For Marcus Rossberg and Philip Ebert, eds. *Abstraction*, Oxford University Press forthcoming.

Abstraction and Epistemic Entitlement:

On the Epistemological Status of Hume's Principle

CRISPIN WRIGHT

§1 The abstractionist programme of foundations for classical mathematical theories is, like its traditional logicist ancestors, first and foremost an *epistemological* project. Its official aim is to demonstrate the possibility of a certain uniform mode of a priori[1] knowledge of the basic laws of arithmetic, real and complex analysis, and set theory (or as much set theory as anyone might soberly suppose to be indeed knowable at all.[2]) It is a further issue whether a successful execution of the abstractionist project for a particular branch of mathematics would amount to a local vindication of logicism in some interesting sense of that term. Traditional logicism aimed to show that mathematical knowledge could be accomplished using only analytic definitions and theses of pure logic and hence is not merely a priori if logic is but is effectively a *proper part* of logic. Abstractionism, however, adds abstraction principles to the base materials employed in the traditional logicist project—principles that, at least in the central, interesting cases, are neither pure analytic definitions[3] nor theses of pure logic as conventionally understood. Thus, whatever significance they may carry for the prospects for logicism, in one or another understanding of that doctrine, the epistemological significance of technically successful abstractionist projects must turn, one would suppose, on the epistemological status of the abstraction principles used, with any demonstration of a

---

[1] I write in the context of something of a renaissance of scepticism about the theoretical interest of the notions of a priori knowledge and justification. I do not think the grounds for this renaissance are strong, but I will not challenge them here.

[2] My own, probably minority, view is that due sobriety will exclude the thought that the key principles required in the generation of the Cantorian transfinite hierarchy—Power and Union—are properly regarded as so *known*. But discussion of that must be deferred to another occasion.

[3] At least, not if we restrict the term, "definition", to principles that provide for the systematic eliminative paraphrase of all contexts in which the definiendum occurs. Hume's Principle, for instance, does not provide such resources for all the types of occurrence of the numerical operator required if arithmetic is to be founded upon it.

priority in particular being dependent on whether those principles can themselves rank as knowable a priori even if they are neither definitions, nor truths of logic, strictly understood. My primary focus here will be on this natural thought.

§2 There is, of course, another major question about the significance of the successful execution of an abstractionist project, even for someone who is content that the possibility of a priori knowledge of a branch of pure mathematics would have thereby indeed been demonstrated—a question a version of which confronts any project in reconstructive epistemology. It is the question: what does it do for the status of our *ordinary* mathematical beliefs if, irrespective of how we actually arrive at them, some philosopher is able to work out an ideal route—very different to anything we actually do—and a persuasive case that *if* we were to follow it, we would indeed wind up with (much of) the knowledge that we think we actually have in the relevant region? After all, one may in fact possess only a posteriori knowledge of things that can be known a priori. Can considerations be marshalled to make a case that successful abstractionist constructions can somehow "rub off" on the status of, for instance, our actual arithmetical knowledge? I think this question is long overdue much more explicit attention than it has received in the 'neo-Fregean' literature to date, but I shan't say more about it here, except to remark that it already faced Frege.

§3 I shall concentrate on the best worked out case: Hume's Principle and its credentials as an epistemological foundation for number theory. The issues may seem straightforward: Do we (can we) know Hume's Principle at all? Do we (can we) know it a priori? If so, how? One quite attractive thought, advanced by Richard Heck,[4] is that Hume's Principle, at least when the range of its higher-order variables is restricted to finite concepts, both represents a correct analytic digest of our intuitive conception of the structure of the natural numbers and provides a neat and natural account of how the content of arithmetical statements allows for their direct application in empirical contexts. (The application is provided for directly by the inclusion of appropriate empirical concepts within the range of the higher-order variables of

[4] Heck (1997). Related ideas are, as he remarks in ch. 1 of Heck (2011), developed in Demopoulos (1998) and (2000). See also Demopoulos (forthcoming).

the principle.) (Finite) Hume's Principle thus has a strong case to be a correct codification of everything essential to pure and applied arithmetical thought.[5]

Attractive as this line may be, it puts the cart before the horse as far as the abstractionist project is concerned. For that project, the justification for Hume's Principle cannot turn on its claim to reflect and encode an antecedent body of arithmetical knowledge. Rather its epistemological merits have to be accessible not just a priori but *in advance* of that body of knowledge. In particular, it has to be possible, at least in principle, to *learn* of the truth of the Dedekind-Peano axioms by reasoning based on Hume's Principle.

