

STANLEY TENNENBAUM'S SOCRATES

Curtis Franks

Plato was not present on the day that Socrates drank hemlock in the jail at Athens and died. Phædo, who was, later related that day's conversation to Echecrates in the presence of a gathering of Pythagorean philosophers at Phlius. Once again, Plato was not around to hear what was said. Yet he wrote a dialog, "Phædo," dramatizing Phædo's retelling of the occasion of Socrates' final words and death. In it, Plato presents to us Phædo and Echecrates' conversation, though what these two actually said he didn't hear. In Plato's account of that conversation, Phædo describes to Echecrates Socrates' conversation with the Thebian Pythagoreans, Simmias and Cebes, though by his own account he only witnessed that conversation and refrained from contributing to it. Plato even has Phædo explain his absence: "Plato," he tells Echecrates, "I believe, was ill."

We look to Socrates' death from a distance. Not only by time, but by this doubly embedded narrative, we feel removed from the event. But this same distance draws us close to Socrates' thought. Neither Simmias nor Cebes understood Socrates' words as well as Phædo did by the time he was asked to repeat them. Even Phædo failed to notice crucial details that Plato points out. Had we overheard Socrates' conversation, we would not have understood it. We look to Socrates' death from a distance, but to understand Socrates, we don't need to access him—we need Plato.

At Stanley Tennenbaum's instigation, Kurt Gödel and Sue Toledo held a series of conversations between March of 1972 and July of 1975. Gödel retired from his permanent position at the Princeton Institute of Advanced Studies the following year and died shortly afterwards. These conversations were among Gödel's last. Toledo's notes of them are fragmented, often cryptic. Rarely are matters pursued at length. Occasionally, Toledo records merely that a theme or a theorem came up and not what was said about it.

One is struck initially by the wide range of topics one finds in these brief notes, from recent results in mathematical logic to the meaning of ancient texts and the subtleties of modern philosophical thought. But more striking is the way that Gödel's discussions of these disparate topics elide into and support one another, even spanning breaks of several months. Tennenbaum called himself a "disciple" of Gödel. He saw mathematics as a testament to the dignity of the human mind, and he saw in philosophy a way to reflect

on the proper cultivation of that dignity, mathematically and otherwise. Both of these visions, he claimed, he acquired from Gödel. They are vivid still, here, in a form we owe to Tennenbaum, among Gödel's final sustained thoughts.

“Phædo”'s classical readers referred to it by the alternate title, “On the Soul.” Phædo appears only in the framing narrative. He is silent while Simmias and Cebes ask Socrates what he means when he says that the poet, Evenus, if he is wise, should follow him as soon as possible in death. He only listens as Socrates explains that his imminent death is not an evil and argues that the soul continues to exist after the body dies. Socrates is discussing the soul. Phædo is in the background.

It is not surprising that Socrates' reasoning is unconvincing to contemporary readers (among the several unpalatable ideas on display is the infamous doctrine that all knowledge is recollection of what was forgotten at birth). The strange thing about the dialog is that Socrates has an unusually hard time persuading his own friends. Plato devotes roughly twenty-five pages to Socrates' attempt to demonstrate the immortality of the soul. The progression is repetitive, cyclic. Simmias and Cebes object continuously. When one of them accepts one of Socrates' points, the other typically does not. Socrates changes his argument a few times. Usually he tries to demonstrate that the soul is immortal by its nature. Occasionally (80e, 81b, 82c) he argues for the seemingly contrary claim that only the souls of philosophers outlive their bodies, because immortality is conditional on one's conduct in this life. When neither of these tactics proves fully convincing, he rehashes some of the images and analogies from the beginning of the discussion in a slightly different order. The result is particularly elegant, but ultimately unpersuasive: Simmias finds more compelling the notion that the soul depends on the body because it is properly understood as a harmonious arrangement of the bodily; Cebes disagrees with Simmias on this point but finds Socrates' metaphor about the cyclic relationship between life and death untenable.

At this point in the dialog, Phædo interrupts the flashback and says to Echebrates:

When we heard what they said, we were all quite depressed, as we told each other afterwards. We had been quite convinced by [Socrates'] previous argument, and [Simmias and Cebes] seemed to confuse us again, and to drive us to doubt not only what had already been said but also what was going to be said, lest we be worthless as critics or the subject itself admitted of no certainty.

(88c)

Echecrates replies that even hearing of the discussion after the fact, he feels that he shares in their despair. Socrates' argument, he says, "was extremely convincing" but has "fallen into discredit." He pleads that Phædo relate Socrates' response precisely, as he is interested both in what new argument Socrates devised at this point and in whether he remained composed in the face of these objections. Phædo says:

I have certainly often admired Socrates, Echecrates, but never more than on this occasion. That he had a reply was perhaps not strange. What I wondered at most in him was the pleasant, kind, and admiring way he received the young men's argument, and how sharply he was aware of the effect the discussion had on us ... (88e–89a)

What did Socrates say that impressed Phædo more than any clever argument he had ever presented? Setting aside Simmias and Cebes' objections, he says, "first there is a certain experience we must be careful to avoid" (89c). Only here does Phædo enter into conversation with Socrates. "What is that?" he asks. "That we should not become misologues," Socrates replies, "as people become misanthropes. There is no greater evil one can suffer than to hate reasonable discourse" (89d). He elaborates:

