

Hegel on the Life of the Understanding

David Morris, Department of Philosophy, Concordia University, davimorr@alcor.concordia.ca

Published in *International Philosophical Quarterly* 46 (2006): 403-419

https://secure.pdcnet.org/pdc/bvdb.nsf/journal?openform&journal=pdc_ipq

Abstract

This article clarifies Hegel's argument in "Force and the Understanding," in his *Phenomenology of Spirit*, by developing Hegel's underlying point through discussion of recent and ongoing issues concerning explanation in natural and psychological science. The latter proceeds via a critical discussion of the problem of other minds and the "theory theory of mind". The article thereby shows how and why Hegel's analysis of the understanding inaugurates a crucial transition in his *Phenomenology*, from consciousness to self-consciousness and life. Putting Hegel's underlying points into conversation with recent science shows how his point—that scientific understanding is not abstract but embedded in human life—still speaks to science.

Hegel's "Force and the Understanding," in his *Phenomenology of Spirit*, is a challenging chapter in an already difficult work, not the least because it deals with now arcane moments in the history of science. The chapter, though, is crucial in the *Phenomenology*, if not in the history of Western philosophy. In the *Phenomenology*, "Force and the Understanding" prepares for Hegel's transition from consciousness to self-consciousness, desire and recognition.¹ It culminates Hegel's novel argument, in the section on "Consciousness," that the topic of the philosophy of knowledge could not be consciousness, or faculties of cognition, but must be the complete form of self-conscious life, which as subsequent chapters show, is lived out through historical relations betwixt humans. As H.S. Harris puts it in his masterful *Hegel's Ladder*, the *Phenomenology* has a "second beginning" after "Force and the Understanding."² This chapter, then, is central in the history of Western philosophy, for it undermines the central tendency of modern philosophy, namely securing abstract reason or cognitive faculties as the foundation of knowledge. On Hegel's argument philosophy must rather

describe how knowledge and reason find their basis in the unfolding institutions of human life. Hegel's revolutionary change in direction, and criticisms of his particular pursuit of this direction, resonate in movements as diverse as pragmatism, existentialism, Marxism, or Foucaultian and Derridian philosophies.

This paper has a twofold aim. The first is to clarify Hegel's argument in "Force and the Understanding," not by giving a line by line exposition of Hegel's chapter and all of its internal intricacies³, but by striking to Hegel's underlying point through discussion of recent and ongoing issues in science. The second is to thereby show how and why Hegel's analysis of the understanding inaugurates a change in direction, from consciousness to self-consciousness and life. Putting Hegel's underlying points into conversation with recent science not only helps clarify Hegel, it shows how his point—that scientific understanding is not abstract but embedded in human life—still speaks to science.

1) Theoretical Consciousness is Self-Consciousness

Hegel's central result in "Force and the Understanding" is that theoretical consciousness is a form of self-consciousness. A recent non-technical essay, "Can Science Explain Everything? Anything?," by Steven Weinberg, a Nobel Prize winning physicist, lets us approach Hegel's result via the perspective of a working scientist who is still plagued by the problem that drives "Force and the Understanding."⁴

Weinberg's essay is prompted by a remark of a scientist friend that "Well, of

course you know science does not really explain things—it just describes them.” The problem here is a traditional one in philosophy of science, namely that scientific understanding teeters on the razor’s edge between explanation and description. In his essay “Hegel’s ‘Inverted World’,” Gadamer calls this teetering the dialectic of explanation, and rightly flags it as central to Hegel’s “Force and the Understanding.” Gadamer gives the example of laws of phonetic change: they really just give a descriptive formula of a linguistic phenomenon; as descriptive, the laws are identical with the phenomena; so it is hard to see how they count as something different from the phenomena, something that could explain them.⁵ Or as Ernst Rutherford famously declaimed (cited in Weinberg), “In science there is only physics; everything else is stamp collecting.” Disciplines like biology just collect data, so biological “laws” do not really explain the phenomena, they are just descriptions of the phenomena put in general form.

Presumably Weinberg’s friend thinks that all science, even physics, is “stamp collecting,” and to counter this Weinberg needs to show how science produces something other than descriptions, something that would count as explaining the phenomena—and explaining them objectively. Weinberg turns to the example of thermodynamics, the science of heat. He argues that “Thermodynamics itself is never the explanation of anything—you always have to ask why thermodynamics applies to whatever system you are studying, and you do this by deducing the laws of thermodynamics from whatever more fundamental principles happen to be relevant to that system.” If you just work out the laws that describe how heat is manifest in a gas, you are stuck with description. If you want to explain heat, you need to deduce those descriptive laws from the

statistical mechanics that describe the movement of gas molecules. Weinberg rescues physics from philately by imposing what I call a deduction criterion.

But Hegel’s analysis in “Force and the Understanding” shows that the deduction criterion does not escape the dialectic of explanation. This is because of the way that consciousness accesses the phenomena. The problem is this: According to Weinberg’s deduction criterion, explanation requires both a law that describes the appearance of the phenomenon, and a way of deducing that lawful appearance from what I call a more fundamental substratum. But the empirical source of our claims about both these things, the lawful appearance and its substratum, is one and the same phenomenon. The sensuous phenomenon is the appearance of both the explanandum (that which is to be explained) and the explanans (that which is doing the explaining). We look at the phenomena of heat. This prompts us to describe the laws of its dissipation over time. By digging further into the same phenomena, we find that we can understand that law statistically in terms of the aggregate behaviour of moving molecules. For our understanding, one and the same sensuous phenomenon is evidence of both a lawful appearance and its fundamental, explanatory substratum. A question immediately arises: why is it that heat understood statistically in terms of the aggregate behaviour of moving molecules counts as more fundamental than and explanatory of heat described in terms of laws of its sensuous appearance? Let us consider a dialectical tension at work in this question.