So, for abstractionist purposes — and assuming we are not, as to his cost was Frege himself, in the market for its derivation from yet more ultimate and basic principles — it seems we need an account of how Hume's Principle might be known *non-inferentially yet a priori* by someone so far innocent of the axioms of arithmetic. One immediately salient candidate account, accordingly, is the proposal that Hume's Principle's a priori credentials are those of a *successful implicit definition*: in effect, a stipulation whose effect is so to fix our concept of the meaning of the sole hitherto undefined term in its statement —the cardinality operator —that the truth of the principle comes to be knowable a priori just in virtue of our understanding of the prior logical vocabulary, our acquired understanding of the content of that expression, and our grasp of the syntax of the statement itself.

Bob Hale and I have previously defended this proposal.[6] Defending it requires making out a connection between implicit definition and a priori knowledge in general. It involves explaining how, in at least some cases, stipulation, or acceptance, of an implicit definition can so constrain the understanding of the very sentence(s) used in the definition that a priori knowledge of the truth of what is expressed is the result; and it then requires defending the more specific claim that Hume's Principle, and other good abstraction principles, rank as *pukka* implicit definitions of this kind.

Here I will revisit and qualify this proposal. Although I continue to have confidence in the idea that implicit definition can constitute one source of basic—non-inferential—a

---

[5] I shall omit the qualification "Finite" in what follows.

[6] Hale and Wright (2000).

priori warrant, I no longer want to rest on the claim that a complete vindication of the possibility of a priori knowledge of Hume's Principle in particular can be accomplished just on that basis. But nor, now, do I think that a complete vindication of the possibility of a priori knowledge of Hume's principle is anyway *required* before knowledge a priori of the basic laws of arithmetic, based on Hume's Principle, may legitimately be claimed. I now reject, in other words, the natural thought noted above that abstractionist foundations for arithmetic, based on Hume's Principle, can have their intended epistemological significance only if Hume's Principle is first known a priori. I will try to make good on that, perhaps slightly shocking, claim.

§4 My confidence in the idea of implicit definition as a possible source of a priori knowledge in general is based in part on a vision of it as located in a wider epistemological setting which I have proposed in other work.[7] Let me briefly outline it here. In his notes *On Certainty*, Wittgenstein writes the following:

> 163. ..... We check the story of Napoleon, but not whether all the reports about him are based on sense-deception, forgery and the like. For whenever we test anything, we are already presupposing something that is not tested. Now am I to say that the experiment which perhaps I make in order to test the truth of a proposition presupposes the truth of the proposition that the apparatus I believe I see is really there (and the like)?

Then, a little later, he observes that

> 166. The difficulty is to realise the groundlessness of our believing.

Compare

> 253. At the foundation of well-founded belief lies belief that is not founded.

And

> 337. One cannot make experiments if there are not some things that one does not doubt. But that does not mean that one takes certain presuppositions on trust. When I write a letter and post it, I take it for granted that it will arrive— I expect this.
>
> If I make an experiment I do not doubt the existence of the apparatus before my eyes. I have plenty of doubts, but not *that*. If I do a calculation I believe, without any doubts, that the figures on the paper aren't switching of their own accord, and I also trust my memory the whole time, and trust it without reservation.

These remarks are characteristically gnomic but one point that seems salient in them is that to credit oneself with a warrant for a particular proposition acquired by the exercise of a certain

epistemic routine always rationally requires a willingness to make certain kinds of what Wittgenstein is calling *presuppositions*.[8] Central among these presuppositions will be the proper functioning on the specific occasion of certain relevant cognitive capacities (eyesight, memory, and so on) and the suitability of the circumstances for their effective operation (no switching of figures, or forged documentation, etc.) But more, I take it that Wittgenstein is asserting not merely that such presuppositions are unavoidable but that one cannot, in the end, do better than to *take some such things for granted*: the claim to epistemic achievement must rest, in the end, on *groundless* presupposition.

