris in 1929, and published in 1931.) But never mind; it doesn’t matter for now. It might have been the discussion in *Cartesian Meditations*, although Husserl’s approach there doesn’t look very much like what Sartre describes.

In his treatment there, Husserl employs a notion of objectivity in the sense in which we speak of an “objective” science as one that is “the same for everyone.” In this sense, an “objective” fact is one that is the same for everyone — or at least would be the same for

everyone who was in a position to look on it. Thus, the fact that there is such and such a crater on the back side of the moon is an “objective” fact; in principle anyone could put himself in a position to verify it. The difficulties are merely technological ones.

By contrast, the subjective in this sense is what is not the “same for everyone.” Not everyone can verify it. If I have a pain, that is a subjective fact, not an objective fact — although if I grimace and groan, those outward manifestations of pain will be objective facts observable on principle by anyone.

This notion is not the same notion of objectivity we discussed earlier in Sartre’s *The Psychology of Imagination*, or in our discussion of “pure reflection,” which did not really have an object, you will recall, but only a “quasi-object.” In those contexts, “objectivity” was a matter of being presented in profile, not all at once, and of being tentative and risky, of making promises that are not guaranteed. In short, “objective” there meant “testable, with the outcome in doubt.”

This new sense of “objectivity” includes all that, but is narrower or stricter. It is not only testable with the outcome in doubt, but testable by everyone — with the same results.

This notion of objectivity goes back to Kant. In his *Prolegomena to Any Future Metaphysics*, he wanted to capture the notion of objectivity as used in the natural sciences, where not only is the notion of experiment important, but the notion of repeatable experiment — with everyone’s getting the same result.

Sartre would be perfectly willing to accept this notion of objectivity. It’s not one that he has emphasized previously, but we’re not going to quibble over terminology. It’s a perfectly coherent notion, and there’s nothing wrong with it.

So, we have this concept of “objectivity” as “the same for everyone,” not just “for me.” Note the “for everyone.” Those have to be “other minds.”

Here is the move as Sartre interprets it: Husserl thinks it is of the very nature of consciousness to be intentional, to be consciousness of an object. The relation between consciousness and its objects is not an external relation but an internal one; it profoundly affects consciousness. Consciousness is made of it.

Now of course the object of consciousness is the World, or some part of the World. It is objective. And objectivity, as we have just seen, involves by definition an appeal to other minds. Thus, consciousness is related by an internal relation of intentionality to an external, objective World the very objectivity of which appeals to other minds.

Thus, for Husserl, the presence of other minds is, as it were, built in to consciousness.

As it stands, this just looks like a plain fallacy of equivocation. We start with the notion of an intentional object, and then slip into the notion of scientific “objectivity.” In fact, Husserl’s discussion is not so sloppy as this. But Sartre does not go into details.

In any event, notice how the whole question continues to proceed in terms of knowledge, in terms of positional consciousness’s relation to its objects. (This much does seem to be true of Husserl’s treatment.)

Sartre thinks it won't work. First of all, there is the fallacy just noted. But that is not Sartre's own objection. He argues that this theory fails for basically the same reasons Kant's position wouldn't work. Both Kant and Husserl have the notion of a Transcendental Ego that constitutes its objects — including the whole objective world. If my Transcendental Ego constitutes a world that contains a reference to other Transcendental Egos, it must itself have built in that reference to other Transcendental Egos. But then we are back to Kant's problem. The Transcendental Ego constitutes things only from its own point of view. It could never constitute a world that contains references to other points of view that on principle it could never share.

Furthermore, Husserl defines being in terms of knowledge, as we have seen. He is an idealist. Recall the "Introduction" to *Being and Nothingness*, where Sartre describes how Husserl defines the being of a thing (for example, the perceived cube) in terms of an infinite series of phenomena, given and promised. It is all done in terms of phenomena — that is, in terms of objects of positional consciousness. Positional consciousness is primary.

Thus, a Transcendental Ego that belongs to someone else and that on principle cannot be a phenomenon for me, cannot by rights be a being for me. (I am not sure whether this same problem arises for Husserl in connection with my own Transcendental Ego. This is because I am not sure exactly how Husserl regards reflection.)

Husserl, therefore, cannot escape solipsism any more than Kant or Descartes could.

Hegel

Hegel, on the other hand, makes some progress on this question, which is why Sartre treats him after Husserl even though that is out of chronological order.

For Hegel, the Other is needed not just for the constitution of "the World," but for the make-up of consciousness itself.

When I am conscious of myself, I am conscious of myself as not the Other. I am me, and not you. It is only in terms of the Other that I am conscious of who I am — namely, not them. I only really get a grasp on who I am by knowing who I'm not.

Furthermore, and this is where Hegel makes progress over Husserl, this is not just a matter of knowledge for Hegel; it is also a matter of being. It's not just that I come to know who I am by contrasting myself with others; that's how I come to be who I am.

My relation to others is thus a kind of internal negation. I, insofar as I know myself, am profoundly affected by my not being anybody else; it is what makes me up.

Throughout this entire discussion, Sartre is almost certainly thinking of the famous Master/Slave passage from Hegel's *The Phenomenology of Spirit*.

For Hegel, the most fundamental element of the Slave's sense of himself is that he is not the master. This is how he comes to realize who he is. But that's not all. It's not just a matter of how the Slave comes to know who he is; it is also a matter of how he comes to be who he is. That is, it is what makes him a slave.

This passage in Hegel is probably the most influential text in all philosophy since Kant. It was the basis for Marx's analysis of the bourgeoisie and the proletariat. It was the foundation for Nietzsche's own discussion of Master morality and Slave morality in *The Genealogy of Morals*.. And here we see it influencing Sartre.

Nevertheless, Sartre thinks it is wrong. The problem is that, despite Hegel's disclaimers, everything is still really formulated in terms of knowledge. Despite what he says, Hegel is not really talking about how the Slave comes to be the Slave; he is not talking about what kind of being I am. He is talking about how I know who I am — but then THINKS that amounts to the same thing!

In short, Hegel is still proceeding at the level of reflection. He is talking about positional consciousness of the self. If we confuse the non-positional consciousness we have of ourselves (the kind of self-awareness that is what we are) with the reflective, positional consciousness we have of ourselves (the kind of self-awareness that is what we know of ourselves), then we will of course think that this reflection will accurately and exhaustively reveal what we really are. And that is idealism.

But if, with Sartre, we think reflection alters the consciousness reflected on, then the knowledge we have of ourselves will not exhaustively tell us about the kind of beings we are.

Without going into the details of Sartre's critique of Hegel here, we can nevertheless see right away a symptom that Hegel has gone wrong. He is giving us an explanation of how we know who we are. And he begins with the claim: I am I — and no one else.

That is, I am what I am, and not what I am not.

But we know better by now: I am not what I am, and I am what I am not.

Heidegger

Despite the fact that Sartre begins this section of the Chapter by saying that Husserl, Hegel and Heidegger all retain the idealist assumption that measures being by knowledge, it becomes clear in his actual discussion that this is not his main criticism of Heidegger. He thinks Heidegger has it mostly right. Heidegger puts the whole thing in terms of being, not of knowledge. (He also thinks that in the end Heidegger is as idealistic as Kant was; it's just that he doesn't start there.)

Heidegger begins with a kind of definition: He says that human reality (= Dasein = being-in-the-world) is being-with (= Mitsein.) The famous slogan is: Dasein ist Mitsein. This is the universal and necessary structure of human being.

Copyright © 1996 by Paul Vincent Spade. All rights reserved. Permission is hereby granted to copy this document in whole or in part for any purpose whatever, provided only that acknowledgment of copyright is given.

Notice: We are talking about being here. There is nothing said about knowledge. And to that extent, he thinks Heidegger is on the right track.

But Sartre thinks Heidegger's account is altogether too general. It won't allow us to account for the individual facts of our relations with others. Of course, any theoretical account is going to have to be in general terms; we just don't have time to fill in all the details for every individual human being, even if we knew them all. But that's not Sartre's criticism. The point is rather that Sartre thinks Heidegger's theory will positively get in the way of any such account, actually prevent it. (I must confess, I find this part of Sartre's discussion hard to see.)

But there's more than that. On Sartre's view, it is a brute fact — it is part of our facticity, there is no sufficient reason for it — that we encounter others in our experience. In short, it is not a matter of definition; it is not a necessary truth. It is a contingent fact. There could exist a human being without others, contrary to what Heidegger says. But in fact that's not the way it is. That fact is a contingent fact, not a necessary structure of human reality.

So in a sense, the point Sartre is making here is another form of his insistence that general principles are not explanatory devices. Even if it is generally true (and it is) that human existence is a communal existence, that Dasein ist Mitsein, that's no explanation of the particular facts; they are not grounded in that general principle. On the contrary, it is just the other way around: it is because all the particular facts of individual human existences are the way they are that the general principle holds at all.

Summary

Let us summarize the results of this § 3 of the chapter. For Sartre:

- (1) We cannot prove the existence of others, if by that we mean grounding it on general principles. This much is contrary to Heidegger.

This does not of course mean we can't be sure of the existence of others; it's just not a matter of proof. We can be sure of it. In fact, Sartre thinks we can be just as sure of it as we can be sure of our own existence. (And therefore the "problem of other minds" is solved. Recall how the problem was originally framed as a problem of the disparity between the certainty I can have of myself and the certainty I can have of others.)

We can be sure of it in exactly the same way too — by the cogito. The non-positional, pre-reflective awareness we have of ourselves is also (as a matter of fact, not of necessity), an awareness of others.

This non-positional awareness is of course an awareness of what we are, not of what we know. It is in this sense that Sartre thinks the problem of other minds is a problem of being, not of knowledge.