If the substratum were simply identical with sensuous appearances (if the explanans is identical with the explanandum), then the so-called explanation would really just be a description of appearances, and Weinberg’s friend is right. So we must be able to claim

that the fundamental substratum is something different than lawful appearances.

But if the fundamental substratum is different from what sensuously appears, it must be supersensible relative to the appearances that it explains, it must be a theoretical object. Yet it is one and the same sensuous phenomenon that gives evidence both of its law of appearance and of its supersensible substratum. More, the supersensible substratum is supposed to explain what appears. On both counts the supersensible substratum must be inherently linked to what appears. What appears is nothing other than the appearance of the supersensible substratum. The phenomenon of heat is nothing other than the phenomenon of statistical mechanics of molecules. But then the substratum is not different than lawful appearances, and the substratum is not entirely supersensible, since it appears as the phenomenon that it explains.

With one hand we draw a distinction within what appears, claiming there is something supersensible behind the scenes that makes what appears be the way it is. With the other hand we erase that distinction, by claiming that what appears is nothing other than the appearance of a substratum behind the scenes. This tension is at the heart of Hegel's dialectic of explanation, and it means that *it is our consciousness that posits a difference, within appearances, between the explanans and the explanandum*. To cite Hegel:

...the difference, then, is posited by the Understanding in such a way that, at the same time, it is expressly stated that the difference is not a *difference belonging to the thing itself*. This necessity [of positing the difference], which is merely verbal, is thus a recital of the moments constituting the cycle of the necessity. The moments are indeed distinguished, but at the same

time, their difference is expressly said to be *not* a difference of the thing itself, and consequently is itself immediately cancelled again. This process is called '*explanation*.' (M154, W/C 109)

The difference between the explanans and the explanandum, for example, between heat and molecular motion, is not a difference *belonging to objects themselves*, it reflects our way of describing things in our effort at explaining them. This is plainly right: in my hot cup of tea there are not two different things, heat dissipating in this and that way, and molecular motions, they are identical. Yet I differentiate the two in trying to explain heat.

Weinberg gives two examples that support the Hegelian point that it is our consciousness that draws the distinction between explanans and explanandum. (1) Sometimes it is not clear what explains what. Weinberg points out that if you are Newton deducing the law of gravitation from the fact of Kepler's laws of planetary motion, then Kepler's laws explain Newton's laws of gravitation. But nowadays it would "feel absurd" to say that Kepler's law explains Newton's, since we take Newton's laws to be more general, and students today learn to deduce Kepler's laws from Newton's, not the other way round. Yet as Weinberg remarks, the criteria of generality are ambiguous. What counts as general is not specified by things themselves but by what we are trying to explain. And here there is another peculiarity. (2) "We often say that something is explained by something else without our actually being able to deduce it," for example, we claim that the "value of [...] proton mass is entailed by quantum chromodynamics, even though we don't know how to do the calculation." The peculiarity is that the understanding is oriented by an unknown, purely theoretical, supersensible object that is nonetheless fundamental. How could a

theoretical term be manifest as fundamental if we do not know how it works in explaining the phenomena? How could such a fundamental substratum have objective status if it is posited as theoretical, as reflective of the demands of our explanation?

It looks, then, like explanation goes in circles, explaining the explanandum by way of the explanandum redescribed; or that in explanations, the vector between explanans and explanandum sometimes reverses according to our interests; or that explanations sometimes appeal to an explanans that we do not fully understand. On Hegel's analysis, the underlying problem is this: In the course of explanation, the understanding seeks, *via* its consciousness of the world, to discover a unifying, fundamental, objective substratum that explains the world, but the very structure of explanation is such that what it actually grasps is something that reflects consciousness. Such a something does not seem to satisfy the initial aim of understanding and explaining, it does not seem to provide an objective stabilizing ground or criterion for explanation. From the scientific perspective, at least as it is usually portrayed, this is particularly horrifying. It is important to note that all this is a consequence of a commitment to the empirical: we access empirical facts only through a consciousness embedded in the world, and we cannot escape this; if our theoretical claims are driven by the facts, then they too arise in light of our consciousness of the world; the difference between facts and theoretical objects is thus drawn within our consciousness of the world; and so it is a difference that reflects consciousness, it is not something purely objective.

But in Hegel's analysis, this problem—that the criterion of explanation reflects consciousness—turns out to be the

solution to the dialectic of explanation. That is, the solution is simply to acknowledge that the initial demands of understanding and explaining are wrong-headed, that a satisfying explanation could never come to rest in an object purely alien to consciousness. The understanding is satisfied by an object that appears as already reflecting consciousness's way of articulating it. *The satisfaction of the understanding is a form of self-consciousness, something more than mere consciousness of the world.*

Given the centrality of this result to Hegel's "Force and the Understanding," it is worth deepening it through a summary in terms of Hegel that also sheds light on his argument. First, "Force and the Understanding" is a chapter in Hegel's section on "Consciousness," where Hegel analyzes philosophies that claim that faculties of consciousness are the source of comprehensive knowledge of the world. In his first two chapters he quickly demolishes the claim that a faculty of immediate sensory reception could give comprehensive knowledge, for at the very least our sensory faculty requires already given universal structures (such as space and time), and what is given in these structures must undergo perceptual synthesis that articulates and categorizes the given.⁶ The faculty of perception, then, is well suited to 'postage-stamp collecting,' to cataloguing and describing regularities in the phenomena. But what is required for knowledge is something further, namely understanding, which grasps the underlying unity of the observed phenomena. For example, perceptual observation can tell you that salt, which has the properties of forming white, cubical crystals, and tasting tart on the tongue, also has the property of promoting rust. But until you understand how all these properties are necessarily interrelated, you do not have comprehensive knowledge of

salt. Perception cannot let you know that it is salt *qua* salt (whatever it is that makes salt be salty, white, cubical and tart) that promotes rust; perception cannot rule out, for example, that it this particular sample of salt *qua* contaminated, or *qua* having peculiarities found only in salt from this mine (say, peculiar allotropes), or *qua* being heated by the sun, that promotes rust. (Perception cannot sort out accidents from essences.) What is needed for comprehensive knowledge of salt as rust promoting is, for example, understanding that the very same chemical structure that forms cubical crystals in salt necessarily and inherently also makes it look white, taste tart and promote rust.