You know how those in particular who spend their time studying contradiction in the end believe themselves to have become very wise and that they alone have understood that there is no soundness or reliability in any object or in any argument, but that all that exists simply fluctuates up and down as if it were in the Euripus and does not remain in the same place for any time at all. ... it would be pitiable ... if a man who dealt with such arguments as appear at one time true, at another time untrue, should not blame himself or his own lack of skill but, because of his distress, in the end gladly shift the blame away from himself to the arguments, and spend the rest of his life hating and reviling reasoned discussion ... (90b–d)

Socrates eventually proceeds with his investigation of the immortality of the soul. Around 107a–b, the discussion winds down. Socrates has introduced a few novelties in the argument. Largely, though, he preserves

substantial patches of what he said earlier. Cebes announces that he is convinced. Simmias is more cautious: “I myself have no remaining grounds for doubt after what has been said; nevertheless, in view of the importance of our subject and my low opinion of human weakness, I am bound still to have some private misgivings about what we have said.” Astonishingly, Socrates endorses Simmias’ subtle reservation:

You are not only right to say this, Simmias, . . . but our first hypotheses require clearer examination, even though we find them convincing. And if you analyze them adequately, you will, I think, follow the argument as far as man can, and if the conclusion is clear, you will look no further.

Thus unfolds Plato’s treatise on the soul. Socrates’ proof falls just short, but he proceeds to hypothesize about the nature of the existence of disembodied souls and adds that

[n]o sensible man would insist that these things are as I have described them, but I think it is fitting for a man to risk the belief—for the risk is a noble one—that this, or something like this, is true about our souls and their dwelling places, since the soul is evidently immortal, and a man should repeat this to himself as if it were an incantation, which is why I have been prolonging my tale. (114d)

Still speaking, he takes the cup of poison, “and then drain[s] it calmly and easily” (117c). Crito weeps, but Socrates speaks on and everyone gains their composure in time to witness his death.

Socrates’ friends managed to reconcile themselves to his passing. His death was not a bad thing. Socrates’ final words brought them to this understanding, but the conclusion he emphasized was not that the soul is immortal. When Phædo recounted this episode to Echeocrates, he said that Socrates’ intense insistence that reasoned discourse is noble even when not entirely satisfying was his greatest teaching. It allowed them not to succumb to doubt “about what was going to be said” as they had come to doubt “what had already been said.” Socrates even was eager to point out explicitly that his argument was in some respects weak, that although they had taken it as far as possible under the conditions, it would benefit still if its initial hypotheses were further clarified at a later time. Socrates’ proof that the soul is eternal convinced nearly everyone; his admission that he

was not entirely satisfied with it fixed the soul's immortality in his friends' minds like an incantation.

The first block of Toledo/Gödel notes is dated March 3, 1972. The conversation from that day is idiosyncratic in that its only topic, Edmund Husserl—his thought but also his person—does not recur on any later date.¹ And yet the centrality of Husserl in Gödel's thinking is unmistakable here. Gödel and Toledo don't talk about anything else, and in none of the later conversations do they discuss anything with so much care.

Many readers will be surprised to find that Gödel says essentially nothing about that part of Husserl's work most often associated with formal logic, his early writings on logical form and content including *Logical Investigations*. He is interested in a comparably esoteric subject, Husserl's emphasis on *epoché* following a transformative philosophical "discovery" in 1909. In fact, the discovery itself, its nature, interests Gödel even more than the philosophical mode that Husserl adopted in its wake. He calls it the moment of "grasping the system of primitive terms and their relationships." This moment is not unique to phenomenology—Gödel identifies the same transformation in Descartes, Schelling, Leibniz, and Plato. He calls it the definitive moment in the "life" of every "real philosopher." Whatever it amounts to, it defines a thinker's *life*, and not only his or her *thought*. Husserl, Gödel notes, underwent "both intellectual and personal" crises in 1909, and his transformative discovery resolved both.

Toledo and Gödel discuss Husserl attentively. They are interested in the nuance of how his thought progressed more than the content of what he says on any particular occasion. They see Husserl inviting us to come along with him, showing us how to transform ourselves in the most fundamental ways. After 1909, Husserl stopped producing "fully worked out phenomenological investigations." He introduced the stylistic convention of deliberately complicated language that forces us to think about each of his words. He wanted us to "use his experience to get to" where he was. Husserl had attained a kind of understanding that "cannot be transferred from one person to another," so his writing is not an attempt to explain anything to us. Gödel calls it both the understanding of the primitive terms and the understanding of "the absolute." We are blind to this absolute. We are not aware of how we are working with our primitive terms, so we are bound to distort any image of the absolute that Husserl could present. Husserl must force us to

¹There is one exception: The notes from July 22, 1975 begin with the observation that the content of some of Husserl's unpublished manuscripts, which had recently been made available, was uninteresting compared with his published work.

notice our own primitive terms, to diagnose our own ways understanding, “to see what the real reasons for [our] beliefs are.” The transformation that accompanies this self awareness is not conceptual. To emphasize the impossibility of conveying it, Gödel compares it to religious conversion and calls it “almost physiological”—reports that suggest that Gödel believes he has experienced it.