I think it is clear that he is right. That is not to deny that, if one chose, one could investigate (at least some of) the presuppositions involved in a particular case. I might go and have my eyesight checked, for example. But the point is that in proceeding to such an investigation, one would then embark on a further enquiry which would entrain further presuppositions of the same general kinds (that my eyes are functioning properly now, when I read the opthalmologist's report, perhaps with my new glasses on; or that my ears are functioning properly when he tells me of his findings.) It is a necessary truth that wherever I achieve warrant for a particular proposition, —even warrant brutely externalistically conceived, if that is your taste, — I do so courtesy of the proper functioning of my cognitive powers and the amenability of the prevailing circumstances; so whenever I, as I think, get in position to *claim* to have achieved warrant, my claim must rest on my accepting specific presuppositions—about the proper functioning of my cognitive powers and the amenability of the circumstances—for which I will very often have, in context, no specific previously earned warrant. This is unavoidable. I may, in any particular case, set about earning such a warrant in turn—and *that* investigation may go badly, defeating the presuppositions that I

---

[7] Wright (2004), (2004b) and (forthcoming).

[8] *Voraussetzungen*.

originally made. But whether it does or doesn't go badly, it will have its own so far unfounded —*unbegründet*[9] — presuppositions. Again: whenever cognitive achievement takes place, it can be claimed only in a context of *specific* presuppositions, some of which will not be the consolidated product of any cognitive achievement to date.

A first reaction to this point, once registered, is to feel that one's cognitive situation is suddenly extremely precarious. If all claim to epistemic warrant rests on ungrounded presuppositions, haven't we just disclosed the materials for a new and rather ugly-looking sceptical paradox? For presumably our confidence in the things which we take ourselves to have verified in a particular context can rationally be no stronger than our justified confidence in the presuppositions of our having verified anything at all. But now it appears that it will always be the case that some at least of these presuppositions are unevidenced and simply taken for granted. Suppose, for instance, I set myself to count the books on one of the shelves in my office and arrive at the answer, 26. Then presumably the warrant thereby acquired for that answer can rationally be regarded as no stronger than the grounds I have for confidence that I counted correctly, and that my senses and memory were accordingly functioning as required throughout. Yet I will have done nothing, we may suppose, to justify my confidence in these specific presuppositions. Indeed, even if I have—even if I have subjected them to independent check—that independent check will have had its own presuppositions. So there must, it seems, be, as it were, ancestral presuppositions of the original enquiry for which I will have no specific evidence or claim of verification. Yet as soon as one such untested presupposition enters the chain, recognition that it does so—so runs the sceptical thought—should divest me of confidence in the specific enquiry for which it is a presupposition and thereby, in the style of falling dominoes, of confidence all the way

---

9 *On Certainty* §253.

back to the original enquiry. How then can I responsibly claim to have achieved any genuine warrant at all?

There is, of course, a great deal to say about this sceptical line of thought. But for my purposes here I must simply offer the essence of what I take to be the correct line of reply. Since there is, necessarily, *no such thing* as a process of warrant acquisition for each of whose specific (ancestral) presuppositions warrant has already been earned, it should not be reckoned to be part of the proper concept of an acquired warrant that it somehow aspire to this incoherent ideal.[10] Rather, we should view each and every cognitive project as unavoidably involving certain elements of adventure. In the end, I have to *take a risk* on the reliability of my senses and cognitive powers in general just as I take a risk on the continuing reliability of the steering, and the stability of the road surface every time I drive my car. For as soon as I grant that I ought ideally to check the presuppositions of an enquiry, even in a context in which there is no particular reason for concern about them, then I should agree *pari passu* that I ought in turn to check the presuppositions of the check—which is one more enquiry after all—and so on indefinitely. So then there will be no principled stopping point to the process of checking and the original project will never get started. The right conclusion is not that the acquisition of genuine warrant is impossible, but rather that the sceptical anxiety is guilty of naiveté about what getting a warrant properly requires. Warrant may be taken to be acquired whenever an investigation is undertaken in a *fully responsible* manner. Responsibility in general, however, cannot require more than taking all precautions which may reasonably be required. Ergo epistemic responsibility, in particular, cannot, *per*

---

[10] It may be felt that a defence is here owing of the description, "incoherent". Has more been justified than merely "unattainable"? I will not attempt to argue the point in detail here. Suffice it to say that ideals of any kind, once recognized to be necessarily unattainable, can be rationally retained as ideals only when they can exercise some kind of asymptotic approachability — that is, roughly, when "Get as close as you can" is a sensible practical imperative. But "Eliminate all untested ancestral presuppositions" is not in that case: one never gets any closer to its completion, since each testing throws up new such presuppositions.

impossibile, involve an investigation of every presupposition whose falsity would defeat the claim to have acquired a warrant.