In Hegel's analysis, this sort of knowledge requires cognition of objects in terms of underlying forces and laws that go beyond what is immediately perceived. But this plunges us into a dialectic of unity and diversity. In comprehending the diversity of sensuous appearances as the expression of underlying, unified forces, we need to invoke a multiplicity of different, interrelated forces. And in order to unify this diversity of forces we appeal to the concept of an underlying law. But law expands into a multiplicity of different laws that are nonetheless conceived in terms of one unified law of everything.⁷ In terms of our framework, the important point is that this expansion of the dialectic of unity and diversity pushes consciousness into thinking of supersensible, that is, theoretical, objects, because unity and diversity have the sort of "immediately cancelled" difference discussed above, a difference posited by the understanding that could not be located in sensuous objects.⁸

The dialectic of unity and diversity, then, leads into a dialectic of the sensible and the supersensible, which Hegel takes up in his superbly difficult discussion of the "inverted world." In effect, this new

dialectic aims to give "immediately cancelled" differences some sort of objective form; instead of admitting that it is consciousness that is drawing distinctions between the explanans and explanandum, the understanding tries to project this difference into the world, in terms of a difference between a supersensible explanans and a sensible explanandum. Hegel's point, though, is that the dialectical tensions underlying this explanatory strategy drive the understanding from: (1) explaining things via supersensible, theoretical objects that are reflective merely of perceptual appearances, in what Hegel calls the first "supersensible world"; to (2) explaining things via a "second supersensible world" or "inverted world" in which perceptual appearances are already 'theory laden' such that appearance *itself* already reflects *consciousness*.⁹

To illustrate, using an example from science: In 1733, in one of the important episodes in the history of electrical science, the French chemist Charles Dufay, observing that rubbed glass and resin behaved in quite different ways, hypothesized that there are two kinds of electricity, which he called vitreous and resinous. Dufay's hypothesis illustrates the explanatory strategy at work in the first supersensible world: Dufay's two kinds of electricity are supersensible¹⁰, theoretical objects that shadow and explain—reflect—two sorts of sensible electrical phenomena; and the supersensible objects, vitreous and resinous electrical fluids, are distinguished by way of characteristics of the perceived objects they are to explain, namely, they are distinguished as appearing in glass versus resin. In a sense, this strategy is not much of an advance over the sorts of claims that perception can make, and as explanatory it is not much better than Moliere's doctor claiming that opium puts people to sleep because it has a dormitive principle. But the

claim does advance science because it discerns two different sorts of electrical phenomena, as differences that are to be understood, not merely perceived.

The history of science shows that giving a satisfactory explanation of these phenomena drives the understanding to the second supersensible world, in which the phenomena themselves are redescribed. Satisfactory scientific understanding is achieved when the two kinds of electricity are no longer distinguished by the fact that they appear in glass versus resin, but by being positive and negative electrical charges. This redescription is the first step in grasping the underlying unity of charges. The key point is that the distinction between positive and negative is *not* drawn in or by immediate sensible appearance itself. What is directly given to perception are different behaviours of glass and resin, not something marked with the conceptual distinction between positive and negative signs, yet you understand that the perceived behaviours of glass and resin are explained by appeal to a conceptual distinction. And you thereby understand that what is at work in sensible appearance is positive and negative charges. You are no longer, then, looking at things *qua* differentiated from one another by merely perceived, immediately given determinations (e.g., the difference between brittle, cold, clear glass, and sticky-friable, warm, cloudy resin); you are looking at them *qua* differentiated by characteristics grasped by the understanding (the difference between positive and negative). Again, when you tingle your tongue by touching it to the top of a nine-volt battery, sensory experience does not give you something labelled as a relation between “positive” and “negative”; but to explain your experience, you must first redescribe it as a matter of tonguing positive and negative poles that are themselves interrelated (because, it turns out, the negative pole has a surplus of

electrons relative to the positive pole, and the surplus flows through your saliva, tingling your tongue, etc.).

The status of the sensible objects of explanation has thereby changed: they are no longer things entirely alien to your thinking, differentiated by sensible characteristics that happen to pile up in the postage-stamp album of the world; they themselves manifest a distinction between positive and negative charge, a distinction that *reflects the understanding*, a distinction that is not alien to you.

There are several reasons why Hegel speaks of objects of this sort as constituting (for consciousness) the “inverted world.” The first, which is well explained by John Russon, is this.¹¹ In the world-view of the first supersensible world, the supersensible, theoretical object, is “the real thing” that stands behind the perceived world: we read from the sensible to an invisible supersensible behind it (from rubbed glass to vitreous electrical fluid). This is the “classic” vector from the sensible to the supersensible that we might find in traditional interpretations of Plato, for example. The world-view of the second supersensible world this vector is inverted: we read from the supersensible (from concepts like positive and negative) to the sensible world as reflecting the supersensible; and we read supersensible, theoretical articulations right *in* the perceived world, as unified with it, not standing behind it unseen (this button of metal *itself* is the negative pole of the battery).¹² And so the second supersensible world also inverts the usual world-view that things in the perceived world are in themselves bereft of intelligibility. Here we might think of Holmes and Watson: for Holmes, the intelligibility of things is visible right in them; seeing the red mud on a client’s shoes is much like reading a book that says “the client has been in this part of

London today.” This startles Watson’s and up-ends his world-view, because for Watson the pages of this book are buried; all he sees is mud.