- (2) The fact that this is a matter of non-positional awareness means that the Other is not (at least not at first) given to me as an object — that would be a matter of positional consciousness. Our primary relation to others is thus not one of knowing, but of being. We encounter others; we do not constitute them. That would be a matter of positional consciousness again.

This much is contrary to the idealism of Husserl and Hegel. In short, Husserl's whole attempt in *Cartesian Meditations* to explain how we constitute others is hopeless. It's not so much that it gives the wrong answer, but rather it asks the wrong question. It's all put as if the main question were how we can know others exist.

- (3) Our relations to others is an internal relation, not an external one. It has to be an internal relation if it is going to be a matter of our being rather than our knowing.

This much is contrary to Descartes and Kant.

The Look

In effect, what we now have at the end of § 3 of the Chapter is an outline of the parameters any successful theory of other minds must fit. In § 4, Sartre begins to set out his own theory.

This section, entitled “The Look” (pp. 340–400) is one of the most famous sections of the book — probably second only to the chapter on “Bad Faith.”

In this section, Sartre gives a justly famous description (pp. 347–349) of a man peeping through a keyhole into a room. To begin with, the man is totally “absorbed” by what he sees in the room. He is on the non-reflective level. And he's all alone; he's not especially aware of the presence of other people. (If he's watching other people in the room, he is in any event not really regarding them as other consciousnesses.)

But now, all of a sudden, the man hears a footstep behind him, and he suddenly realizes he's being watched. He's been caught! Suddenly the whole situation changes radically for him. He's suddenly aware of himself as being seen!

The change, of course, is that now he's aware of the presence of someone else — another consciousness, who is watching him. In short, the difference between the two situations, before and after, is exactly the difference between an isolated consciousness, all by itself, and a consciousness in the presence of others — what Sartre calls being-for-others.

(Note: This is not some third kind of being, in addition to being-in-itself and being-for-itself. What Sartre is doing here is introducing a new element into his analysis of being-for-itself. He remains a dualist, not a triadist.)

Now it's crucial to understand that when the man suddenly realizes he is being watched, he does not necessarily shift to the reflective level. No doubt he soon will move to the reflective level, but what Sartre wants to focus on is that delicate moment after he realizes he is being watched but before he begins to reflect.

Now, you might say, how can this be? We said the man realizes he's being watched. Doesn't that realization involve a consciousness of a situation in which he himself is a component? And isn't that the definition of reflection that we've been dealing with as long ago as *Transcendence of the Ego*? Are we now going to tamper with our earlier definition of reflection?

No, not at all. We retain the notion of reflection we have been using all along. There is nothing funny going on there. In order to see what is going on, let us take another example. This one is Sartre's example of the man in the park (pp. 341–346).

This example is perhaps a little easier to see the point of, and moreover has as a kind of fringe benefit the fact that it explains why it is so difficult to look someone in the eye — the staredown.

Suppose you are in a park, minding your own business. Everything is fine; there are no special problems at the moment. A few paces off there is another person, sitting on a bench reading a paper and minding his own business too.

Everything is normal. Everything is just as we have described it up till now in *Being and Nothingness*. The whole world constituted by your consciousness, including that other human body, is arranged to refer to a particular point of view — your point of view. Everything refers to you; everything is organized around you — the eye of the camera that is always present but is never seen as a phenomenon on the screen.

In short, the whole situation is a matter of your phenomena, along with the promises of further phenomena that would also be yours if you did such and such. We are talking about your phenomena throughout.

But now, suddenly, that other man puts down the paper firmly, looks up and stares directly into your eye. You are startled; you become unnerved. Why?

It's because all of a sudden the world comes on to you differently. There is something threatening about this man's ominous stare.

It's not as though you're afraid he's going to attack you, or anything like that. Let's suppose the man is old and feeble, so that there's no question of any physical danger in the situation. Still, you continue to be unnerved by his stare. Why?

Well, Sartre says, it is not that he is threatening you with bodily harm. Rather, it's more serious than that. He's a threat to the order and arrangement of your whole world.

In the very fact of recognizing that there is another consciousness behind those alien eyes, you recognize that there is another point of view on things, a point of view that ON PRINCIPLE you can never occupy. All of a sudden, the world comes on to you as referring not just to your point of view, but to another one too — to another camera. The world is no longer just nicely ordered and arranged around you. It's now arranged around him.

Everything stays the same, of course. The trees are still the same color, the bench is still there. And yet it's profoundly different. And notice, there's nothing here that's reflective yet.

Everything is still the same, and yet something has dissolved. The world is now his world, a foreign world that no longer comes from you but from him. For example, the values that appear in the world are suddenly his values — values that you can never get in a position to see.

Furthermore, you suddenly recognize that he can see that peculiar vantage point that you are. In other world, he can see you. That peculiar, private point of view that is you — which you yourself always are but can never accurately see (the invisible camera) — can now be seen by that other man.

Once again, this still doesn't have to be reflective. It's as if, in your movie, another movie camera suddenly came up and stared into the lens of your movie camera. You're still not reflecting — that is, you still don't see your movie camera on the screen. But you are very definitely aware of being seen.

Does that other man approve of what you are doing in the park? Is he secretly condemning you? Does he find you ugly, awkward, out of place?

You can't tell! Your world is suddenly haunted by the Other's values, over which you have no control. There is another freedom loose in the world, a freedom that does violence to your own.

You suddenly realize that the other person can see you as an object. That peculiar vantage point that you occupy but can never see, that point of view that you try in bad faith to turn into an inert object, an in-itself for you — the Other succeeds in seeing all that as an object. You can never see yourself as others see you. The attempt to do so, the attempt to see yourself as an object, is bad faith.

But the Other sees you as an object. Thus you are for-others what you never succeed in being for-yourself. And so you are exposed; you are vulnerable.

You try to be noble, let us say, you try to be good. But you never make yourself noble or good, just like that. You can never define yourself in that way.

But the Other can do it to you. He decides whether you are noble or good. He passes judgment, projects his values on things — including you. He sees you as you really are.

Am I funny? Only if he thinks I am. Am I ugly? Only if he thinks I am.

Now you might well ask: Who is he to define who I am? Why should his evaluation of me affect me like that? Why should I accept his point of view?

Well, there's no good answer to that, I suppose. But the point is, I DO accept his point of view! I feel ashamed, for instance, or proud. And of course, those feelings by their very nature refer to others' values.

At one point, Sartre says it is impossible to be ashamed alone. And he's right! Of course, it's possible to do something secretly that you're ashamed of. But that's not what he is denying. That feeling of shame already puts you in touch with other people and their values, even if they don't happen to be right there on the spot.

Thus I recognize myself in the Other's judgments of me — even though I may not know what they are. His judgments cut me to the core. Why should his judgments be able to hurt me unless I recognized myself in them? And yet they are completely beyond my control.

Note once again: All this goes on at the pre-reflective level. I don't have to step back and think about myself in order to be unsettled and threatened in this way. In fact, if I do step back and reflect, that disarms the situation. I withdraw from the other person's look and turn to reflection. At that moment I might question his judgment, think he doesn't really know me, etc.

But, the fact remains, before I reflected, I felt his look, and I recognized myself there — not positionally (that would be reflection) but non-positionally.

Of course, as soon as we say 'non-positionally', we're talking about the being of consciousness, not its knowledge. And this is a symptom of the very basic way in which Sartre has recast the problem of other minds. For Sartre, the fundamental question is: how do I come into contact with other people? The question is not: how do I KNOW other people exist? It's not an epistemological question at all; it's an ontological one.

For Sartre, the fundamental way I come into contact with other minds is not by knowing they are out there, but by means of feelings of shame, pride, etc.

All my life I aim at defining myself, at becoming what I am, at being a definite in-itself. (Remember the desire to be God, from the Chapter on "Bad Faith.")

But all that is bad faith. I cannot succeed in it. I am forever separated from myself as I am, the real me. It is a goal I cannot reach.

And yet the Other does it for me, whether I like it or not, at one stroke, by a single glance. He makes me what I am. He defines me.

Yet, while that is surely me he defines (I recognize myself in his judgments, or else why would they bother me), I am still separated from myself. That is, I have no control over what he makes of me. I am still separated from myself as I am.

I can try to win the Other's approval by being friendly, smiling a lot, behaving in ways I think will win his approval. In other words, I can try to manipulate his freedom to get him to judge me in the way I would like.

But it's still his freedom. He may approve of my efforts, or he may regard them all as sycophantic, sneaky ways of trying to win his approval. I can't control which of these alternatives he will choose.

This is still a very difficult theory to come to terms with. Let's try to get a better handle on it by returning to a problem we discussed a short while ago: Why should the Other succeed where I fail? Why should the Other's "Look" define me any more than my own estimation of myself can succeed in defining me?

The situation here is a very complex one, and threatens to dissolve into one of two oversimplified cases. We might want to say one of the following (note: Neither of these is correct, as we shall see):

- (1) The Other's viewing me as an "object" is a distortion, much like the distortion that appears in impure reflection. It's the kind of distortion that inevitably occurs whenever we try to think about a for-itself. We discussed this when we were talking about *Transcendence of the Ego*. It is what Sartre meant when he said "Man is always a wizard to man." The Other, by viewing me as an "object" that is nevertheless conscious, is viewing me as a kind of magical object, and that of course is wrong.

On the contrary (we might go on), the fact of the matter is that I cannot be defined by the Other's look in this way. I am a story that is still being told, and the attempt to judge the story before it is finished is just premature. There is nothing yet to define.

In short, this response holds that the view of the for-itself that we have been talking about up to now is still correct, and the Other is simply in error if he thinks he can pin me down and define me so easily.

Or we could say:

- (2) It is just the other way around. The Other does have a kind of authority when it comes to defining me, to deciding what kind of person I am. And this is for the reason we gave a while ago: I recognize myself in his judgments of me — whatever they are. They stick!