The implicit principle that makes the ultimate groundlessness of at least some of the presuppositions (ancestrally) involved in any cognitive project look like a fast track to scepticism is that any acquired warrant is no stronger than the weakest of one's independently acquired sets of grounds for each of its (ancestral) presuppositions. The crucial point I am tabling is that this principle is wrong. The right principle is rather something like this: that any acquired warrant may rightly be regarded as no stronger than the weakest of one's independently acquired sets of grounds for each of its presuppositions—*excepting* those presuppositions to which one is independently entitled *without specific enquiry*.

§5 Very well. But which are the presuppositions to which one is, supposedly, entitled without specific enquiry? And what confers that status on them? It will suit our purposes here to work with one specific notion of epistemic entitlement that I have proposed in other work.[11] First, let's be a little more exact about the relevant notion of a presupposition. Define a cognitive project as a pair consisting of a question and a way of trying to answer it. And let us say that

> P is a *presupposition* of a particular cognitive project if to doubt P (in advance) would rationally commit one to doubting the significance or competence of the project.

Then the relevant kind of entitlement—*entitlement of cognitive project*—may be proposed to be any presupposition, P, of a cognitive project meeting the following two conditions:

> (i) There is no extant reason to believe not-P

and

> (ii) Someone pursuing the relevant project who accepted that there is nevertheless an onus to justify P would thereby implicitly commit themselves to an infinite regress of justificatory projects, each concerned to vindicate presuppositions of its predecessor.

---

[11] Wright (2004) and (2004b).

No doubt that will stand refinement, but the general *motif* is clear enough. If the attempt to vindicate the presuppositions of a cognitive project would raise presuppositions of its own of *no more secure an antecedent status*, … , and so on indefinitely, then we are rationally entitled to—may rationally trust in—the presuppositions of the original project without specific evidence in their favour.

This is a defeasible entitlement, of course. It is lost if a presupposition ceases to satisfy condition (i)— if sufficient evidence emerges against P to justify doubting it. It is also hostage to the standing of the cognitive project concerned. An entitlement of cognitive project is a warrant to trust in P *if* one undertakes the project. There may be epistemic or other reason not to do so. Cognitive projects may be badly conceived—the method of answering may be flawed, or maladapted to the question concerned; or, like any other projects, a cognitive project may be pointless, or dangerous, or wasteful. It is, however a non-negotiable part of our rational nature to undertake enquiry.[12] The alternative is a form of intellectual (and thereby bodily) suicide. That simple consideration, I propose, issues in a *right*: to undertake cognitive projects absent reason to think that they will be ineffective, and absent overriding moral or other practical reason to abstain. Say that a project that meets that condition is *unimpugned*. Then when a cognitive project is unimpugned, and when P is a presupposition of it that meets conditions (i) and (ii) above, it is rationally permissible to undertake the project and rational for one who undertakes the project to trust that P is satisfied.

Let me run through the key ideas one more time. Entitlements of cognitive project are all presuppositions in the sense I outlined —statements, that is, of conditions such that a doubt about their obtaining would be rationally sufficient for a doubt about the competence or significance of the particular cognitive project in hand. What makes such a presupposition into an entitlement for someone undertaking the project, it is proposed, is a combination of three factors: first, that the project is unimpugned — that it is one that the agent has an undefeated right to undertake; second, that no information is possessed which would warrant

---

[12] Enoch and Schechter (2008) defend the interesting claim that enquiry is actually a rational *obligation*. It will, naturally, be supportive of my argument if they are right.

doubt that the presupposition was met; and third, that the attempt nonetheless to verify that it was met would implicate further presuppositions of no more secure an antecedent standing . . . and so on indefinitely. In such circumstances, to run the original cognitive project, and to take its findings on board, is, to be sure, to run a risk—but an unavoidable risk, a risk of a kind that it is part of being a living rational agent to be prepared to undertake and which go with the right to enquiry itself.