Further, given this sort of reading, in the inverted world, supersensible terms can invert their values. In contemporary electronics, the negative pole of a battery is understood as having a surplus of electrons relative to the positive pole, and when a circuit is closed, the surplus electrons flow from the negative to the positive pole. But the assignment of values and distinctions such as positive/negative and surplus/deficit, can be inverted according to how we need to read the phenomena. For example, when explaining semiconductors such as PNP transistors, it is helpful to think of the N (negative) region as having a *deficit* of holes rather than a surplus of electrons (holes themselves are *deficits* of electrons, of negative charges, creating by doping the crystalline structure of the semiconductor lattice so as to displace electrons); and to think of *holes* (absences) moving around when current is applied to the N region. There is nothing to prevent us from analyzing current in batteries as a flow of a surplus of electron deficits from the positive to the negative pole. When we read our understanding in the visible phenomena, what is visible to the understanding is inherently invertible.

Put another way, when we explain from the world-view of the first supersensible world, we ask “What are the theoretical objects that explain the perceived facts?” We answer the question by seeking a supersensible, theoretical explanans that, it turns out, is in fact merely the explanandum redescribed (e.g., a vitreous electrical fluid that explains observed charges in glass). This explanatory strategy repeats the dialectic of explanation in a new register. In contrast, when we explain from the world-view of the second supersensible

world, the inverted world, we ask “How do the phenomena *themselves have to be* if they are explicable and intelligible to us?” In posing the question this way we admit that first of all we have to redescribe the explanandum as already reflecting the understanding, rather than having the sensible drive us to seek a supersensible behind it. To explain the two kinds of electricity we have to step back from what perception directly tells us (that glass and resin respond differently to being rubbed) and *first of all have the insight that we should redescribe the explanandum* as constituted by negative and positive charges, as phenomena *already inherently reflective of theoretical distinctions, such as the distinction between positive and negative.*

As Hegel puts it, “The Understanding’s ‘explanation’ is primarily only the description of what self-consciousness is” (M163, W/C 116), for what we observe in the phenomena are distinctions that reflect our own consciousness and criteria. Our consciousness, then, is not simply receptive of objects, nor is it simply synthesizing and unifying what is given to it; we comprehend objects that first appear independent of consciousness, as in fact reflective of consciousness. This sort of consciousness is what Hegel calls self-consciousness. Kepler’s laws can explain Newton’s laws or the other way around, there is really nothing in the object that resists this reversal; what resists this reversal, what makes the explanation go one way and not the other way, is our way of being conscious of these relations; and so when we are conscious of Newton’s laws as explaining Kepler’s we are in fact also self-conscious of the interests and criteria we are bringing to bear in our explanations.

The initial problem broached was that the dialectic of explanation erodes resistance points, objective grounds, for drawing a

distinction and stabilizing the relation between the explanans and the explanandum. We have now found a resistance point, but in being conscious of it we are implicitly describing our way of being self-conscious, we are implicitly conscious of our own criteria shaping our consciousness of the world. Consciousness, as Hegel puts it, overarches (*übergreift*, M166, W/C 120) its object, comprehends things that appear independent of consciousness as in fact reflecting and satisfying its own interests, and as therefore dependent on consciousness. What satisfies explanation, makes an explanation count as an explanation, is not anything purely objective or alien to us, but an object that appears as already reflecting consciousness's way of articulating it, so, as said above, the satisfaction of the understanding is a form of self-consciousness. Consciousness identifies itself with and through its object, and it is in this identification that it finds the satisfaction of explanation.

2) The Life of Theoretical Consciousness *via* "The Scientist in the Crib"

In the chapter following "Force and the Understanding," Hegel discusses self-consciousness in terms of life and desire. The transition to life and desire is one of the most difficult yet crucial transitions in the *Phenomenology*. So we now need to see why consciousness's overarching identification with its object is to be conceived in terms of life. I approach this point through discussion of a form of scientific consciousness that is often overlooked...because it is not usually taken to be scientific consciousness at all, namely the infant's consciousness. H.S. Harris gives a precedent for this shift in showing how Hegel's turn from consciousness to self-consciousness can be interpreted in

terms of family life, and also suggests it in his discussion of understanding as learning.¹³ But the proximate stimulus for my shift is a recent book, *The Scientist in the Crib*, in which three prominent child psychologists, Alison Gopnik, Andrew Meltzoff and Patricia Kuhl give a (non-technical) summary of their research. They argue that children are not simply passive vessels to be filled with knowledge, they are active in learning about the world. Indeed, children, even very young infants, act in the manner of scientific investigators of the world who learn by making and testing theories.¹⁴

Specifically, what I want to look at is the child's consciousness of others. For the authors of the *Scientist in the Crib*, especially Gopnik, the child's experience of other minds gives strong proof that children are theoreticians. This claim is in part motivated by the classical problem of other minds. As the authors of the *Scientist in the Crib* put it, in a pungent, non-technical formulation of the classical problem:

All that really reaches us from the outside world is a play of colours and shapes, light and sound. Take the people around the table. We seem to see husbands and wives and friends and little brothers. But what we really see are bags of skin stuffed into pieces of cloth and draped over chairs. There are small restless black spots that move at the top of the bags of skin, and a hole underneath that irregularly makes noises. The bags move in unpredictable ways, and sometimes one of them will touch us. The holes change shape, and occasionally salty liquid pours from the two spots.