But remember (this second response goes on), we are not talking about reflection here. So my recognition of myself in the Other's judgments is not a matter of how I think of myself. The recognition is not a matter of positional awareness of myself in reflection. It is a matter of non-positional awareness — that is, it is a matter not of what I know but of what I am.

We've said all this many time before. But notice what it means here. To say that the Other's judgments can affect what I am in this way is just to say that he's right!

In that case, it would seem to follow that the picture of the for-itself we have developed up to now in *Being and Nothingness* is wrong. I do have a definition after all. I am not an uncompleted project.

Those are two kinds of theories we are perhaps tempted to adopt. But neither of them is correct, according to Sartre. The problem is that both theories assume that each point of view excludes the other one — the view that the for-itself has no definition is incompatible with the view of the for-itself as defined by the Other.

For Sartre, the fact is that neither of these two points of view excludes the other one. Both are correct, and both must be taken into account.

What we have here is what Sartre calls a "metastable" situation. Recall the notion of a "metastable" situation from the chapter on "Bad Faith." A "metastable" situation is one that combines two polar opposites in an unsettled way so that they keep threatening to fly apart.

Bad faith itself was a "metastable" enterprise. Now we believe, now we don't — because we are making ourselves believe. Here, in the discussion of other minds, is another "metastable" notion: how properly to think of the kind of being the for-itself is.

A metastable situation combines opposites. It is contradictory. But it is not like the contradictory combinations we get in the notion of God, for instance, or of the unconscious, or of the Transcendental Ego, or of the "magical." Those combinations are not just contradictory but impossible and cannot exist. (How do we tell that? Eidetic abstraction shows us that things just cannot go together like that.)

But metastable situations do exist. Bad faith is real, contradictory or not. And so is this complex situation with other minds.

You may ask, couldn't I be mistaken about all this? And the answer is: Yes of course, if by that you mean you can't prove the existence of others. But you might as well say you could be mistaken about your own existence. You can't prove that either. If you tried to prove it, you would have to resort to the reflective level, in which case what you are talking about will be some distorted substitute for what you thought you were proving. But that doesn't mean that maybe you don't exist! It just means that the certainty of your own existence is not the kind of certainty you can prove; it's the kind of certainty you are. (I'll have a little more to say about that claim in a moment.)

So too with other people. I can be quite mistaken about what other people think of me; we get that sort of thing wrong all the time. And I can even be mistaken about where those other consciousnesses are located.

Sartre talks about soldiers crawling up a hill toward a farmhouse. They don't know whether it's occupied or not; it may be totally empty. But they feel the presence of

others. Those windows are like eyes. They are just as threatening, just as unnerving as the man's stare in the park.

Of course they can be mistaken about that. For that matter, perhaps the farmhouse is occupied by their allies, and there's been a horrible confusion. But that's not the point. The point is not to identify where those other consciousnesses are, or whether they are my friends or my enemies (to identify their values with respect to me). All those are questions of knowledge. But even if I get all that wrong, the fact remains: I am quite certain I am not alone in the world. And, just like the certainty of my own existence, this certainty is not something I can prove; it's a certainty I am. (Again, I'll say more about this in a moment.)

In short, my being as a for-itself is not just an isolated being; it is a social being. I am aware of that non-positionally, just as I am aware of myself non-positionally in all the other respects we have discussed in this course. This social being of mine is not a necessary structure. I could have existed even if no one else did. (This is contrary to Heidegger, recall.) But as a matter of quite certain fact, that's not the way it is.

Now of course this whole situation is complicated still further by the fact that, while the Other is looking at me, I am looking at him! The situation is mirrored from the other side.

What we are going to see here is another, very important instance of Sartre's general tendency to see things in stark alternatives.

We've run into this several times before:

- (1) The in-itself and the for-itself. Those are the only two regions of reality. There is no compromise, no middle ground. That would be God.
- (2) The "deterministic" world and the "magical" world in Sartre's discussion of the emotions. You are always in the one mode or the other, never both.
- (3) In his book *What Is Literature?* (we have not talked about this), Sartre sharply distinguishes poetry from prose. All writing is either the one or the other. There is no borderline case.
- (4) In fact, even the "metastable" notions we've just been talking about show this tendency toward stark alternatives. They all combine sharp contradictories. But there is no middle of the road, no reconciliation.

Well, here is another such stark alternative in Sartre's theory:

- (1) I can either regard the Other as an object, as one special object among many in my world. In that case, I am safe; I keep control of the situation, and the Other's threat is disarmed. Or

- (2) I can regard him as a subject, as a consciousness defining me by his look. In that case, I lose control of the situation; I become an object for him. My being-for-itself becomes a being-for-him — a being-for-others.

For Sartre, at any given moment I can be adopting the one or the other attitude toward the Other. But I can never combine them.

In short, there is a tension, a struggle between us. There is, so to speak, a staredown. Which one of us is going to be the one to define the world?

The one who turns his eye away first loses the contest. He has yielded. He becomes an object.

This little “eyeball to eyeball” confrontation is a miniature of interpersonal relations in general. My relations with other people are all of the nature of a staredown.

Even love is just a particularly devious form of this vicious struggle, just as the attempt to be sincere is a particularly insidious form of bad faith, as we saw in Part I, Ch. 2.

At the end of his excellent play *No Exit*, Sartre has the famous line “Hell is other people.” This is not a throwaway line. He means it!

It is perhaps worth recalling here once again Hegel’s analysis of the Master and the Slave in *The Phenomenology of Spirit*. I mentioned this before as a very influential passage on Sartre. Here we see that influence quite clearly.

Before we go on, let me turn back to a point I made a moment ago. I said that, for Sartre, we cannot prove the existence of others. That is true if by “proving” it we mean “grounding” is somehow on a kind of universal and necessary principle. For Sartre, other minds do exist, and we can be quite sure they exist. But they don’t necessarily exist. On the contrary, the existence of others is a thoroughly contingent fact. Solipsism is perfectly possible, for Sartre. It just isn’t true, and we can be quite certain it isn’t true. But it might have been true.

Furthermore, Others don’t exist because of some general law of principle — for example, Heidegger’s *Dasein ist Mitsein*. On the contrary, it’s just the other way around; it’s only because of the quite contingent fact that other for-itselfs exist that it’s true in general (although not necessarily) that *Dasein ist Mitsein*.

It’s in that sense, then, that we cannot prove the existence of others. But note: It in no way follows from this that we cannot prove the existence of others if by that we mean a perfectly sound and persuasive argument for it on the basis of true (although contingently true) and certain premises.

Now, you may object on the basis of what I said earlier: To make it a matter of proof like this is to make it a matter of knowledge, isn’t it, a matter of positional consciousness? And I thought you said our relation to others is not primarily one of knowledge but of

being. Answer: That's true. It's not primarily one of knowledge but of being, but that doesn't mean we can't then go on to know it too.

So let me now try to give you a Sartrean proof of the existence of others. As good phenomenologists, of course, we shouldn't be *relying* on arguments and proofs. But we're not doing that.

All right, here we go:

Go back and recall Sartre's objection against the "idealistic" theory we associated with Kant. (The association with Kant is really just my bookkeeping device to distinguish "idealism" in the sense of § 2 of the chapter, "The Reef of Solipsism," from "idealism" in the sense of § 3 of the chapter, "Husserl, Hegel, Heidegger." Sartre himself doesn't link this up with Kant explicitly. In § 2, it was a matter of putting the problem in terms of *constituting* others; in § 3 it was rather a matter of measuring being by knowledge. There was some overlap, to be sure.)

In any case, the problem for "idealism" in § 2 was: How can consciousness — whether presided over by a Transcendental Ego or not — *constitute* a world with other people in it? Everything a consciousness constitutes is *inevitably* constituted from "its own point of view"; there is no other way to do it. But of course, a world with other *people* in it is a world that involves *other* points of view. How can consciousness project on the screen any *references* to any other film but its own, to any other vantage points but the eye of its own camera? And the point of Sartre's objection is just: *you can't do that*. This step of the argument seems to be a *necessary* premise, on Sartre's theory.

Now think about the example of the man at the keyhole, or the man in the park. In those cases, or in cases of *shame* or *pride*, the world we experience nevertheless *does* include those references to others. That's the whole point of those examples: in actual fact, our world *does* contain these references to alien points of view. This step is a *contingent* premise, but one each of us can verify for himself or herself.

How did those references get there in my world? How did my world come to include those *unguaranteed* and *untestable* promises to others? *I* didn't put them there (that's premise 1). And they surely *are* there (that's premise 2). Therefore: Someone else put them there.

That's the argument. *All* structure comes from the constituting activity of consciousness. But there's *some* structure in my world that *I* can't constitute; therefore, some other consciousness must constitute it.

So stated, the argument seems ironclad.

Now let me make a few other observations.

First of all, notice the sense in which this is *not* a matter of *positional* consciousness of the other. Take the case of the man at the keyhole first. In the first instance, before the man begins to react, he does *not* experience another consciousness. What he experiences is a *promise* of, a *reference* to, another consciousness. He doesn't even experience a

profile of, an *Abschattung* of, another consciousness, in the way I, for example, experience a profile or *Abschattung* of the cube. In the case of the cube, I perceive *part* of the whole, and the references and promises are to the remainder. In the case of another consciousness, I don't experience *part* of that consciousness, together with a reference and promise to the remainder. I experience *the world*, together with references and promises to something that is not *in* the world at all, but — like me — is an onlooker, a witness.

Now you say: but what about the case of the man in the park? There, at least, when the man looks up and stares me in the eye, I *see* him as an object of *positional* consciousness.

But not necessarily. What I see is a world that refers to another consciousness. A conspicuous part of that world at the moment is this other organism, this other body. And I take it that the other consciousness is located in that body.

But for Sartre, that's too quick. While I can be absolutely sure there is another consciousness besides my own loose in the world, and perhaps more than one, I can *never* be sure *where* they are. Recall the example of the soldiers crawling up the hill toward a house.