Gödel focuses on this philosophical strategy. It sounds like a first step in a larger program: after completing what Gödel calls our “self analysis of [our] own cognition,” we expect to move on to the adoption of the right conceptual grid, to take off the glasses that have distorted everything and try on the new set that Husserl is offering. But if Husserl was interested in this second phase, Gödel is not. The self analysis, what Gödel says is analogous to psychoanalysis, is the crucial move. We are not aiming at the acquisition of any new facts, not even of “conceptual” facts (There are very few “theorems” in Husserl’s writing, Gödel says). We want to see the “whole world . . . in a different light”—what Gödel calls “insight.” To do this, we don’t need to adopt new primitive terms, we only need to recognize the ones that we are working with and to become aware of how we are working with them. Gödel doesn’t seem committed even to the *possibility* of changing our primitive terms. We want simply to find them, to make them clear. This means that we should become aware of how it is that we use them, how we *in fact* handle them as opposed to how we ought to do so.

This is the key to understanding a potentially disorienting part of Toledo’s notes. Gödel says that after 1909 Husserl’s thinking was Kantian, that his philosophical method was “critical” in the sense associated with Kant’s transcendental idealism. The significance of this is that our primitive terms could be anomalous, out of sync with the noumenal world, moreover that we might not “handle them correctly” in any case. Indeed, it is highly unlikely that our primitive terms are “the right ones” in any deep sense and practically certain that our handling of them is flawed from a transcendental point of view (“it’s not surprising criticism would be needed,” Gödel says, given the fact that we form our ideas in childhood well before the way we work with them can meaningfully be called thinking). The term “*kritik*” is supposed to suggest this: our basic ways of thinking are subject to criticism. But in a statement that is somewhat jarring, Gödel says that he admires the fact that Husserl does *not* supply the criticism that is doubtless warranted. “There is no reason to assume that we always handle [these concepts] correctly,” but notice, Gödel says, that Husserl didn’t mention this. His *epoché* “is essentially an exclusion of criticism, of any concern about truth and falsehood.” We are not supposed to understand the quest for primitives as a quest for

the fundamental features of reality, but “rather as an analysis of the natural way of thinking about” that reality. We don’t want to evaluate that way of thinking—and *epoché* ensures that we can’t—but to become aware of it because we can’t gain insight about our world or our activities without awareness of how it is that we think about the world. Prior to Husserl’s transformation, he might still have considered the possibility that our deployment of primitive terms tracks the form of the world, or at least that it could be made to do so. After his transformation, he realized that assumptions of this sort are careless. But he went even further by insisting that insight is to be attained not by correcting this “problem” but by ignoring it and learning to be concerned about something else.

Gödel criticizes “analytic” philosophy for its insensitivity to this sort of insight. Its practitioners “try to make concepts clear by defining them in terms of primitive terms,” but, Gödel says, they don’t bother “to make the primitive terms clear,” by which he means that they don’t bother to first get clear about how we handle those primitive terms. “Moreover,” Gödel says, “they take the wrong primitive terms.” What kind of criticism is this? It might seem that Gödel is pointing out that analytic philosophers look at the world from the wrong angle, that this is their shortcoming. But this is not what Gödel is saying. He thinks they err by deceiving themselves about how they in fact look at the world. What they take to be primitive terms, the terms in which they try to define all of our concepts, are not actually the fundamental components of our thinking. Perhaps they ought to be, but they are not. Husserl’s understanding of the absolute came with the realization that dissecting concepts according to how they ought to be seen is a distraction from the sort of insight that can transform our thinking and resolve our personal crises. Analytic philosophers err by working with a false self-image, or by thinking that self-understanding is irrelevant to conceptual analysis. So they analyze everything in terms of concepts that are not fundamental for us and are left with no insight into our concepts.

So far I have said very little about the argument in “Phædo” itself, but one detail is pertinent. This is the hypothesis of opposites, a crucial instance of which incited Cebes’ initial objections. Socrates poses it thus:

[F]or all things which come to be, let us see whether they come to be in this way, that is, from their opposites if they have such, as the beautiful is the opposite of the ugly and the just of the unjust and a thousand other things of the kind. Let us examine whether those that have an opposite must necessarily come to

be from their opposite and from nowhere else, as for example, when something comes to be larger it must necessarily become larger from having been smaller before. (70e)

The hypothesis is revived against Cebes' objections and takes on a central (if occasionally cryptic) role in Socrates' discussion throughout the dialog. But where did it come from?

Phædo's account of Socrates' last day begins early in the morning. The Athenian jail guard doesn't let Socrates' friends in initially. When eventually they do enter, Socrates has recently been released from his shackles. Still lying in bed, Socrates addresses Crito. Then he sits upright, "ben[ds] his leg, and rub[s] it with his hand," saying as he does:

What a strange thing that which men call pleasure seems to be, and how astonishing the relation it has with what is thought to be its opposite, namely pain! A man cannot have both at the same time. Yet if he pursues and catches the one, he is almost always bound to catch the other also, like two creatures with one head. I think that if Æsop had noted this he would have composed a fable that a god wished to reconcile their opposition but could not do so, so he joined their heads together, and therefore when a man has the one, the other follows later. This seems to be happening to me. My bonds caused pain in my leg, and now pleasure seems to be following. (60b-c)

We learn immediately after this passage that Socrates has spent his time writing poetry—specifically transposing Æsop into verse—since his incarceration. This activity apparently has played into his current observation, or at least into its preliminary formulation as a principle. When Socrates later turns to discuss the immortality of the soul, he reaches for this principle—presumably not because it is evidently connected with the statement he is trying to prove, but for the plain reason that it has recently been on his mind. He has discussed the soul's immortality in the past, but no one present can recall the arguments from those occasions (73a, 76b, 88c). The "theorem," as Gödel would call it, is known to Socrates, but he hasn't yet found the memorable proof. He has been reading Æsop. He recently saw a familiar experience in a new light. He tries another angle.