This is, of course, a madman's view of other people, a nightmare. The problem of Other Minds is how we somehow get from this mad view to our ordinary experience of people. (4-5)

According to this formulation of the problem of other minds, perceptual consciousness could not give us an encounter with other minds, for perception only gives us bags of skin, and other minds are something more than bags of skin. To encounter other minds we need something more than perception.

Gopnik argues that the something more is a *theory*. The infant is a theoretician who comprehends her world, and other minds appear in that world as the explanation of phenomena. Like quasars, other minds, according to Gopnik, are not direct observables, but theoretical postulates necessary to explain observables. As Gopnik puts it in her technical work:

The developmental evidence suggests that children construct a coherent, abstract account of the mind which enables them to explain and predict psychological phenomena. Although this theory is implicit rather than explicit, this kind of cognitive structure appears to share many features with a scientific theory. Children's theories of the mind postulate unobserved entities (beliefs and desires) and laws connecting them, such as the practical syllogism. Their theories allow prediction, and they change (eventually) as a result of falsifying evidence.¹⁵

Gopnik conducted experiments in which children answered questions about what other people are thinking, perceiving or intending. She argues that the claims children make about other minds, and the errors they make, have the sort of systematic, predictive and revisable character that belongs to scientific theories. The child is not directly responding to the perceptual data, but is interpreting the data in light of a hypothesis; fundamental shifts in the child's claims are not due to changes in the perceptual data, but to shifts in the theory. In sum, the child's experience of

other minds is neither immediate nor perceptual, it is mediated by an at least implicit theoretical comprehension (the language of mediation and immediacy is Gopnik's). This position is known as the "theory theory of other minds": a theory of other minds in which possession of a theory of mind is fundamental to overcoming the gap between one's own mind and another's mind.

Put in Hegel's terms, according to Gopnik, the child's consciousness of other minds is a form of theoretical (rather than perceptual) consciousness. What I wish to show is that in fact this is a form of self-consciousness stabilized by living interests, by the infant's desire. I do this by showing how the infant theoretician's knowledge of other minds is untenable if configured as in "the first supersensible world" discussed above, and can only stabilize in an "inverted world," a world-view in which the phenomena already reflect the infant's interests and consciousness.

To begin, let us note that in Gopnik's "theory theory" perceptual data prompts the child's application and revision of a theory of mind. In order to comprehend the perceived behaviour of complex "bags of skin," the child hypothesizes something that is not perceived, but theoretical, namely the mind of another person. As in Hegel's analysis, the data of perception pose a problem that demands further comprehension: the very diversity of the perceptual phenomena compels the child to attend to a comprehensible, supersensible, theoretical unity, not a direct observable.¹⁶ On this assumption, the child's experience of failing to comprehend the perceptual data is what prompts a revision of its hypothesis. This defines the dynamic of the understanding that we are studying.

Consider the child at the dinner table, trying to comprehend perceptual phenomena such as smiles and frowns, that is, regular

patterns in “bags of skin.” We can imagine the child acting on the hypothesis that the smile-phenomenon is explained because it manifests something supersensible, called approval. We do not need to claim that the child explicitly and self-consciously *thinks* about this hypothesis, for the child who merely *acts* on this hypothesis will experience the diversity of smile-phenomena as in each case explained by and manifesting a comprehensible unity called approval. That is, we can consider theoretical consciousness in terms of the acts and experience it generates, although below, to simplify the exposition, I speak of the child as explicitly thinking about theoretical terms. For the child, what I am here calling approval is a phenomenon having the structure of what Hegel calls a force: it is the manifest expression of a supersensible comprehensible unity. Approval is manifest in the smile in the way that vitreous electrical fluid is, for Dufay, manifest in charged glass, and the child is pursuing the explanatory strategy of the first supersensible world.

But people do not just smile, they do other things with their faces. And to comprehend approval is to comprehend contrasting expressions. Not only does the child have to comprehend all instances of smile-phenomena in terms of approval, she has to comprehend approval as manifest only in smile-phenomena, not other facial gestures, such as frown-phenomena. Further, smile-phenomena and frown-phenomena are interrelated: smiles and frowns spread from face to face, or lead to their opposites. The child cannot stop with comprehension of smiles-as-approval, she has to comprehend an interrelation between different instances of approval, and between approval and disapproval. The child must comprehend a unified network of expressive phenomena. This is Hegel’s point that forces are necessarily manifest as a unified sphere of

interrelated forces. (M150-151, W/C 105-106)

But sometimes we express the opposite of what we really feel: for example, we smile even if we disapprove. This would not be apparent to a child who comprehends smiles as isolate expressions, but it will be apparent to one who comprehends expressions as a unified network. Against that comprehensive background, a reversal of expression is confusing. Either the child’s way of comprehending the interrelation of approval and disapproval is wrong, and the child cannot explain the phenomena, or it is not possible for the child to directly read back from the sensible phenomena (such as smiles and frowns) to a comprehensible, supersensible unity (such as approval and disapproval). In either case, if the child persists in the strategy of reading back from the phenomena to a supersensible explanans, then making sense of the confusion requires the introduction of another, hidden variable.

The child could act on the hypothesis, for example, that some “bags of skin” have a sweet, positive character and others have a sour, negative character. (I introduce the terms “sour” and “sweet” here to help illuminate Hegel’s usage of these terms in paragraphs such as M158, W/C 111-112.) Thus a sweet character may smile even though disapproving of something, etc. Character, by virtue of its role in the child’s explanatory strategy, is supersensible. If it were sensible, it would be part of the problem, another possibly confusing expression, rather than a supersensible factor that helps clear up confusion. But, for similar reasons, if character is a supersensible thing that sorts out confusions about other supersensible terms such as approval or disapproval, then *character cannot be directly read back from someone’s expression*. Because of its explanatory role, character is *underdetermined* by the phenomena. In

other words, the child who is trying to sort things out by introducing the contrast between sour/sweet or negative/positive character can no longer pursue the strategy of the first supersensible world, because the register of sensible expressions can no longer determine variables in the supersensible world.