The point is: When I suddenly experience the world as referring to "*another point of view*" than my own, I don't simply mean another *spatial* point of view. After all, the argument against "Kantian" idealism rested on there being references to points of view I can *never on principle make my own*. But of course different *spatial* points of view aren't like that. I can always just *walk over there*, go up to the farmhouse and look out those windows, *move the camera*, etc.

Rather, we're talking about a much *richer* notion of "point of view" that involves things I can *never* have access to: your *values*, your *hopes*, your deepest *intentions*. In short, what Sartre sometimes calls your "*interiority*."

Concrete Relations with Others

Let's go on. I've already given you Sartre's line, "Hell is other people." That is, our fundamental relations with others are always relations of *conflict*. It is a question of who is ultimately going to "control" the "perspectives" in the world, you or me.

Furthermore, this struggle (as you might expect by now) is one no one really ever wins. Defeat lies in victory, in much the same way that belief never really succeeds in believing.

Let's look at this more closely. Why can't anyone ever really win the staredown, the conflict in interpersonal relations?

This brings us to Part III, Ch. 3: "Concrete Relations with Others."

The struggle can never really be won because each side wants contradictory things. Take my side, for instance. What do I really want out of my encounter with you? I want two things, and I will never have them both:

- (1) On the one hand, I want what only the Other can give me. The Other is the only one who can tell me who I am, what I am. Hegel was right here (although for the wrong reasons, as we discussed earlier).

The Other is the only one who can take the point of view on me that I would like to take on myself, but cannot. He is the only one who sees me as I am.

And that is something I very much want to preserve. I need his recognition. I want him to tell me “Yes, you really are the kind of person you have been aiming at being all along.” “You’re OK.”

And this is not some trivial, short-term benefit we are talking about. The kind of reassurance, the kind of ultimate justification (in almost the theological sense) we are talking about here is the most profound value in my whole world. It is the whole goal of my life, of my striving to be God. And the Other person is the only one who can in any sense give it to me. (Of course, even the Other cannot give it to me in any ultimately satisfactory way. I will never reach that goal.)

So I need the Other’s recognition. In a sense, we use other people like mirrors; they tell us what we look like.

In *No Exit*, the three characters wake up in hell. And hell is a drawing room. One of the significant features of the room is that there are no mirrors in it. At one point in the play, one of the two women asks the other “Is my lipstick on straight?” In the context of the play, it is a tense moment. She needs the other woman to tell her she looks all right. Will the other woman tell her the truth? Will she even answer the question? There’s no way to control that. Yet that recognition is not something you can just do without.

What this means is that in my struggle with the Other, I could not win by just killing the Other. In the case of the man in the park, I cannot just club the poor man to death. That would remove his threat, to be sure. But I would also lose what he has to offer me. And that’s something I want very much. I would lose my only chance at getting the most important thing in my life.

So, on the one hand, I want the Other to be a consciousness that can give me that recognition. And of course he can’t be a consciousness without being free.

So one part of what I want from him requires that the Other be free.

- (2) And yet, while I want what the Other can give me, I don’t want him to be free to give it or to refuse it. I want to be in control. His freedom is a threat to mine, as we saw with the man in the park. So

what I try to do is to preserve his freedom, but capture it, ensnare it. I try to make the Other view me as I want to be viewed.

Again, this is not some idle wish. It is fundamentally important to me. I not only want him to judge me; I want the whole thing to come out right! This is my one chance. So I want to make sure it turns out the right way. I don't want him to be free to do it some other way.

So I want incompatible things. I want the Other to be a free CONSCIOUSNESS — since only then can he take that objective view of me. But I don't want him to be a FREE consciousness. I want to rule his freedom.

This is why the struggle cannot ever be finally won. And yet it is an inevitable struggle.

For Sartre, there are two basic ways in which this social drama is played out, two fundamental patterns, although the details of course are as varied as people are. (Note: Once again, we find Sartre's tendency to think in terms of rigid dichotomies, stark alternatives.)

- (1) First way: I can try to get the Other freely to deny his own freedom. This is futile, of course, since by denying it he would be exercising it. Or
- (2) Second way: I can try to force the Other to affirm his own freedom. But this is futile too, since if he is forced, he isn't free.

Note: It is perhaps appropriate to recall here the analysis of the notion of distance Sartre gave in Part I, Ch. 1. The distance between Bloomington and Indianapolis, for example, could be thought of as the positive road, terminated by the two cities at either end. Or we could think of the two cities as the positive things here, and the road as what negatively keeps them apart. This kind of Gestalt flip-flop situation is very much like the kind of thing we see here in the two possible patterns of our relations with others.

In each case, I am aiming at the Other as a free consciousness capable of recognizing what I want him to recognize in me, and also as an unfree thing (an object), subject to my freedom. Of course, such a thing would be a for-itself (since it is a consciousness) and also an in-itself (since it is not free).

We've seen that kind of combination before. In the end, what I want the Other to be is God. And what I want from him is redemption, justification.

The two approaches are alike in that each gives with the one hand what it takes away with the other. The differ in that what the one way starts off with and ends up denying, the other way starts off without and ends up affirming.

This is still all very general, of course. Let's descend to some particular cases. The general patterns above are played out in an infinite variety of ways.

Examples of the First Approach

Take the first approach (getting the Other freely to deny his own freedom). Sartre describes how this may take the form of hatred, sadism, even (curiously) of indifference towards others.

In all these cases I try to capture the Other's freedom by the direct approach, by brute force. (Don't object that force was characteristic of the second pattern, not the first. Watch how it works out.) I treat the other person as an object. I abuse him, humiliate him, torture him if I can. All of which is a way of making the Other "just" an object for me. I treat him as a thing to be manipulated, subject to my whim.

On the other hand, it is not enough for me to treat him that way. I must try to get him to treat himself that way too. It is not enough for me to humiliate the Other; he must humiliate himself.

Take the extreme case of torture or sadism. Sartre was especially interested in this, since during the War he was concerned with the question of how long one could hold out under torture.

I am a sadistic tyrant, let's say, who torture my enemies in order to establish my power and authority. It's not a question of getting information out of them; it's a question of proving who's boss. If they resist, I'll just torture them all the more. I'll make them change their minds! I'll force them to recognize who they're dealing with.

But that is the key. They have to change their minds, to make a decision. They have to recognize my superior authority. The choice is theirs — otherwise, it's not really good enough. They are free to withhold their recognition.

It is not enough for me to force my subjects to bow and grovel and all that. That's only a matter of externals. That is just show. They must genuinely give me recognition. And that recognition is what I cannot compel.

When the poor man being tortured looks up and looks his torturer in the eye, the torturer fails. He has not succeeded in capturing that freedom after all.

Even if the tortured man "breaks" and does recognize me the way I want him to, he breaks because he chose that moment not to endure any longer. It was not my choice. He must freely deny his own freedom — which is hopeless.

Notice how, when you think about this situation, it continually threatens to dissolve into an example of the second pattern, of trying to force the Other to affirm his freedom. You can look at the situation according to the one pattern, and then, just by a little shift of mental focus, see it as exhibiting the other pattern. Once again, this is very much like the earlier analysis of distance, and shows that we are dealing with one of those "metastable" notions.

The case of indifference towards others is only slightly more complicated. If I adopt an attitude of indifference toward other people, I don't care about them; they are just objects; they don't get in my way.

But indifference is never really indifference, just as belief is never really belief. Indifference is a posture we adopt, the posture of not caring what others think of us. But of course the very fact that we take such great care to adopt the attitude of indifference betrays us. We are so careful and studious about it because we care very much what others think of us. We want them to think of us as indifferent, as "invulnerable" to their "Look," to their judgments.

But of course that won't work. What if one of them doesn't think of me as so invulnerable after all? What if one of them finds himself in a jam and comes to me to ask for my help? What if he says, "Spade puts on a hard exterior, but I don't care. I'm going to ask anyway." He refuses to recognize me as indifferent.

On the other hand, suppose he accepts my indifference. Suppose he says, "Spade's such a crusty old crank, he doesn't care about anyone. I'll go ask someone else for help."

That's his decision. I can't compel it. If he wants to think of himself as a mere object in his relation to me, that's fine. But I cannot force him. He must freely deny his status as a "subject" for me, and allow himself to be just an object that doesn't matter to me.

Examples of the Second Approach

On the other hand, let's look at the second pattern (where I try to force the Other to affirm his own freedom). Again, these examples can all dissolve into examples of the other pattern, if you look at them the right way.

This second pattern may take the forms, for instance, of masochism and (curiously) of love.

On this approach, I play the game just the opposite. Instead of trying to demean the Other, make an object of him (although a free object, as though that were possible), I now try to demean myself, and force the other person to adopt toward me the attitude that I am a mere object. I try to force him to choose to view me in that way. That of course would be a forced choice, which is impossible all over again.

For instance, take the case of masochism. In masochism, I treat myself as an object. I allow myself to be humiliated and abused. But, in order to succeed, the Other has to humiliate and abuse me. And suppose he doesn't want to! I want him to assert his freedom over me; in fact, I want to make him assert his freedom. But of course if I really could make him do that, he wouldn't be free and it wouldn't work.

Take also the interesting case of love. John loves Mary, let's say. He loves her so much that he is willing to anything for her. He is willing to obey her slightest command. In short, he is willing to be her slave, to be just an object for her.

But Mary has other ideas. She's not particularly interested in John. She notices him, and all that. (She can hardly not notice him; John sees to that.) But she doesn't care to have John jump at her slightest whim. His slavishness, his politeness, his attentions — all that offends her. It's just too much! In short, John cannot force her to play along with his game.