There is no record of a conversation between Toledo and Gödel for two years following the Husserl discussion. The next set of notes, dated June 13, 1974, is split between one discussion of David Hilbert's papers from 1928

and another one of Plato’s dialog, “Euthyphro.” The two discussions are not evidently related. The comments on Hilbert’s foundational program are briefer, but thematically connected with the material in the four conversations to follow. The slightly more extended discussion of “Euthyphro,” by contrast, seems initially out of place in the broader context.

Surely the most striking remark recorded here is that Hilbert’s program “was completely refuted,” though by Gerhard Gentzen’s work and not by Gödel’s own results. This is exactly opposite the customary appraisal of these matters. Gödel’s results from 1931 are widely recognized as a refutation of Hilbert’s project, for they show that the consistency of certain precisely delimited mathematical theories cannot be proved using only those same theories’ means, whereas Hilbert had sought “finitary” proofs of the reliability of abstract, infinitary mathematical techniques. Notoriously, in his 1931 paper, Gödel cautioned against drawing this conclusion. Maybe there are perfectly concrete principles of inference, he suggested, of the sort that Hilbert would countenance, that surpass the techniques present in any formal system and that suffice to prove any such system’s consistency.

In 1936, Gentzen presented his work on arithmetical consistency in this light: he had shown how to supplement manifestly “finitary” techniques with principles of transfinite induction so that these combined resources suffice to prove the consistency of any formal theory. One need only take induction principles through sufficiently high ordinals according to the complexity of the theory one is investigating. Can’t such consistency proofs be counted as finitary? “We might be inclined to doubt the finitist character of the ‘transfinite’ induction [through ϵ_0 used in his proof of the consistency of Peano Arithmetic (PA)],” he wrote in *Gentzen 1938*,

even if only because of its suspect name. In its defense it should here merely be pointed out that most somehow constructively oriented authors place special emphasis on building up constructively ... an initial segment of the transfinite number sequence And in the consistency proof, and in possible future extensions of it [to theories stronger than PA], we are dealing only with an initial part, a “segment” of the second number class I fail to see ... at what “point” that which is constructively indisputable is supposed to end, and where a further extension of transfinite induction is therefore thought to become disputable. I think, rather, that the reliability of the transfinite numbers required for the consistency proof compares with that of the first initial segments, say up to ω^2 , in the same way as the reliability

of a numerical calculation extending over a hundred pages with that of a calculation of a few lines: it is merely a considerably *vaster* undertaking to convince oneself of this certainty . . . (p. 286)

Yet Gödel suggests just the opposite, that—far from being a way around the implications of his incompleteness theorems—Gentzen’s work fully refuted Hilbert’s program. Why?

In the notes from July 6 of the same year, Gödel says: “We can try to see how far we can get finitistically in ‘seeing’ transfinite induction. Certainly we can get to ω^2 , perhaps even to ω^ω . This may differ from individual to individual, or depending on training. But Hilbert wanted a proof for everyone, not just for those with special training.” This is a novel idea in the evaluation of Hilbert’s program.² Gentzen had said that it is not evident at what point principles of transfinite induction lose their constructive nature. Gödel replies that this is irrelevant. “One could consider an idealized finitary mathematician, one who could consider completely any finitary process, no matter how complicated. In this case, one *might* be able to obtain an adequate characterization of finitary mathematics,” he says. “We would like to know about this idealized case. ϵ_0 might be finitistic in this case.” But for us even to see the idealized case, “we must introduce abstract concepts.” Thus “this is no help for Hilbert’s program,” he says, “where we have to use the means at our disposal.”

The notes from August 21, 1974 show Gödel approaching the same issue from a different angle. In place of the question “Where in the progression of transfinite ordinals do things lose their finitary nature?” Gödel considers the question of the naturalness of primitive recursive well-orders used to define elementary order-types and notes that “if you allowed your well-ordering to be sufficiently wild, ω^2 could be used for any system” (The force of this observation is compounded by Gödel’s claim that, “for Hilbert, at ω^ω induction would still be finitary.”). Gödel here is referring to the appeal to meta-mathematical notions in definitions of small order-types. A concrete example will illustrate the point (the example is due to Kreisel): First define a primitive recursive predicate $P(x) \leftrightarrow \exists y \leq x \text{Prf}_{PA}(\ulcorner \perp \urcorner, y)$. Then define a binary relation \preceq as follows:

²In chapter 2 of *Franks 2009* a case is made that Hilbert ought to be understood in this way, as insisting on a “proof for everyone.”

$$\begin{aligned}
2n \preceq 2m & \text{ iff } n \leq m \\
2n \preceq 2k + 1 & \text{ iff } \neg P(n) \wedge P(k) \\
2k + 1 \preceq 2m & \text{ iff } P(k) \wedge P(m) \\
2k + 1 \preceq 2l + 1 & \text{ iff } P(k) \wedge P(l) \wedge l \leq k
\end{aligned}$$

Observe that \preceq has order-type ω if PA is consistent and contains a finite sequence bounded by a strictly decreasing infinite sequence otherwise. It is fairly straightforward³ to prove the consistency of PA by an elementary induction on \preceq , hence, since PA is consistent, on a recursive well-order of type ω . However, this well-order is unnatural (Gödel calls it “wild”)—we would never be able to define ω in this way without appealing to our intuition that PA is consistent.