This leads to an obvious problem. What determines the character-value of a given “bag of skin”? Is Georg’s behaviour to be explained by thinking that he is a frowny person who is really sweet underneath, or a frowny person who is playing at being sweet underneath whilst really being sour? How can you tell the difference between a sweet/positive and a sour/negative character, given that a sweet/positive character can become sour/negative, or someone who is fundamentally sweet/positive can at times behave in a sour/negative manner? In facing people like Georg, mightn’t you experience them as having a shifting appearance, first appearing sweet/positive, then sour/negative, and so on?

It is easy to grasp what it would be *like* to be conscious of the world in this way, especially as a child. Do they really approve of me? How am I to make sense of the conflicts between what they appear to approve and what they really approve? What is their character, really? They keep changing their minds; I do one thing one day and they’re happy about it, I do the same thing the next day and they’re angry. How can I tell what they’re really like? We have all been in something like these familial shoes, I’d wager, perhaps that is why they bronzed them: congratulations, you figured it out. But how did we figure it out?

The child here is in something like the position of the scientist who realizes that a positive-negative type contrast needs to be deployed in explaining the phenomena, but also realizes that the valuation/assignment of positive and negative sides might vary with

the phenomena, for example, might differ in thinking about batteries versus transistors. There is a difference of course, for in our example, it is, as it were, hard to tell if Georg is a “battery” or a “transistor” and the problem is that he can switch between the two. Moreover, scientists belong to communities which cut through such problems by invoking a rational criterion enjoined by consciousness itself, for example, Ockham’s razor, a principled, rational demand for economy in explanation, a demand that does not stem from the object, but from the interrelation of subject and object as developed by a community that has procedures for agreeing upon standards. But the child does not have such a principle, or at least is not ready to grasp it or justify it as a member of a rational community. More, the child is not quite a member of such a community, for what we are looking at now concerns how the child learns to become a member of any community at all in the first place.

What the case of the child lets us see is that what stabilizes the child’s theorizing is not a principled, rational demand, but rather a more basic demand: the simple *need* to have some sort of stable comprehension of “bags of skin.” Imagine a child vacillating between different attitudes toward Georg’s character. The child’s storm of indecision is experienced by the child as in fact storming across Georg’s face: Georg sits there, across the table, face unchanging, at first looking to the child sweet/positive, then sour/negative, and so on, as the child’s view of Georg changes. The problem is that this storm offers no certainty, security or comprehension, quite the opposite. What calms the storm is the child’s consciousness that *she is responsible for it*. She realizes it is not simply Georg’s expression that is changing Georg, but *her* view of Georg. The other part in calming the storm is that Georg *himself has* a character and takes

responsibility for it: if Georg's expressions were entirely arbitrary and irrational, if he did not work to have them express something like a steady character that can be mapped in terms of contrasts such as sweet/sour or positive/negative, there would be no way for the child to comprehend Georg. Gopnik, Meltzoff and Kuhl admit that the view of other people as "bags of skin" is a madman's view of the world, but if people around the child did not take responsibility for their expression—trying to teach expressions as expressing character so that the child may learn about expressions and character—then the child would be living in a mad world and would go mad.

We can say, then, that the child reaches an understanding of Georg's character only because she is engaged in an interested¹⁷, interpretative act that overarches (*übergreift*, M166, W/C 122) Georg. She is not simply engaged in a reading that is propelled backward from the phenomena to the hypothesis of a supersensible character behind the scenes, she is engaged in an activity of making sense of the phenomena by reading character *in* the phenomena, and she can do this because the phenomena already conform to this kind of reading. The child, like Holmes, is living in an "inverted world" in which perceived objects, as it were, wear their intelligible character on their sleeves; but whereas Watson has the option of living so as to not be able to read the intelligibility that Holmes finds in things, the child who cannot read other bodies as intelligible others is living in a possibly mad world.

All of this, however, conceptually entails that the problem of other minds, as treated by the "theory theory," is badly put. There never is a stage in which all the child encounters is "bags of skin"; and her encounter with something more than that is not mediated by the simple addition of a theory to the facts. If the child is ever to

make sense of them, the "bags of skin" themselves already have to appear as mindful bodies who are mindful of the child. The child's ability to learn about others depends on the child's desire to learn, a desire geared to the adult's desire to teach. Or, equally, what leads to understanding is not a purely theoretical process divorced from its object, but a pre-theoretical need to understand that is satisfied only by taking responsibility for the way the object is comprehended, which is possible only if the object itself is responsible to comprehension, if it works in such a way that probing it, questioning it, working with it, provokes responses that fit with our ability to comprehend. In this respect the analysis resonates with classic phenomenological criticisms of the problem of other minds.¹⁸