But he tries. He gives her gifts, he slips anonymous love notes under her door at night (taking care that his handwriting gives him away). He serenades her under her balcony at night. He calls her up. In short, he is trying to impose himself on her consciousness, to become the most important thing in her world. He is trying to force her to recognize him.

What will happen? Well, Mary is free; she may adopt any one of a number of attitudes toward John. She may be completely put off by all this fuss. If so, John fails.

She may decide this is a pretty good deal, and take John for all he is worth. She may take advantage of his masochism. But if she does, that is her free choice. John has not forced her, and that is what he wanted.

Or worse (and this is a variation), she may decide she thinks all John's protestations of love are very sweet, and that John himself is just cute as a button, and fall in love with him. Then what happens? Then she, out of love, is willing to do anything for John, to obey his slightest command, to see to his slightest whim.

In that case, John's whole plan has backfired. John was trying to force Mary to assert her freedom over him; he tried this by playing the role of a mere object — he is so "unworthy," and all that. But instead, Mary ends up adopting the role of a mere object with respect to him.

Sartre thinks all human interrelations are just so many variations on these two basic themes. It is a hopeless situation, a struggle without any possible resolution. As he says at the end of the book, "Man is a useless passion."

Existential Psychoanalysis

Let us skip ahead now to Part IV, Ch. 2, § 1: "Existential Psychoanalysis." The rest of Part IV is excellent, but you should now be in a position to read it on your own.

Sartre thinks his theory of consciousness, as developed in *Being and Nothingness*, has certain obvious implications for psychotherapy. And so, toward the end of his book, he sketches a theory of "Existential Psychoanalysis." It is important to recognize that this is only a sketch. Sartre certainly had no clinical training, and doesn't pretend to be filling out a complete theory here. As he says, existential psychoanalysis is still awaiting its Freud.

In this section, we get — once again — Sartre's explicitly contrasting his own theory of the mind with Freud's. We have seen this already in *Transcendence of the Ego*, in the chapter on "Bad Faith," and in *The Emotions*.

Copyright © 1996 by Paul Vincent Spade. All rights reserved. Permission is hereby granted to copy this document in whole or in part for any purpose whatever, provided only that acknowledgment of copyright is given.

Sartre thinks a great deal of traditional psychology, and most of standard Freudian psychoanalytic theory, is subject to a fundamental methodological criticism. They try to understand the individual in terms of general principles. And Sartre thinks this is wrong.

This is part of Sartre's overall reaction against what you might call "essentialism" or "universalism," which I pointed out very early in this course. It is part of his reaction against Hegel's explaining the concrete in terms of the abstract and the general, against Husserl's use of the eidetic reduction (although Sartre does not reject the eidetic reduction itself) and his emphasis on universal essences. One of the main themes of Sartre's version of existentialism (and lots of others too) is to reverse the traditional ordering of general to particular, and to exalt the individual.

Sartre thinks this traditional mistake can be illustrated by many biographies written from a psychological viewpoint. He quotes from a biography of Flaubert (p. 713 — we discussed this passage earlier, but I'll quote it again anyway):

... A critic, for example, wishing to explain the "psychology" of Flaubert, will write that he "appeared in his early youth to know as his normal state, a continual exaltation resulting from the twofold feeling of his grandiose ambition and his invincible power The effervescence of his young blood was *then* turned into literary passion as happens about the eighteenth year in precocious souls who find in the energy of style or the intensities of fiction some way of escaping from the need of violent action or of intense feeling, which torments them.

There are two problems with this for Sartre:

- (1) It reduces Flaubert to the intersection of general principles ("as happens about the eighteenth year"). It misses all the uniqueness of Flaubert.

As Sartre says (p. 714):

Why did ambition and the feeling of his power produce in Flaubert *exaltation* rather than tranquil waiting or gloomy impatience? Why did this exaltation express itself specifically in the need to act violently and feel intensely? Or rather why does this need make a sudden appearance by spontaneous generation at the end of the paragraph? And why does this need instead of seeking to appease itself in acts of violence, by amorous adventures, or in debauch, choose precisely to satisfy itself symbolically? And why does Flaubert turn to writing rather than to painting or music for this symbolic satisfaction; he could just as well not resort to the artistic field at all (there is also mysticism, for example). "I could have been a great actor," wrote Flaubert somewhere. Why did he not try to be one? In a word, we have understood nothing....

You cannot exhaust the individual by generalities in this way. We have seen this point before. But there is a second criticism too:

(2) This approach has to stop at arbitrary “givens.”

What does this mean? Well, let us take an example (one Sartre himself uses). Pierre wants to go rowing this afternoon, on this particular stream. But why? This is the fact we are trying to explain. Here we go, then:

He wants to go rowing this afternoon on this stream, because he is in general fond of rowing; it's one of his favorite sports. But why does he like rowing? Because he likes open air sports in general. And why is that? Because he likes to be outdoors under any circumstances. And on and on.

Here the particular is supposed to be explained in terms of the general. But where do we stop? We keep adding more and broader categories and principles to the story. If we never stop, if each step of the explanation is to be explained further by a higher, even more general step, then we have obviously not succeeded in finally explaining Pierre's present desire to go rowing on this stream this afternoon. Each step just delays the final explanation to the next step.

On the other hand, if we do stop at some point — say, at the point of Pierre's liking to be outdoors under any circumstances — and say “There, this is the end of the story, there is no further explanation to be had,” then why stop there? Why not go one step further? The explanation, as it stands, leaves this fact as just a brute given.

Of course Sartre agrees that we are in the end going to have to accept some things as just brute givens like this. Recall his slogan “Being is” that affirmed a violation of the Principle of Sufficient Reason. So his criticism here is not that we have a brute given, but that the brute given is an arbitrary one.

For Sartre, we should accept as a “brute given” only what really is a brute given. That is, we should stop our attempts at explaining further and further only when we have reached a point where the question “Why stop there rather than somewhere else?” is no longer an appropriate question.

For Sartre, this brute, irreducible fact cannot be found in general explanatory principles but rather in what he calls the individual's original project. And what is that? The original project is what the individual is fundamentally trying to make of himself.

The original project is Sartre's answer to Freud's notion of a complex. Let us look more carefully.

Sartre thinks classical Freudian analysis is not entirely in agreement with itself. The clinical practice is fine. Sartre has no quarrel with Freudian practice. Freud, after all, could actually cure people, and there's no denying that. But, for Sartre, the Freudians' theory of what they are doing is wrong. Of course this means that Sartre thinks the Freudian theory doesn't really fit its own practice.

We can perhaps see the point by asking why the psychoanalytic patient needs the analyst. Why can't the patient cure himself? (At least not usually. Freud, after all, psychoanalyzed himself.) What is the psychoanalyst trying to do?

Well, the Freudians say they are trying to bring to consciousness the deep-seated complex responsible for the patient's behavior, so he can understand it and deal with it.

Sartre, of course, thinks that whatever there is in the patient's mind is already conscious, even if only non-positionally. There is no unconscious in the Freudian sense. There is nothing unconscious about consciousness.

(You may want to refer back to our discussion of *The Emotions* to review how the Sartrean distinction between positional and non-positional consciousness is not just the same thing as the Freudian distinction between conscious and unconscious. The difference between Sartre and Freud here is not just terminological.)

Furthermore, Sartre continues, the patient might even know what the particular fact is that is involved — what is going on, as the Freudians would put it, in the “unconscious.” The patient might very well know this in the sense that he is already conscious of it in a positional way, that he has made whatever it is an object of consciousness, reflected on it. After all, for Sartre, any act of consciousness can be reflected on. There is nothing especially difficult about reflection, and we certainly don't need an analyst to help us here.

For example, I might know quite well that certain actions of mine are motivated, say, by a self-destructive urge. That's just a matter of reflection. I might not want to face that fact about myself, since it may be unpleasant. But there's nothing especially hard about doing it.

But what the patient cannot do by himself — and what he typically needs the help of a trained analyst for — is to know what it is that is motivating him, in the sense of knowing the meaning of the conscious act on which he is reflecting. He doesn't know how to decipher its signification, to use the terminology of *The Emotions*. It's like not knowing how to read the perceived cube we have discussed so many times. It's like not being sure what you should think of as “promised” and what not.

For example, I may realize that some particular action of mine is motivated by a self-destructive urge. By why is that there? What does it mean for what I am likely to do tomorrow? Is there a common thread between this self-destructive urge and certain moments of profound joy I experience, for instance? That kind of “reading the runes” is what the patient needs the analyst's help for.

This meaning or signification, the whole of which my self-destructive urge is but one facet, one profile, is the Sartrean analogue of the Freudian “complex.”

And it is what Sartre calls the “original project.” The original project is the ultimate project or goal that is the transcendent meaning of all the patient's acts, the original plan that amounts to what the person is trying to make of himself. Recall Sartre's famous phrase from “Existentialism Is A Humanism”: “Man makes himself.”

Copyright © 1996 by Paul Vincent Spade. All rights reserved. Permission is hereby granted to copy this document in whole or in part for any purpose whatever, provided only that acknowledgment of copyright is given.

The “original project” amounts to something like the old theological notion of providence — God’s plan for me. But, on Sartre’s theory, this plan does not come from God; I do it. And it is not fixed. I can change it (although that is as rare as what we call a religious “conversion” experience).

Just as the whole cube is “signified,” “meant,” “promised” in each perspective view, so too the original project is “signified,” “meant,” “promised” in each action of mine.

Psychoanalysis, then, aims at something a little like an eidetic abstraction — but not exactly. It is like an eidetic abstraction in that it is the “meaning” of all the profiles, all the perspectives. But it is not like eidetic abstraction in that it is not just a matter of general principles. It is very particular: The question is “What am I up to?”