So how ought we distinguish natural definitions from unnatural ones? Not, Gödel thinks, by an analysis of the distinction between meta-mathematics and ordinary mathematics or any other such distinction, but by *self-analysis*, by getting clear about our own finitary constructions rather than investigating the ideal case. Thanks to the juxtaposition of these remarks with those from two years earlier, the influence of Husserl on Gödel’s evaluation of Hilbert’s program is unmistakable. The essential link appears in the course of the discussion of *kritik* from March 3, 1972:

An adequate proof-theoretic characterization of an idealized intuitive evidence (this conception being obtained by giving up this restriction to things *we* can understand) will comprise inferences that are not intuitive *for us* & which certainly allow a reduction of the inductive inferences to an essentially smaller ordinal.

In the remarks from August 21, 1974, Gödel asks whether “Gentzen, in his second paper, still considered it an undecided question as to whether there was a finitary consistency proof for first order arithmetic.” Gödel is referring to *Gentzen 1938*, the crucial passage of which is quoted above. (In Gentzen’s “first paper” *Gentzen 1936*, he argued more forcefully for the claim that his consistency proof was finitarily acceptable. In this “second paper” Gentzen expresses dissatisfaction with that argument, but still clearly suggests that the question is open.) Gödel’s question seems inexcusably cautious until it is understood in this way: Gentzen clearly considered this question undecided in 1938, but he was referring to proofs that were “essentially” finitary as opposed to proofs that were finitary *for us*. This

³One may consult §7.1.9 of *Girard 1987* for the proof and a general discussion.

is what Gödel means in the passage from June 13, 1974, when he says, “what is fundamental in finitism is that things must be able to be given, not that finite collections are being dealt with.” Gödel wonders whether Gentzen could seriously understand foundational programs in terms of such an idealized grounding of mathematical activity, as opposed to a more anthropomorphic grounding in our actual primitive terms. As for the latter conception of mathematical foundations, Gödel sees Gentzen’s work, rather ironically, as being quite decisive. In the notes from July 26 of that year, Toledo wrote: “[I]f we look at [Gentzen’s] proof of induction up to ϵ_0 , we see that it is *the* proof of it. And it is an impredicative proof . . . , which [perhaps from some angle] looks finitary,” but implicitly appeals to abstract notions. Gödel means that once we understand this proof, we recognize both that it is the natural proof and that it exceeds our actual ability to work with things given to us in intuition. “That Hilbert’s goal was impossible became clear after Gentzen’s method of extending finitary mathematics to its utmost limits,” he proceeds to say, because this method made evident that the natural proof of PA’s consistency exceeded those limits.

In the dialog “Euthyphro,” Plato dramatizes a possibly fictitious encounter between Socrates and a priest named Euthyphro. The two meet outside the king-archon’s court, where alleged affronts to the Olympian gods are adjudicated. It is clear that Euthyphro and Socrates know one another, though their exact relationship is left vague.

Euthyphro is surprised to find Socrates here, given his habit of remaining aloof from civil matters. Socrates explains that he is not here on his own initiative but has been charged by a man named Meletus with harming Athens by spreading heretical ideas. (These are the charges on which Socrates will soon be convicted, imprisoned, and eventually executed.) He has come to the court for the preliminary stages of his hearing. Euthyphro is aghast. Socrates, he says, far from posing a danger to the state, is “the very heart of the city” (3a).

Euthyphro’s presence outside the court is, by contrast, unremarkable: his religious station makes his testimony relevant to court procedures. All the same, Socrates asks him about his current court business. Euthyphro explains that he is pressing charges of murder on his father because of an episode of negligence that resulted in the death of a former house servant. When Socrates points out that it is considered scandalous for a son to prosecute his own father, Euthyphro replies that he is confident that public opinion is wrong on this point, that he is sufficiently “advanced in wisdom” to see past societal conventions to the fact that justice demands equitable

treatment of relatives and strangers (4b).

Socrates' interest is piqued. In an ironic plea that Euthyphro takes seriously, he says that because Euthyphro is such an expert in divine affairs he would like to enlist under his tutelage so that the charges being brought against him might be deflected to his new official adviser in spiritual matters. Euthyphro agrees to this arrangement without hesitation. Socrates insists that they begin right away and asks Euthyphro to explain to him the nature of *hosion* (ὁσιον = piety, holiness, sacred matters).

Euthyphro first tries to illustrate *hosion* by pointing to his own plan to prosecute his father as an example: "You want to know about piety? I'll show you piety. Watch what I'm doing, these circumstances. Learn piety by seeing a pious person in action!" Socrates says that this isn't how he expects to learn things and demands instead a "formal" definition. Euthyphro, seemingly reluctantly, agrees to try this out: "If that is how you want it, Socrates, that is how I will tell you" (7e). He first tries to define *hosion* as "what is loved by the gods." Socrates finds a problem with that definition, in that the gods of the day had competing loves. So Euthyphro amends his definition slightly by saying that *hosion* is what all the gods agree to love. This leads into the famous causal dilemma: Does the gods' mutual love of a thing make it pious, or does the piety of a thing earn the gods' love?

This dilemma is easily associated with Plato's alleged ontological doctrine that an abstract phenomenon's nature is uninfluenced by an agents' knowledge of or decisions about it. Euthyphro's attempt to define *hosion* in terms of the gods' love doesn't do justice to its alleged eternal "form." If whatever the gods agree upon as pleasant thereby becomes holy, then indeed holiness does not *have* an eternal form. It must rather be, Socrates argues, that a thing's holiness attracts the gods' love of it. If so, then defining *hosion* as "what the gods love" is like defining it as "what you and I are talking about right now"—one has identified merely an "affect or quality" of holiness, not its essence (11a).