In other words, the need to explain is stabilized and satisfied by something pre-theoretical. Gopnik, in other research, is aware of this. In fact, Gopnik suggests that explanation is driven by something that might be called desire, even going so far as to talk about "explanation as orgasm," in the sense that the satisfaction found in explaining things is akin to the satisfaction of an orgasm.¹⁹ Gopnik's point is that without some kind of drive for satisfaction, theorizing would be rudderless. Gopnik, however, in effect eliminates what Hegel calls self-consciousness from the drive to explain and from the child's pre-theoretical relation to others, by conceiving both in terms of evolved capacities, that is, in terms of faculties of consciousness, a kind of possession of the subject. For Hegel, instead, self-consciousness is to be understood as a dynamic and open ended relationship that is not a possession, but a dialectic, a process. In fact, Hegel's analysis of understanding would imply that no possession of the subject could stabilize the theoretical circles that have repeatedly

arisen in our discussions above, for no internal criterion posited by the understanding is univalent or stable; in terms of concocting theoretical structures, paranoid and baroque systems of explanation are always possible, and what cuts through these and gives us a comprehensive relation to an intelligible world is *relational*, a relation to a world that is intelligible and reflective of our pre-theoretical needs and desires.²⁰ The understanding only finds theoretical ballast in its pre-theoretical relation to something beyond it. Coming back to recent science we could say that we, after all, evolved in this world and universe, and our evolved ecological relation to our world and our developmental milieu is part and parcel of our theoretical ability. Put in stronger form: we could not know, in any robust sense, a universe in which we did not grow up or evolve. Put in more everyday form, to gain knowledge in new situations, we grow into them. Even more bluntly: knowledge doesn't just run on theorizing, but on growing up and learning, becoming responsible to things in multiple dimensions.

What makes theoretical consciousness possible, then, is a form of life, a pre-theoretical institution, that gives something like an Ockham's razor bred in bone, a razor necessary because a child, before becoming a member of an explicitly self-conscious, rational scientific community, needs to relate to others as others. A child who did not act with the implicit pre-theoretical conviction that the dinner table situation must make sense, and that therefore distinctions that make sense to the child (such as sweet/sour or positive/negative character) must apply to the world, would go mad. And children can be driven mad if they grow up with dinner table situations that refuse to allow such distinctions to apply.

3) Conclusion

Understanding, Hegel shows, is not a merely theoretical act, nor is it an act of consciousness over against an utterly alien object. Understanding is an act in which we identify with our object, and are conscious of what is at stake in our own consciousness—identify ourselves—through our identification with our object. It is also at the same time necessarily an act in which we differentiate ourselves from our object, for it is only by being conscious of something different than ourselves that we settle our understanding and identity. Put in terms of the “resistance point,” the resistance point that stabilizes explanation could neither be in consciousness on its own, for consciousness cannot resist itself; nor could it be in the object on its own, for independent of a criterion of consciousness the phenomena manifest by the object can be explained in many ways. Rather, consciousness resists itself with and through its object, which is to say it is self-consciousness.

The analysis of Gopnik's “theory theory” of other minds lets us see how this relation to the object is structured by something pre-theoretical, by a form of life that has its own imperatives and desires. The infant in the crib is a kind of scientist, but her scientific grasp of the human world is first of all stabilized by needs, imperatives and desires, especially the first need of a scientist, namely the need for others. This is the first need for it is only as a member of a scientific community that can confirm, disconfirm, analyze, criticize and test results and claims, that a finite theoretician can ever actually become a scientist who can claim knowledge. The rest of Hegel's *Phenomenology*, amongst other things, proves this last point, and also shows how the imperatives and desires that drive our need for others would, in the dialectic of

mutual recognition, need to be satisfied through the institution of a rational scientific community and the moral and spiritual community that makes such community possible. The important point here, though, is the exposition and articulation of Hegel's point that theoretical consciousness is a form of self-consciousness structured by a pre-theoretical living relation to the world. All scientists are first of all children, scientists in the lab or mind are first of all scientists in the crib, and if we did not first of all, in the crib, live and experience an imperative to grasp the natural and human world, there would be no scientists.

Taking this point about the life of understanding back to Weinberg's initial problem, we saw that the question whether a claim is a description or an explanation is not to be settled by a resistance in the object itself, but by a resistance that we introduce into our process of comprehension. What grounds such resistance points? A claim explains if it comprehensively describes something that matters to, that has resisted, a rational community. This does not mean that explanations are "purely subjective" or that individual objects explained are "constructed." It just means that objectivity itself—what we comprehend by this concept—reflects the life of a rational community. And this does not mean that objectivity is a pure fiction, since life and communities themselves have objective weight, a resistant flow. The basis of rational objectivity is a shared life in a shared world, not a criterion given in and of itself external to our lives. On Hegel's argument, scientific objectivity cannot be established by abstracting from all the situatedness and idiosyncrasies of life and rising to the point of view of a universal reason that would be everywhere the same. Rather, it entails a commitment to nurturing the interpretative practices of the community life that is the resistance point that grounds

our claims about what really matters.²¹ And if that community were already everywhere the same, it would afford no resistance, and would afford no possibility of the phenomenon that we call science, a phenomenon of free and finite individuals trying to comprehend the world.

Hegel's analysis shows that understanding is no mere faculty, but a self-conscious activity oriented by living interests, an activity that surpasses singular interests through the life of a community. Articulated to its fullest, then, the life of the understanding is such that when a scientist defends a claim, she not only changes our view of the world, she changes the community of scientists. But members of that community are not scientists merely, and are not only scientists. Science affects the comprehension and life of a much broader community. This applies to the child-scientist too, for the child's growing comprehension of the world affects the comprehension of those around her. But adult science has a much more radical effect, for there is a greater freedom in it, and a greater responsibility. In our day and age science affects just about everybody and just about all life. Official science, it seems to me, prizes the moment of transforming comprehension of the world, of showing how the actual is rational, and it urges the purity, sovereignty and priority of this moment over the moment of transforming the community of comprehenders. Hegel's point is that the two cannot be detached, that scientific comprehension certifies itself through its recognition in what we all know and do, that the rational is actual. Reason and life cannot be pried apart, so we need a science that does not detach itself from life, but is committed and responsible to its life in the community of reason, that takes explanation not as an abstract practice but as an ingredient in our desire to comprehend our world, a desire that issues into

dimensions of life that are not merely scientific, but social, cultural, historical and moral.²²

¹ References to this work will be given in the form M [paragraph # in Miller's edition, G.W.F. Hegel, *Phenomenology of Spirit*, trans. A.V. Miller (New York: Oxford University Press, 1977)], W/C [page # in Wessel and Clairmont's edition, G.W.F. Hegel, *Phänomenologie Des Geistes*, Hrsg. v. Hans-Friedrich Wessels und Heinrich Clairmont ed. (Hamburg: Felix Meiner Verlag, 1988)].