So this is what the patient needs the analyst to help him with — to understand his original project. How is this done in practice? Well, the way to do it, Sartre says, seems to be a matter of comparing the various particular bits of behavior — your dreams, your day to day actions, recollections from your childhood, etc. — and to try to decipher the meaning, to see what they are all “saying.” In short, the method is very much the same as the way we come to understand the cube, to know it — namely, to walk around it, to touch it, to get as many “profiles” on it as you can. The whole thing is very tentative and provisional, to be sure; it’s more an “art” than a “science.”

Of course, this sounds very much like what the Freudians do, with their free association, interpretations of dreams, etc. But, we said earlier, Sartre has no quarrel with Freudian technique. It’s their theory that is wrong.

There are two lines of criticism Sartre raises at this point:

- (1) The Freudians think the meaning of a particular action can be deciphered according to more or less fixed rules, a kind of code.

Certain objects are phallic symbols. Certain characters are “really” your mother, even though she doesn’t look anything like her. And so on. The Freudians do this sort of thing because they have a theory of behavior that is basically causal. For them, as we saw in the discussion of *The Emotions*, behavior (including the actual conscious acts of awareness) is caused by things going on in the unconscious region of the mind. And causality, of course, operates according to certain rules. That is what science is all about.

For Sartre, on the contrary, man is free. And this means there are no set rules of interpretation. You can’t decipher the meaning or signification of a bit of behavior by looking it up in some kind of Freudian symbol-book.

All this means, of course, that psychotherapy for Sartre is much more an art than a science. You have to play it by ear. Of course, in practice, that is exactly what the Freudian does too. It’s just that his theory says otherwise.

- (2) The Freudians think that what Sartre calls the “original project,” and what they call the “complex,” is always the same in the end.

Although the details may vary from person to person, depending on the particular circumstances of the case, the same fundamental principles are operating in every case. It is always a matter of the Id's striving to satisfy itself. Or rather it is a complex buried in the Id and ultimately to be explained in terms of Id drives.

For Sartre, this misses the irreducible uniqueness of each person. Such an approach tries to explain the individual in universal terms; after all, the Id, the Pleasure Principle, is the same for everyone.

And now, people, things get difficult. Listen hard.

Sartre too thinks that each original project fits a general pattern. But the general pattern on Sartre's theory does not provide an explanation. This general pattern is what Sartre calls the human condition in "Existentialism Is A Humanism."

For Sartre, each individual project is fundamentally a project to be. What I am aiming at ultimately is really to be, that is, to be, whole and entire, in the stable manner of an in-itself. I want to be conscious, of course. But I don't want any of the risk, any of the responsibility, any of the incompleteness of the for-itself. What I ultimately want, what I am ultimately aiming at, is a combination of the in-itself and the for-itself. My original project is thus an attempt to be God. This much is common to all original projects. But of course it comes about in an infinite variety of ways.

The original project has to be fundamentally a project to be, a desire to be, if it is going to fit the requirements Sartre insists on. Remember, if we are going to get at the individuality of a person ultimately, we have to be prepared to accept certain things as just brute givens, but only such things as really are brute givens — such that the question "Why stop there? Why not go one step further?" is inappropriate.

Is this what we have here? Sartre thinks so. We have reached a point where we can go no further. Pierre's present desire to go rowing on this stream today is a manifestation of his fondness for open air sports in general, which in turn shows his desire to be outdoors under any circumstances. And on and on. But his fundamental project is his desire to be, his desire to "make something of himself." For Sartre, you cannot go any further than that. "To be" is the most basic and ultimate you can get.

Just what is going on here? Isn't this the same kind of thing the Freudians and the psychological biographers were guilty of, explaining the particular in terms of the general? Isn't Sartre himself doing just what he is complaining that others do? What is the difference?

Let us look more closely.

Both Freud and Sartre can distinguish three levels here:

<u>Sartre</u>	<u>Freud</u>
---------------	--------------

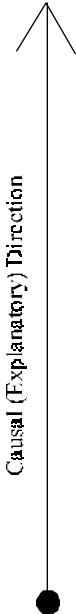
(1)	The empirical desire, the particular fact of consciousness	The empirical desire
(2)	The Original Project (= the person)	The Complex
(3)	The Desire to be God (= the general structure of all original projects)	The Pleasure Principle (= the general structure of the Id)

Let us look at the relations between these various levels. For Freud, the relation between (3) and (2) is a causal one. The general principles of psychology, the Pleasure Principle and whatever other general principles might be involved, along with various environmental factors, etc. — all those things combine to form the complex. It is a purely mechanical process. Given those ingredients, you will always end up with that particular complex. Similarly, the complex, which is hidden in the Id, is related to the empirical desire, which is what we are trying to decipher, by another causal link. The complex speaks to consciousness in the form of conscious impulses, desires, etc., according to the circumstances you find yourself in. But the point is that all the links in the Freudian picture here are causal links such that, given the causes, the effects necessarily follow.

On a schema like this, where is there any room for brute givens, for contingency? Clearly, only at level (3), the level of general structures. Everything else is explained in terms of those general psychological principles. It all comes back to the laws governing the Id. But if you ask why the Id is in every case made up this way rather than according to some other pattern, the Freudians will say “That’s just the way it is.”

So, in the Freudian picture, the explanatory direction goes up the diagram. Level (1) is explained by level (2), and level (2) by level (3). But (3) is where you stop.

	Sartre	Freud
(1)	The empirical desire, the particular fact of consciousness	The empirical desire
(2)	The Original Project (= the person)	The Complex
(3)	The Desire to be God (= the general structure of all original projects)	The Pleasure Principle (= the general structure of the Id)



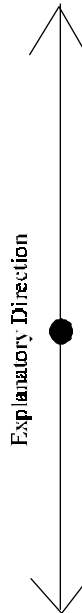
Contrast Sartre's view. For Sartre, none of the links between these various levels of analysis is a causal link. The relation between (2) and (1), for example, is not a causal link. The original project is signified by empirical, particular fact of consciousness, but it does not cause it any more than the cube causes the three sides we see. The relationship here is not one of causality, but rather the relation between one profile and the whole of which it is a profile. The empirical, particular fact of consciousness is for Sartre just one view on the whole that is the original project. The relation here is more like a whole/part relation than like a causal relation.

Similarly, the relation between levels (3) and (2) is not a causal relation for Sartre. The abstract structure "the desire to be God" does not cause my original project. Given that abstract structure — and for that matter, given any other general principles that might be involved — there is still all the room in the world for a variety of quite different original projects. The "desire to be God" is just the abstract notion of "trying to make something of yourself." All the uniqueness of an individual comes in what he is trying to make of himself. And the general structure "trying to be God" does not determine that.

Furthermore, that general structure does not explain anything. Any action, any project, fits into the general pattern "trying to be God." So, to say that a person's life fits into that pattern is just to say he is doing something. That is no explanation of what he does.

Thus, for Sartre, the brute given, the irreducible contingency, is located at level (2), not at level (3). The brute given is located at the level of the particular, not at the level of the general. There is still room for human freedom.

	Sartre	Freud
(1)	The empirical desire, the particular fact of consciousness	The empirical desire
(2)	The Original Project (the person)	The Complex
(3)	The Desire to be God (= the general structure of all original projects)	The Plesasure Principle (= the general structure of the Id)



Let's go over this again.

For both Freud and Sartre, we have general structures, general features that apply to everyone. And we also have particular, individual minds and their individual actions and desires. What is the relation between those two levels?

Most scientific theories — Freudianism among them, but also including various theories of physics, etc. — think of the general structures as somehow primary and basic. Individual things and events are viewed as following from these general principles.

Why does the apple fall? Because of the law of gravity. The particular event follows from the general principles.

Thus, the connection between the general principles and the individual events is a necessary connection. Given all the ingredients, we have to have such and such a result. Given the law of gravity, and given the mass of the apple, and given the density of the air, and given all the other things, the apple has to fall at the rate it does.

If this is so, and if we still try to find a place for brute facts in this setup, obviously the only place for them is at the level of the general principles. Everything else depends on them. But why are the general principles this way rather than some other way? Why does gravity decrease with the square of the distance, rather than with the cube of the distance (which would seem more natural, since we are talking about three dimensions)? It is at that level that we find brute facts. The law of gravity might have been different, but it isn't, and that's just a brute fact.

Sartre thinks in the case of human beings (and probably in all other cases too, if I understand Sartre correctly) this kind of approach is unsatisfactory, for the reasons we've seen. First, you can't get enough general principles to determine the particular case. But second, even if you could, the fact that you stop with these general principles rather than with more general ones yet is totally arbitrary.

For Sartre, contingency must be located at the level of the individual, not the general, at level (2) rather than level (3).

The reason we have the particular original projects we do is not because we are trying to be God. It is not even because we are trying to be God plus a bunch of other "becausees." It is the other way around. I don't have my particular original project because I am trying to be God. Rather, it is because all original projects are the way they are that we can generalize and say that all original projects fit the general pattern "trying to be God." The explanatory starting point is at level (2), not at level (3).

It is perhaps worth noticing that the three levels we distinguished above in the Sartrean and the Freudian psychoanalytic theories are very much like the three stages or levels we saw in Part II of *Transcendence of the Ego*. Thus:

<u>Sartre</u>	<u>Freud</u>	<u>Transcendence</u>
---------------	--------------	----------------------

(1)	The empirical desire, the particular fact of consciousness	The empirical desire	Momentary repugnance for Pierre
(2)	The Original Project (= the person)	The Complex	The state of hatred
(3)	The Desire to be God (= the general structure of all original projects)	The Pleasure Principle (= the general structure of the Id)	The quality of hatefulness in general

In all three cases, (1) is a profile on (2), which in turn is simply a special case of the more general (3).

(Note: In *Transcendence of the Ego*, there was a fourth level too, the Ego or psyche, underlying and unifying the whole business. We have nothing quite like that here.)

In *Transcendence of the Ego*, we were told that this picture is false. But now something very much like it is said to be the truth. What is the difference?