§6 If this notion is in good order, the next question is: what is the range of such entitlements? We have in effect already noted two categories of potentially entitled presupposition.

First, in all circumstances where there is no specific reason to think otherwise, the stated conditions will entail that we are normally each of us entitled to take it, without special investigative work, that our cognitive faculties (senses, intellect and memory) are functioning properly, at least to the extent that we would need (ancestrally) to depend upon them to conduct an effective check.

Second, there will be a class of entitled presuppositions concerning the general epistemic cooperativeness of the cognitive environment. My eyes may, on a particular occasion, be functioning well enough yet my acquisition of visually-based knowledge may be frustrated by the character of the background conditions: perhaps the local environment is populated by barn facades, mules cleverly disguised to look like zebras, and hi-tech robotic *doppelgänger* of my friends and colleagues. Or again, my intellect and senses may be functioning properly yet my attempt to acquire knowledge by some routine calculations may be frustrated if they require more written pages of formulae than I can simultaneously attend to and the written figures which I am not attending to mysteriously mutate or disappear. Our cognitive faculties are merely abilities and, like all abilities, their successful exercise depends upon the conducive nature of the prevailing circumstances. That circumstances are appropriately conducive is clearly a presupposition of any cognitive project in the sense defined: to have reason to doubt it in a particular case would indeed be to have reason to doubt the significance or competence of the project in question. There will thus be an entitlement to take it that the prevailing circumstances are indeed conducive to the successful

operation of the cognitive faculties involved in the case of any unimpugned project where there is no antecedent reason to suppose that they are not, and where to attempt to investigate the matter nevertheless would throw up further, no more secure presuppositions of the same sort. In general, such an investigation would indeed expectably have that regressive character. I may, for example, investigate the stability of the figures on the paper, or scrutinise the environment for barn facades, but in doing so I will have to take it for granted that the prevailing conditions are generally conducive to the successful operation of the faculties deployed in these further investigations— that the appearance, sideways on, of a barn façade is not that of some further clever form of illusion, for instance, and that the characteristics of the paper, including possibly mobile figures inscribed upon them, are receptive to check by ordinary vision and memory.

§7 With these ideas on the table, reflect now that there is, plausibly, a *third* kind of entitlement of cognitive project: namely, that we are in general entitled to take it that the concepts in terms of which we formulate a project and its findings are in *good standing*. It's challenging to say what precisely, for these purposes, the good standing of a concept consists in. But there are, anyway, various foreseeable ways in which a concept may *fail* of good standing. It may be, for example, that the introduction of an expression for it into an otherwise consistent language generates inconsistency. Or it may be that it of itself incorporates an internal inconsistency —like Frege's concept of *extension* —or that its application canonically carries implications which canonical grounds for applying it cannot justify —the situation, for example, according to certain sceptical viewpoints, of the concepts of *causation* and *empirical law*. Or it may merely be that the standard explanations of a concept leave it unclear in crucial respects that frustrate its utility. That a particular concept has any of these, or other, failings is something which we can look into and perhaps, if need be, address. The modern development of axiomatic set theory, for instance, represents a systematic attempt to address the shortcomings in the intuitive concept of *set*. But it would be fanciful to suppose that final assurances might be achieved that any particular concept was in definitive good standing. The most that one might hope to do would be to address specific grounds for doubt. And in any case—more important—any investigation of the matter would

presuppose—or ancestrally presuppose—an antecedent conceptual apparatus whose good standing would have to be taken for granted. At some point, then, even the most circumspect of thinkers will have to take a risk on the good standing of a conceptual apparatus in terms of which she exercises her circumspection and carries out whatever tests and safeguards she attempts to impose.

§8 The good standing of relevant concepts is thus potentially a third type of entitlement of project. This is the consideration that connects with the epistemology of implicit definition. To bring out the connection, however, I have to put forward a certain general conception of implicit definition in general.