² H.S. Harris, *Hegel's Ladder I: The Pilgrimage of Reason* (Indianapolis: Hackett Publishing, 1997), chapter 7.

³ For commentaries and analyses of this chapter, the following are particularly helpful: Harris; Hans-Georg Gadamer, "Hegel's 'Inverted World,'" *Hegel's Dialectic: Five Hermeneutical Studies* (New Haven: Yale University Press, 1976); John Russon, *Reading Hegel's Phenomenology* (Bloomington, IN: Indiana University Press, 2004); Joseph C. Flay, "Hegel's Inverted World," *Review of Metaphysics* 23 (1970): 662-78; Joseph C. Flay, *Hegel's Quest for Certainty* (Albany: State University of New York Press, 1984), Jean Hyppolite, *Genesis and Structure of Hegel's Phenomenology of Spirit*, trans. Samuel Cherniak and John Heckman (Evanston, IL: Northwestern University Press, 1974). Also see Michael Baur, "Hegel and the Overcoming of the Understanding," *The Owl of Minerva* 22, no. 2 (1991): 141-58; Martin de Nys, "Force and the Understanding: The Unity of the Object of Consciousness," in Merold Westphal, ed., *Method and Speculation in Hegel's Phenomenology* (Atlantic Highlands: Humanities Press, 1982).

⁴ Steven Weinberg, "Can Science Explain Everything? Anything?," *New York Review of Books* XLVIII, no. 9 (2001): 47-50.

⁵ Gadamer, 43-44.

⁶ In this way Hegel anticipates and influences attacks on the myth of the given, see Wilfrid Sellars, *Empiricism and the Philosophy of Mind* (Cambridge, Mass.: Harvard University Press, 1997) and John McDowell, *Mind and World* (Cambridge, Mass.: Harvard University Press, 1996).

⁷ This problem with force and law is summed up in the beginning of M154, W/C 108-109: "The difference, then, in both cases is not a difference *in its own self*: either the universal, force is indifferent to the division which is the law, or the differences, the parts, of the law are indifferent to one another."

⁸ In terms of contemporary science, e.g., are the strong, weak, electromagnetic and gravitational forces four diverse forces or four aspects of a more fundamental structure? If these forces are determinately interrelated then we cannot say that they are four diverse forces, yet we do differentiate between them. The difference and unity in question is not in things themselves, for what we are differentiating and unifying are theoretical terms, not something present in things themselves.

⁹ There are many expositions of this chapter in the literature, but one of the most lucid and accessible is in Russon, 39-55. Also see Gadamer; Flay, "Hegel's Inverted World"; Harris. The beginning of M150, W/C 105 makes clear that the understanding must get beyond circling between sensuous appearances and the supersensible. This is

the task of M150-157, W/C 105-112, which leads to the inverted world.

¹⁰ The fact that Dufay conceived these sorts of electricity as fluids does not undermine the claim that they are supersensible and theoretical objects. Dufay did not directly observe any sort of fluid at work in the observed phenomena. Rather, we should say that because his theoretical objects are supposed to explain charges that are observed to move fluidly between objects, he conceived his theoretical objects as fluids, without any direct evidence for this conception.

¹¹ See Russon, 40-48.

¹² For more on understanding as “reading,” see Russon.

¹³ Harris, 7; Russon, 40-43.

¹⁴ Alison Gopnik, Andrew N. Meltzoff, and Patricia K. Kuhl, *The Scientist in the Crib: Minds, Brains and How Children Learn* (New York: William Morrow and Company, 1999).

¹⁵ Alison Gopnik, "Theories and Qualities," *Behavioral and Brain Sciences* 16, no. 1 (1993): 44-45, 10.

¹⁶ This is true in Gopnik's scenario even if, as she suggests, the child's theory of mind is ultimately innate. The theory may be innate, but something in the perceptual sphere prompts application (and then development) of the theory.

¹⁷ For a helpful discussion of interest in relation to self-consciousness, see Flay, *Hegel's Quest for Certainty*.

¹⁸ For example, those of Edmund Husserl, *Cartesian Meditations*, trans. Dorion Cairns (Dordrecht: Kluwer, 1991), and Maurice

Merleau-Ponty, "The Child's Relation with Others," in James M. Edie, ed., *The Primacy of Perception* (Evanston: Northwestern University Press, 1964). For a recent survey of criticisms, see Shaun Gallagher, "Understanding Interpersonal Problems in Autism: Interaction Theory as an Alternative to Theory of Mind," *Philosophy, Psychiatry, & Psychology* 11 (2004): 199-217.

¹⁹ Alison Gopnik, "Explanation as Orgasm," *Minds and Machines* 8 (1998): 101-18.

²⁰ A recent theory of child development that places such an emphasis on relationality is to be found in the work of Alan Fogel, e.g., *Developing through Relationships: Origins of Communication, Self, and Culture* (New York: Harvester Press, 1993).

²¹ For a recent argument that arrives at a similar conclusion about economic science, though by different means, see Bernard Hodgson, *Economics as Moral Science* (Berlin: Springer, 2001).

²² This article is based on a paper originally presented at *Force and Understanding: A Conference on Chapter III of Hegel's Phenomenology*, Bishop's University, Canada, 2001. I would like to thank audience members for their perceptive questions at the time, and particularly John Russon and Peter Simpson for their criticisms of my treatment of the inverted world in that version of the paper.