I will leave the answer to you to work out. But here are two clues:

- (1) The quality is not an empty, abstract structure that all states exhibit. Only some of them do. By contrast, all original projects share the general structure “trying to be God.”
- (2) The relations among the three stages in Sartre’s picture are very similar to the three levels in Freud’s picture, which we already know is going to be wrong. Thus, the empirical, particular psychological event is just one profile on the complex (although Freud also describes the relation as a causal one). And the complex is simply a particular configuration of quite general Id drives that apply to everyone. (Again, Freud also thinks of this as a causal connection.)

This suggests that what is wrong with the picture we were given in Part II of *Transcendence of the Ego*, like what is wrong with the Freudian picture, is not the structural interrelations of the three levels — since Sartre himself accepts something very much like that — but rather the interactions between the levels, what is supposed to “explain” what, and so where contingency can come in.

Conclusion

Finally, I want to look briefly at the “Conclusion” to *Being and Nothingness* (pp. 785-798), to try to draw some themes together and tie things up.

Copyright © 1996 by Paul Vincent Spade. All rights reserved. Permission is hereby granted to copy this document in whole or in part for any purpose whatever, provided only that acknowledgment of copyright is given.

The “Conclusion” is divided into two parts: “Metaphysical Implications” and “Ethical Implications.”

The first part, “Metaphysical Implications,” addresses two questions:

- (1) How are we to think of the “emergence” of being-for-itself from being-in-itself? How did it come about that we have not only being-in-itself (which stands in need of nothing else) but also being-for-itself?

Note: The question here is not why we have both. That is, Sartre is not here looking for a sufficient reason. We saw a long time ago that there is no sufficient reason. Rather, what he is asking now is about the process, the “mechanism,” by which this came about.

In the end, Sartre refuses to answer this question here. He raises it and says some things about it, but in the end claims he cannot answer it within the methodological confines of the book.

The question, he says, is a metaphysical question, not an ontological one. But the whole of *Being and Nothingness*, as announced on the very title page, is “A Phenomenological Essay on Ontology,” not on metaphysics.

What is the difference between ontology and metaphysics for Sartre? Well, the distinction is rather artificial perhaps, and certainly not explained very well, but seems to be roughly like this:

Ontology is the study of the general structures of reality. Hence the role of the eidetic reduction in Sartre’s book. Ontology is an appropriate thing to do phenomenologically.

On the other hand, metaphysics for Sartre is something we cannot really do phenomenologically. It involves hypotheses, like the sciences. In other words, metaphysics is conducted from what Husserl called the natural standpoint. Thus (p. 788):

We, indeed, apply the term “metaphysical” to the study of individual processes which have given birth to this world as a concrete and particular totality. In this sense *metaphysics is to ontology as history is to sociology*.
[Emphasis added.]

Think about that last sentence. Again (p. 788 once more), with respect to this first question:

Ontology can not reply, for the problem here is to explain an event, not to describe the structures of a being.

(The latter is what ontology does.)

Ontology can set some limits to metaphysical speculations about this question, but it cannot answer the question all by itself.

- (2) (pp. 790 ff.) The second question in this section on “Metaphysical Implications” is one we already saw raised at the end of the “Introduction.” How are we to think of the relation of the for-itself and the in-itself? Why are they both called being when they are so radically different? Is this just a terminological accident?

This is a very interesting section of the book, and in fact suggests that we may have to renegotiate everything. In the end, Sartre says again that ontology cannot completely answer this question; it is a matter for metaphysics to decide which is the most promising way to think about these things.

One way to think about reality is as what Sartre calls a “detotalized totality.” This is the way things have been described throughout *Being and Nothingness*. This way of looking at things is perfectly all right, but Sartre now thinks it need not be the best way to look at things. Whether it is or not is a metaphysical question.

This view, the view we have been studying throughout this course, is a kind of dualism. Being-in-itself and being-for-itself are two quite distinct kinds of reality, and neither can be reduced to the other.

But it is an odd kind of dualism. For, while the in-itself and the for-itself never mix, are never combined, nevertheless the whole meaning of this duality is given in terms of such a combination — the impossible in-itself-for-itself. That impossible combination (God) is what Sartre here means by a “totality.”

Of course that totality is impossible. We never reach it; it cannot exist. The two ingredients that would coalesce in that ideal totality remain two, separate but always tending toward, referring to, that totality. This is why Sartre calls it “detotalized.” Thus (p. 792):

Everything happens as if the world, man, and man-in-the-world succeeded in realizing only a missing God. Everything happens therefore as if the in-itself and the for-itself were presented in a state of disintegration in relation to an ideal synthesis. Not that the integration has ever taken place but on the contrary precisely because it is always indicated and always impossible.

It is this perpetual failure which explains both the indissolubility of the in-itself and of the for-itself and at the same time their relative independence....

What we have here, then, is a little like Plato’s theory of Recollection or Reminiscence, or like the mediaeval Augustinian theory of “illumination” — or, to a lesser extent, Descartes’ theory of innate ideas.

I see a dinner plate and I see it as “approximately” circular. It is not perfectly circular, of course; it is an imperfect circle. That is to say, the shape it actually does have refers to,

comes on to me as falling short of, a perfect version of itself. The perfect circle is never actually presented to us, but it is always implicit, always just over the horizon, and unifies all these quite different imperfect circles.

Similarly in Plato and in Augustine, the actual world reflects and bears traces of a perfect, ideal version of itself. We don't see that perfect version, but everything points to it — either to the Forms (in Plato), or to God and his divine ideas (for Augustine).

This is exactly the way Sartre has been viewing things throughout *Being and Nothingness* — even to the role God plays in this picture.

But of course for Plato and Augustine, the ideal was real, whereas for Sartre it is impossible.

This raises an important question: What difference does Sartre's atheism really make in the end? Would anything important change in Sartre's philosophy if God were not impossible — if he existed? Of course, the whole tone would probably change; it might be much more optimistic, for instance. But is that all? Or would the whole system come apart?

I don't know the answer to this. It is a good point to ponder in light of what we have just been discussing about Plato and Augustine.

Now, while this way of looking at things is perfectly all right for Sartre, and is the way Sartre himself has developed his theory throughout *Being and Nothingness*, he recognizes in his "Conclusion" that it is not the only way to look at things, and that it is a metaphysical question whether it is in fact the best way.

There is at least one other way, Sartre says.

Recall our discussion of distance from Part I, Ch. 1: "The Origin of Negation." There we said that the road between Bloomington and Indianapolis can be viewed in either of two ways (although not both at once):

- (1) Either as a road (positive), terminated at this end by Bloomington and at that end by Indianapolis (which thus serve negative roles).
Or
- (2) As what separates (negative) Bloomington and Indianapolis (which are here playing positive roles.)

Which way we look at it depends on the Gestalt we take on the road.

So too here. We have been viewing being-in-itself and being-for-itself as the beings here, as "positive" in the sense of what is really real. On the other hand, the phenomenon we have been viewing as not "really real" in this way, but as parasitic and derivative, as no more ultimately real than the events portrayed in the movie on my movie screen.

But we can also view the situation with just the opposite emphasis, just the opposite Gestalt (p. 794):

It is up to metaphysics to decide which will be more profitable for knowledge (in particular for phenomenological psychology, for anthropology, etc.); will it deal with a being which we shall call the phenomenon and which will be provided with two dimensions of being, the dimension in-itself and the for-itself (from this point of view there would be only one phenomenon: the world), just as in the physics of Einstein it has been found advantageous to speak of an event conceived as having spatial dimensions and a temporal dimension and as determining its place in a space-time; or, on the other hand, will it remain preferable despite all to preserve the ancient duality “consciousness-being.”

On this new and alternate viewpoint, we would not have a dualism of the in-itself and the for-itself, but rather a monism of the phenomenon. Sartre warns us that this alternative view must avoid the temptation to turn the phenomenon into simply being-for-itself (resulting in something like Husserl’s idealism, in which everything was ultimately mental), or at the other extreme, to regard the phenomenon as some kind of being-in-itself — as he says, to “look on the phenomenon as a new kind of object” (p. 795).

Notice here how Sartre keeps talking in terms of utility. Metaphysics must decide which is “more profitable,” Einsteinian physics has found it “advantageous” to think in terms of space-time, etc. Here the pragmatic, practical point of view characteristic of the natural standpoint is coming out. Once again, metaphysics is not ontology.

Ethical Implications

Finally, let us look quickly at the second part of the “Conclusion,” the discussion of ethical implications. Here too Sartre does not draw any definite conclusions. The ontological task of the book is completed; the job of constructing an ethics is a different task entirely.

Nevertheless, Sartre does say some things in this section that are extremely interesting and tantalizing, and suggest the ways his thought will continue to develop after *Being and Nothingness*.

Existential psychoanalysis, he says (p. 796), leads us to repudiate the spirit of seriousness — the view that values are absolute givens, that they are objective needs and demands in the world. Thus (p. 796 again):

Objects are mute demands, and he [= man] is nothing in himself but the passive obedience to these demands.

That is the point of view of the spirit of seriousness.

Existential psychoanalysis, he says, is going to rid us of that illusion. Which is to say, Sartre’s own existential theory as developed in this book should rid us of this illusion if

we just can put it into practice in our own lives. It teaches us that we are the source of values, that they are not objective, that no single one of them is necessary, that it is all up to us. Thus (p. 797):

... it amounts to the same thing whether one gets drunk alone or is a leader of nations.

But at the end of this section, he says something extremely interesting (p. 797):

But hitherto although possibles could be chosen and rejected ad libitum, the theme which made the unity of all choices of possibles was the value or the ideal presence of the *ens causa sui* [= “the being that is the cause of itself” = God]. What will become of freedom if it turns its back upon this value?