First, note that there is a sense in which the phrase, 'implicit definition', and the implied contrast with 'explicit definition', is apt to mislead as to the nature of the accomplishment involved when a definition of this kind goes well. An explicit definition introduces an expression for a concept, or complex of concepts, which we are presumed to have independently and which are drawn on in understanding the very statement of the definition. By contrast, there is no presumption that an implicit definition should serve merely to introduce means for the expression of antecedent conceptual resources. Rather, in the most basic kind of case, implicit definition itself *introduces* those very conceptual resources —fits out a recipient with a concept or concepts that they did not have before. We do not, for example, first grasp the concept of (objectual) universal quantification and then come to recognise that it may be implicitly characterised by the stipulation of the normal introduction and elimination rules. Rather we grasp the notion by mastering (informal) practices of inference in accordance with those rules. We are given it by immersion in the practices which those rules encode. Implicit definition is a means whereby we enlarge our *understanding*, not just our vocabulary.

Next we need to reflect that, from one quite natural perspective, implicit definition is the *only* means whereby we can do this. We do not get our concepts in advance, before we have any practice of making judgements involving them. Rather one adds components to one's conceptual repertoire by immersion in, and by thereby acquiring mastery of judgmental practices in which those very components are utilised. Broadly, such mastery involves, at a

minimum, learning what counts as *justification* for—grounds to accept—a range of basic judgements in which a novel concept is involved, and learning what may be taken to be the *consequences* of such a judgement's being correct. The most basic level of mastery of a concept like *pain*, for instance, consists in knowing the circumstances under which it is appropriate to affirm of oneself, or others, that they are in pain, and, in the most general sense, in an understanding of the practical significance of pain —that it is a situation which normally a sufferer will want to change, and that anyone will normally wish to avoid, etc.

It is a matter of familiar difficulty how to divide knowledge of this kind into that which properly belongs with grasp of the concept in question and that which belongs merely with collateral empirical knowledge about it —or even whether such a divide should be attempted at all. However the point remains undeniable, indeed platitudinous, that having concepts is having capacities of judgement in which they are exercised, and that these capacities boil down to an understanding of the conditions under which the judgements in question are appropriate, and an understanding of what hangs on them. All our acquisition of concepts must therefore ultimately be viewed at the acquisition of judgmental competences in which these two basic kinds of grasp are exhibited. Insofar as implicit definition may be thought of as 'definition in use', it is not a poor relative of explicit definition but the canonical means whereby concepts are explained, the very soul of our conceptual education.

Here are the two key points again: first, that implicit definition is properly viewed as a means of conceptual, rather than merely lexical innovation; and second, that all our concepts, even explicitly definable ones, are ultimately grounded in implicit definition —in processes which explain, more or less directly, the proper basis for judgements which configure them and the consequences of such judgements' being correct.

Now, someone might want to protest at this, in a spirit of fastidiousness, that we should restrict the term 'implicit definition' to explicit, *verbal* codifications of such judgmental practices. But we are already familiar with the idea of definition by process rather than statement —ostensive definition is the canonical example —and the epistemological point I am stressing is precisely that what is achieved by the usual kind of, if I may so term them, *explicit* implicit definitions —stipulations that certain sentences

containing undefined terms are to count as true, or the laying down of Gentzen-style introduction and elimination rules—is just the characterisation of practices which, in the more general run of cases, we absorb by immersion and demonstration rather than by description or explicit stipulation.

Here's the upshot. I argued that we have an entitlement of cognitive project to take it that concepts are in good standing, absent specific reason to think the contrary. But our concepts in general are in good standing only if standard basic means for explaining them are effective in establishing them as such. That standard basic means, whether it proceeds by explicit stipulation of grounds for and consequences of judgements which configure a given concept, or whether it proceeds by immersive explanation, as it were, of practices in which those grounds and consequences are acknowledged, is implicit definition. So we are entitled to take it that implicit definition is a generally effective means of conceptual innovation and hence that, in the absence of specific reasons for misgivings, it is productive of concepts in good standing in particular cases. That any particular —statement or process of —implicit definition succeeds in conveying a concept in good standing is a defeasible entitlement.

§9 Consider, then, any statement of an implicit definition. Such a statement may take the form of an explicit placement of conditions on the correct use of a definiendum that it mentions, but the case that interests us is where the definiendum is simply *used*, unquoted, in the context of the statement of the implicit definition—call it *the vehicle*— and where the latter is simply accepted as true, as a correct constraint on the meaning of the expression it serves to introduce. I have just argued that, absent special reason for doubt, we are entitled to regard such a statement as introducing a concept in good standing. But can that entitlement somehow be parlayed into a priori knowledge of the truth of the thought that, when such a concept is indeed successfully introduced, the vehicle will express?