Not equipped to navigate the causal dilemma satisfactorily, Euthyphro proposes a third definition, namely, that *hosion* is part of justice. Socrates presses Euthyphro into specifying which part of justice it is, and Euthyphro says that it is a kind of trading relationship with the gods. We receive from the gods, so we return this favor with certain behaviors. Socrates asks what the gods receive from us through these behaviors, and Euthyphro begins to speak eloquently about the cosmic harmony brought about by a ritually meticulous and duly reverent society at prayer (14b). But Socrates will have none of this kind of talk and cuts him off, insisting again on a "formal"

definition. “What benefit,” he asks, “do the gods receive from our pious behavior?” Euthyphro replies rhetorically, “Do you suppose, Socrates, that the gods are benefited by what they receive from us?” (15a). Socrates accepts this and then asks what piety’s purpose is, if not benefit. How is it repayment for the good that we receive from the gods? Euthyphro says that the display of honor and reverence is pleasing to the gods. Socrates then points out that this claim, understood formally, is just a repetition of their earlier failed attempt to define piety: piety is what pleases the gods. They have covered no ground.

After each of Euthyphro’s failures to sufficiently explain *hosion*, Socrates has urged him to collect himself and to try again so that he can become Euthyphro’s pupil and absolve himself of the charges he faces. Until now, Euthyphro has been a relatively good sport, trying out new angles despite Socrates’ commanding refutation of all his ideas. But this time Euthyphro runs off, somewhat impatient.

One who would decipher this dialog faces a strange amalgam of earnestness and irony. The charges brought against Socrates are grave, and in “Apology” and “Phædo” it becomes clear that they are not just idle threats. Socrates’ reaction to them is befuddling. He says that he would like to deflect the charges onto Euthyphro. Is there really any possibility of doing this? If not, then shockingly, Socrates would appear to be joking at the least appropriate time. On the other hand, if this strategy is viable, then it is equally odd that Socrates would make the arrangement conditional on Euthyphro’s demonstration of expertise. Shouldn’t the priest’s reputation suffice? Moreover, if these terrible charges can so easily be deflected, then it is puzzling why Euthyphro, who evidently is not equipped to define piety in a way that holds up to scrutiny, would agree to the arrangement so unhesitatingly. It’s fairly clear that Euthyphro didn’t take the exercise seriously to begin with and that Socrates never expected Euthyphro to define anything satisfactorily.

What, then, does Socrates hope to accomplish with these antics as his fateful trial draws near? At 5d and again at 6e Socrates asks Euthyphro whether he agrees that *hosion* “presents us with one form.” Both times Euthyphro agrees. In the bulk of the dialog that follows, we see how ill-equipped Euthyphro is at specifying that “form.” Socrates expects this display of ineptitude. His purpose is to expose the ignorance behind Euthyphro’s pretension to wisdom. What seems like a curious preoccupation on the eve of one’s heresy trial is itself, for Socrates, pious behavior—to “go around seeking out anyone, citizen or stranger, whom I think wise . . . [and] if I do not think he is [to] come to the assistance of the god and show him

that he is not wise” (“Apology” 23b).

Euthyphro’s reception of this treatment is noteworthy. After stumbling over the causal dilemma the first time, he says, “I have no way of telling you what I have in mind, for whatever proposition we put forward goes around and refuses to stay put where we establish it.” Socrates replies, “. . . [i]f I were stating them and putting them forward, you would perhaps be making fun of me and say that because of my kinship with [Dædalus] my conclusions in discussion run away and will not stay where one puts them. As these propositions are yours, however, we need some other jest, for they will not stay put for you, as you say yourself.” To this Euthyphro says, “I think the same same jest will do for our discussion . . . for I am not the one who makes them go around and not remain in the same place . . . ; for as far as I am concerned they would remain as they were” (11c–d). When, at the end of the dialog, Socrates points out that their discussion has cycled back to the same conundrum, Euthyphro throws up his hands and leaves. Socrates takes note. When the “form” he seeks to disclose proves elusive, Euthyphro seemingly concludes “that there is no soundness or reliability in any object or in any argument, but that all that exists simply fluctuates up and down as if it were in the Euripus and does not remain in the same place for any time at all.” He doesn’t “blame himself or his own lack of skill but, because of his distress, in the end gladly shift[s] the blame away from himself to the arguments.”

It is well known that Gödel defended a sort of “Platonism” about mathematical truth, a thesis that meaningful mathematical talk is explicable only by there being a mathematical reality whose details don’t depend on our ability to determine them. In *Gödel 1964*, Gödel considered the possibility that “Cantor’s conjecture” (that the continuum has the least possible cardinality greater than the cardinality of the set of integers) might be independent of the standard set-theoretical axioms and remarked that

a proof of the undecidability of Cantor’s conjecture from the accepted axioms of set theory . . . would by no means solve the problem. For if the meanings of the primitive terms of set theory . . . are accepted as sound, it follows that the set-theoretical concepts and theorems describe some well-determined reality, in which Cantor’s conjecture must be either true or false. Hence its undecidability from the axioms being assumed today can only mean that these axioms do not contain a complete description of that reality. (p.260)

In his “Gibbs lecture” he reaffirmed his belief in a mathematical reality independent of human conventions and behaviors: “the objects and theorems of mathematics are as objective and independent of our own free choice and our own creative acts as is the physical world” (*Gödel 1951*, p. 312).