In other words, the picture we have got so far in *Being and Nothingness* is that human beings are simply projects of trying to be God. The impossible ideal of the in-itself-for-itself is the goal and purpose of all our acts. But now he seems to be suggesting that, once we realize that we are the sources of our values, and that we do not have to have the values we do, it is open to us to reject this ultimate value too. Does this mean that we could stop trying to be gods, stop trying to “justify our existence”? As Sartre very properly asks, “What would happen then”?

Can we in fact do this? Sartre isn’t sure here (p. 797):

Will freedom carry this value [that is, the in-itself-for-itself] along with it whatever it does and even in its very turning back upon the in-itself-for-itself? Will freedom be reapprehended from behind by the value which it wishes to contemplate?

That is, can we abandon this “ultimate” value? Or would the very attempt to do so be governed by that very value too, so that the whole process would be self-defeating, and a kind of bad faith? Perhaps that would be the outcome. But Sartre goes on to suggest a brand new possibility (pp. 797-798):

Or will freedom, by the very fact that it apprehends itself as a freedom in relation to itself, be able to put an end to the reign of this value? In particular is it possible for freedom to take itself for a value as the source of all value, or must it necessarily be defined in relation to a transcendent value which haunts it?

The interesting thing is that this is exactly what Sartre suggests in “Existentialism Is A Humanism,” where Sartre takes freedom as the great value — not God. And this seems to be what he comes to mean by ‘authenticity’. It is not Bad Faith; it is not an attempt to be something impossible. It is purely and simply an attempt to be free.

This picture, of course, would seem to require a radical reorientation of everything we have done in this book.

To be free is not to coincide with oneself, not to be self-identical. This new value, therefore, is not just a matter of replacing the old, impossible ideal with something else. It is a matter of replacing the old, impossible ideal with its direct opposite. Thus (p. 798):

This freedom chooses then not to *recover* itself [as we were trying to do when we were trying to be God] but to flee itself, not to coincide with itself but to be always at a distance *from* itself.

So this is perhaps what authenticity is. And these passages are the closest glimpse Sartre gives us of it in this book. He still doesn't explain how it is possible within the ontology he has constructed in *Being and Nothingness*. And, in fact, in this passage, he doesn't even commit himself to saying it is possible. He simply poses it as a question.

I have a suggestion for how to interpret Sartre here. It is purely speculative, but it is the only way I know of to make sense of the new possibility he suggests here without totally rejecting everything we have done for the last eight hundred pages.

The whole ontology we have developed in *Being and Nothingness* indicates that the attempt to be God is inevitable; it is not something we can avoid. What he is saying here in these last pages sounds as if, once we realize that we are the source of values, we don't have to try to be God any more, and can choose freedom itself as our ultimate value. But of course, if that is really what Sartre is saying, then we do have to jettison everything we have done so far.

But perhaps things are not so bad. Consider an analogy from sports. Suppose you are a runner, and that you compete, let's say, in the 100-yard dash. The 100-yard dash, for purposes of our analogy, is going to be like life.

Now, while you are running the race, your goal is to win it. In fact, that is the whole point of racing. It is the ultimate goal, and everything else is done in the light of that goal.

This much is like what we have seen so far in *Being and Nothingness*.

But now suppose you begin to think of racing as something that serves a higher goal. Racing is no longer just an end in itself for you. You run your races for the sake of something else — the glory of competition, say, or to keep fit, or because your competitors are good friends and it's a good way of socializing.

It doesn't matter for the purposes of my illustration what that higher goal is. The point is that winning is no longer your ultimate value. It is subordinated to some other value. You don't try to win just to be winning.

In that sense, you no longer have the ultimate value you did before. You have replaced it with a new one.

Something important happens with this new attitude. Previously, when winning was everything, if you lost the race you had just failed, and there was no way to put a nice

face on it. You had failed to achieve the goal that was the whole point of your participating in the race in the first place.

But with this new attitude, even if you do lose the race, in the end that doesn't really matter. You still have the glory of competing, you still keep yourself fit, you still can get together with your buddies who also enjoy this kind of sport.

The same thing would hold if, for whatever reason, you could not win the race, if winning were strictly impossible. You can still fulfill those higher goals, even if you don't — and cannot — win, and even if you know that.

And yet — and this is the point of my analogy — you don't stop trying to win. In fact, you don't even try any less to win. You still run for all you are worth. If you slack off even a little bit, you are not really racing. You're not giving it your best.

So too — and now we return from our analogy — perhaps what Sartre has in mind in these final pages is a picture where we don't stop trying to be God, and we don't stop trying to be God with all our might. Perhaps we cannot stop trying to be God. To that extent, the whole structure we have developed throughout *Being and Nothingness* is still intact.

But perhaps we can come to view that goal of trying to be God as subordinate to some higher goal — perhaps, as Sartre himself suggests, freedom itself. I may never achieve the lesser goal of trying to be God, but in the process of trying I may very well succeed in reaching this new and higher goal. (In fact, if the new goal is freedom, it is hard to see how I could ever not succeed, since human beings are inevitably free.)

I don't know whether something like this is what Sartre has in mind or not. I think it is a promising interpretation and deserves to be thought about.

Note one consequence of following out this suggestion. It means that, even though I must inevitably fail at being God, which is something I am aiming at with all my might, it does not follow that everything is futile. It does not follow that there is no point in the attempt. On the suggestion we have just seen Sartre making (at least if my interpretation is correct), it does not follow after all that “Man is a useless passion.” The very pessimistic tone of *Being and Nothingness* seems to be turning, here at the very end, more toward the much more optimistic tone of “Existentialism Is A Humanism.”

Note also that this reading of what “authenticity” is makes it look very much like Kierkegaard's “Knight of Infinite Resignation.” (See the discussion in my notes on *Existentialism Is A Humanism*, in the course packet.)

At this point, we're just speculating. But while we're at it, let's speculate some more:

Recall back when we were discussing Sartre's treatment of *value*, and we were talking about some of the implications of *Being and Nothingness* for an “existential ethics.” At that point, I asked *what is “bad” about “bad faith”?*

We said it looked as if Sartre's ethics were headed toward out and out relativism. There are no absolute values, and anything goes. So what's wrong with being in bad faith if I

want to be? After all, there are certain obvious advantages to bad faith — it's definitely more pleasing and reassuring than dealing with various unpleasant facts about ourselves.

It seems to me now that one thing Sartre can say is: There's nothing wrong with being in bad faith, if that's what you want to do. But that never is what you want to do, now is it?

What's wrong with bad faith, in other words, is that it doesn't work. That doesn't mean you can't deceive yourself; you can — and that self-deception can sometimes be maintained for long period of time, even though it is a “metastable” situation that keeps threatening to fly apart.

Rather bad faith doesn't work in the sense that it doesn't succeed at what it's trying to do. And I don't just mean that it doesn't succeed in reaching our ultimate goal of trying to be God. Of course it doesn't do that; but it also doesn't succeed in reaching even its more immediately goal.

The point is: Bad faith is never an attempt to be in bad faith. The parents of the Vietnam MIA's weren't trying to believe-their-sons-were alive-while-realizing-that-probably-wasn't-so. No, they were trying to believe-their-sons-were-alive-without-having-to-work-at-that-belief.

Bad faith is never aiming at forced belief but at a natural belief that answers to the facts. And that's exactly what it fails to achieve.

Sartre's calling bad faith “bad” is thus a little like an Aristotelian/natural-law theory of ethics. For such theories, we automatically and spontaneously aim at certain things, simply in virtue of our having the natures we do and being the kind of things we are. given that nature with those goals, certain things will — just as a matter of act — help us achieve those goals, and certain other things will impede us. From this point of view, ethics becomes a matter of enlightened self-interest. There is none of this Kantian “autonomy” of ethics.

So, in a sense, calling bad faith “bad” might be nothing more than an observation that it's not what you really want after all.

Still, for Sartre we never get what we really want anyway. So why is bad faith singled out for condemnation?

Well, as we now see in the discussion of “Ethical Implications” in the *Conclusion*, maybe we can get what we really want. There is this business about “turning our back” on the ideal of God, and taking freedom as our value. That is a goal, Sartre is suggesting now, that can be achieved. My problem with this is that it is hard to see how we can work this out without simply rejecting the whole framework of *Being and Nothingness*, according to which the desire to be God is inevitable.

Anderson and Detmer, in their books, make a big deal out of this “freedom as a value” business, but in the end, I think, do not really address this problem.

Furthermore, since we are *automatically* free anyway, what sense does it make to *aim* at it? Detmer tries to make a distinction here between *ontological* freedom, on the one hand, and *liberty* on the other. And maybe that's part of the story.

But if taking *freedom* as our value means taking *liberty* as our value, then it's hard to see how we end up with an *achievable* goal after all.

Contrary to Detmer, I suspect that what Sartre has in mind is not *liberty* so much as our old, familiar *ontological* freedom. And of course, since we already have *that* in every action we take, it makes little sense to *aim* at it. Perhaps what Sartre is talking about is not so much *freedom itself* as a kind of *glorying* in our freedom, *aiming* at it instead of *fleeing* it, being *honest* about our freedom: *realizing* that things are up to us, etc. (Intellectual honesty has been the fundamental philosophical value ever since Socrates, after all, so there is something appropriate here.)

Taking *honesty* especially about our freedom and responsibility as our highest value leaves us with a perhaps *achievable* goal. And if that's what *authenticity* is, it's easy to see why Sartre regards it as "better" than bad faith.

There is still an objection: Bad faith is achievable too. After all, people are in bad faith all over the place. So once again, why is authenticity "good" and bad faith "bad"? Answer: Authenticity can not only be *achieved*, it can also be a *goal*. Bad faith, as we just saw a moment ago, can be *achieved*, but can never be a *goal*.

It is not clear how all this works out at the end of *Being and Nothingness*. Sartre says he is going to write another book on these ethical issues, even though he never completed it. But, in any case, that is the end of *this* book.