It is thus remarkable that in Gödel’s probing discussion of “Euthyphro” in 1974, he says nothing about the dialog’s ontological implications. Instead, Gödel reads “Euthyphro” as a call to wrestle religiosity and moral conviction from the dictates of authority. He associates the idea that the gods not only command us to do what is right but that their commands actually determine what is right with “orthodox religion.” Orthodoxy, Gödel suggests, confuses expertise with authority. We look to experts for answers and advice with good reason, but institutions take advantage of our tendency to trust experts: they transform our reasonable habit of turning to experts as *our source for facts* into the scandalous idea that their opinions are *the source of truth*. Orthodox religion arises in the wake of this transformation and thus struggles to conceive of a deity other than on this authoritarian model, i.e., as investing particular behaviors with moral worth by commanding their performance. The usual grounds for objecting to this theology is that such gods’ commands can defy reason, and we are therefore expected to defy ourselves in so far as we are expected to heed those commands. Better, a theology where the gods’ commands direct us to patterns of living and thinking that we can rationally appreciate. This is what Gödel calls “rational religion.” He reads the causal dilemma about the relationship between what pleases the gods and *hotion* as fundamentally an argument between rational and orthodox religious thought.

Gödel sees this very same dilemma appearing a second time in the dialog, this time concerning our duty to the state instead of our duty to the gods. The orthodox view is that whatever the civil authorities deem is *de facto* what you are obligated to do for the state. The rational view is that the civil authorities are beholden to some “exterior” facts about what sort of behavior among its citizens would be good for the state, and are thereby obligated to enforce this, rather than just whatever they please.

According to Gödel, the dialog between Socrates and Euthyphro is about how to break away from the authoritarian view of morality. Both speakers agree that the break is needed. Athenian society has established an implicit taboo on prosecuting one’s own father. Socrates reminds Euthyphro (ironically) of this fact when he hears of his purpose in court, trying to elicit a reaction to the apparently common view that such societal conventions are constitutive of right behavior. Euthyphro passes this test and also another: when he learns that Meletus, the representative of “orthodox authority” has

brought charges against Socrates, Euthyphro has a chance to add his voice to that charge but instead accuses Meletus of harming the state.

But Euthyphro and Socrates agree only this far, and Gödel sees the dialog's significance as coming a step after its characters' mutual rejection of authoritarian ethics. When one recognizes that important truths are not true by convention—not by our individual whims, nor by communal consensus, nor even the dictates of recognized authority—the obvious task one sets oneself to is devising a method to discover these truths. The Socratic proposal is to resort to reasoned discussion. But how ought one react when our conversation falls obviously short of its target? For Gödel, *everything* depends on how we answer this second question. Euthyphro reacts with frustration. “I was right,” he might say, if he weren't too impatient to stick around any longer, “to doubt that this would be an effective way to learn anything about piety. Do you honestly think it will be worthwhile to approach this problem again?” Euthyphro can see pretty clearly that a satisfactory definition is beyond their reach. Their initial attempts to produce one have only made the impossibility of the task more evident. Socrates' willingness to press further appears quixotic. But for Socrates, the realization that an eternal truth is beyond one's understanding is the *beginning* of inquiry. New, more delicate questions emerge: “What can I learn about my own assumptions from this discovery?” “How can I avoid succumbing to authority at this point?” “What must I do to keep conversation on this topic meaningful, now that I cannot sincerely hope to learn what I was originally interested in?”

How can Euthyphro so glibly agree to Socrates' request that he stand trial in his stead? Gödel sees this reaction as an act of betrayal, for it makes evident the fact that Euthyphro doesn't expect his discussion with Socrates to be sufficiently conclusive for the arrangement to materialize. What Euthyphro fails to realize is that Socrates is not testing him for expertise. The condition Socrates places on entering a cooperative with Euthyphro isn't that the priest first define piety but that, despite his inability to do so, he remain committed in his stand against authoritarianism. Gödel calls it half-heartedness, the commitment to “ontological Platonism” without an accompanying commitment to what one might call “moral Platonism.” He faults Euthyphro with letting the fact that the nature of piety defies rationalization keep him from defending Socrates, and reason, in court. Authoritarianism reigns not only when everyone is convinced by it, but also when we despair of standing whole-heartedly against it simply because we realize that we can't conclusively explain what the authorities dishonestly claim they control. Had Euthyphro not backed down and had Socrates' friends

joined in the stand against authority instead of trying to devise schemes for Socrates' escape, then, Gödel claims, Socrates' execution would have been prevented.

Thus the significance of Gödel's belief in a "well-determined" mathematical "reality" is not the plain fact that he held this view nor, ultimately, the reasons he gave to support it. What is crucial is that he stressed his ontological Platonism even in the face of systematic incompleteness, and that he was in turn prompted by his moral Platonism to devise new ways to attain a synoptic view of the collective body of mathematical facts rather than to despair of any possibility of doing so. In the conversation of August 21, 1974, he asks whether all mathematical questions can be solved in the logics of transfinite ordinals. Later that year, in the conversation from November 1, he expresses his belief that a sort of completeness result can be obtained for ordinal logics. His own discoveries from half a century earlier could easily dissuade one from pursuing this sort of problem any further, yet in Gödel's hands they seem only to have led to a richer view of the sort of questions that can be asked.

In the last section of the notes from July 22, 1975, Gödel discusses mathematical intuitionism. He claims that mathematics "seems to have found its primitive elements" in intuitions, and he contrasts the classical and the intuitionistic reactions to this discovery. Each school runs blindly with one of the two tenets of Husserl's thought that Gödel appreciates. The classical mathematician "hunts for axioms" using ideas from outside of mathematics. "But," Gödel says, the "axioms" that result "are about mathematical objects." The classical mathematician, then, retains the possibility of "criticizing" his or her methods, of recognizing that they aren't adequate to their subject matter. By contrast, every statement of intuitionism involves reference to "the mind of the mathematician & his ego." The meaning of such a statement "must" therefore "be completely *within* the ego." Thus intuitionism, by disallowing criticism, will more readily accommodate a self-analysis, leading to clarity about our actual intuitions. One must seemingly choose between assuming the critical stance and clarifying our intuitions.

But Gödel advocates the simultaneous cultivation of insight and critical awareness in mathematics. One should neither rest content with "working with ideas that haven't been fully analysed" nor risk flirting with conventionalism. Both mistakes are carelessly myopic, but Gödel's distaste with the latter peril is also vividly moral. In the notes from August 21, 1974, we read, "an unanswered question is: What is really convincing in mathematics. And, can mathematics be reduced to something completely convincing? ... Although Gödel thinks not, so far there is no convincing proof." In its

least critical form, intuitionism is such an attempt at reduction: “Conventionalism is an attempt to reduce to the ego alone & to arbitrary decisions of the ego,” he says the following year. What makes this reduction unconvincing, though, is its moral failing. Toledo writes, “everything is true by (my) convention.” The idea conveyed parenthetically seems to be that a little pressure on such conventionalist lines brings their latent authoritarian connotations to the surface. What makes our primitive terms worth scrutiny, Gödel thought, is not that they are constitutive of anything and therefore beyond reproach, but simply that they are what we have to work with and that we work with them much better when we know something about them. The nobility of the human mind lies not in its role as arbiter in crucial matters, but in its ability somehow to tap into matters that transcend it.

When Phædo arrives in Phlius, Echebrates immediately asks him if he was present at Socrates’ death, as he is anxious to hear about the event. Phædo lets him know that he was there and that he has plenty of time to talk. Echebrates asks for every detail. Phædo prepares to relate the meandering discussion and the dramatic scene. He will start with Socrates’ observation about pleasure and pain, that “a man cannot have both at the same time.” But even before that he wants simply to convey the feeling of the moment:

I certainly found being there an astonishing experience. Although I was witnessing the death of one who was my friend, I had no feeling of pity, for the man appeared happy in both manner and words as he died nobly and without fear That is why I had no feeling of pity, such as would seem natural in my sorrow, nor indeed of pleasure, as we engaged in philosophical discussion as we were accustomed to do—for our arguments were of that sort—but I had a strange feeling, an unaccustomed mixture of pleasure and pain at the same time as I reflected that he was just about to die. All of us present were affected in much the same way (58e–59a)

Thus at the crucial moment when Socrates’ words must be effective, just as he is about to die, his friends share a feeling that undermines his argument. Though it led them that day to see their teacher’s death not as a bad thing, the hypothesis of opposites is false. But they don’t notice. Had they noticed, it would not have mattered. They would not turn to misology. They saw the theorem well enough and already knew that the principles

that it rested on needed further attention. They don't notice, but Plato does.

Works Cited

Franks, C. 2009. *The Autonomy of Mathematical Knowledge: Hilbert's Program Revisited*. New York: Cambridge University Press.

Gentzen, G. 1936. "Die Widerspruchsfreiheit der reinen Zahlentheorie," *Mathematische Annalen* **112**, 493-565. Translated as "The consistency of elementary number theory" in *Szabo 1969*, 132–213.

Gentzen, G. 1938. "Neue Fassung des Widerspruchsfreiheitsbeweises für die reine Zahlentheorie," *Forschungen zur logik und zur Grundlegung der exakten Wissenschaften*, New Series **4**, 19-44. Translated as "New version of the consistency proof for elementary number theory" in *Szabo 1969*, 252–86.

Girard, J. Y. 1987. *Proof Theory and Logical Complexity. Vol. 1*, Napoli: Bibliopolis.

Gödel, K. 1931. "Über formal unentscheidbare Sätze der *Principia Mathematica* und verwandter Systeme I," translation by J. van Heijenoort as "On formally undecidable propositions of *Principia Mathematica* and related systems I" in K. Gödel, *Collected Works, Vol I: Publications 1929–1936*, edited by S. Feferman, et al. New York: Oxford University Press. 1986. 144–195.

Gödel, K. 1951. "Some basic theorems on the foundations of mathematics and their implications," K. Gödel, *Collected Works, Vol III: Unpublished essays and lectures*, edited by S. Feferman, et al. New York: Oxford University Press. 1995. 304–323.

Gödel, K. 1964. "What is Cantor's continuum problem?," in K. Gödel, *Collected Works, Vol II: Publications 1938–1974*, edited by S. Feferman, et al. New York: Oxford University Press. 1990. 254–270.

Grube, G. M. A (trans.) 2002. *Plato: Five Dialogs*, second revised edition. Indianapolis: Hackett.

Szabo, M. E. (ed.) 1969. *The Collected Papers of Gerhard Gentzen*. London: North Holland.

Toledo, S. "Sue Toledo's Notes of her conversations with Gödel in 1972–5," edited by J. Kennedy. This volume.