There is a temptingly straightforward line of thought to the conclusion that a priori knowledge of such a thought may indeed be the result. It runs like this. When an implicit definition succeeds in introducing a concept in good standing, its vehicle will be true. And when an implicit definition succeeds in introducing a concept in good standing, we will have accepted it as true— since such an implicit definition only works at all, when it does work,

because we accept it as true. So if we are good at implicit definition—good at this kind of conceptual innovation—there will be a reliable correlation between our acceptances of the vehicles of implicit definition and the truth of the thoughts they express. On a simple reliabilist conception of knowledge, then, our acceptances of the vehicles of implicit definition will generally be knowledgeable. And since they will be acceptances *in vacuo*, uninformed by evidence of any kind, the knowledge involved will be a priori.

Such a conception of basic a priori knowledge will not attract those whose preference is that such knowledge should be, as it were, intellectually self-certifying—should be fully authenticated by processes of internal scanning and reflection. But that preference, it may be rejoined, is already under challenges connected with the fallibility of implicit definition and the theoretical difficulties that confront the attempt to articulate the nature of the putative internal scanning involved and to vindicate its credentials. Hence a purely reliabilist— or perhaps a more sophisticated form of externalist— account of the basic a priori may present itself as something that needs to be taken seriously, at least for the object-language vehicles of definition in general. Is this, roughly, how we should best conceive of the possibilities of a priori knowledge of Hume's Principle?

I shall not here argue against the suggestion that it is. But what is certain is that such an externalist turn is unlikely to appease the critic of abstractionism who needs to be persuaded of the epistemic credentials of abstraction principles. Such a critic will want to be shown that, in the best cases—and Hume's Principle had better count as one of those—such principles are indeed known a priori: he will want a *demonstration* of their epistemic credentials, not a mere canvassing of congenial possibilities. What can be said to address such a critic?

§10 Hume's Principle certainly fits nicely enough with the general conception of implicit definition canvassed in the earlier remarks. The right-to-left direction of its ingredient biconditional offers canonical grounds —the obtaining of one-one correspondences between an appropriate pair of concepts —for accepting a numerical identity, whose canonical consequences are then those respectively required by the standard logic of identity and by the left-to-right direction of the ingredient biconditional. So, according to the argument above,

the default position should be that the stipulation of Hume's Principle suffices to fix a concept in good standing of (finite) cardinality and to bestow sense—or a clear enough sense— on the singular terms which it introduces.

This, though, of course, is only the default position. It is, familiarly, open to a number of challenges. There are perfectly legitimate prima facie concerns over the similarity of the principle to Frege's inconsistent Basic Law V, and over the subtler 'bad company' issues brought out by George Boolos[13] —(not all abstraction principles of this general shape can rank as acceptable stipulative implicit definitions; so which are the good cases, and why?) There are legitimate concerns concerning the impredicativity of the first order variables in the fully expanded version of Hume's Principle, unavoidable if the proof of the infinity of the number series is to go through[14]—(does such impredicativity introduce some kind of vicious explanatory circle, as Dummett for one has suggested?[15]) And there are legitimate concerns whether *enough* has been explained —(the real thrust of the notorious Caesar problem concerns the capacity of the stipulation of Hume's Principle to impart a semantically structured understanding of its left-hand sides and thereby to allow us intelligibly, as is again essential for the construction of arithmetic on this basis, to form open sentences out of them and to quantify into, or otherwise bind, their argument places.) Each of these potential pressure points has quite properly received plenty of attention in the literature on the abstractionist programme. But let us here assume—what I anyway believe to be the case—that the neo-Fregean can win these skirmishes and hence re-establish the default position: our entitlement to regard Hume's Principle as a successful implicit definition of the meaning of the numerical operator and the concept of (finite) cardinal number.

---

[13] Boolos (1990). A sophisticated discussion of the problem has developed in the literature. See Linnebo, ed., (2009) for references and several useful contributions.

[14] This claim was queried by an anonymous referee for this volume, citing a result of John Burgess However the claim is good, and the query confused two kinds of impredicativity. What Burgess shows (2005, pp. 113 and following) is that it is possible to derive the infinity of the objectual domain from Hume's Principle in a predicative second-order logic provided non-standard definitions are given of zero, natural number and successor. Even then, though, the referents if the terms introduced by Hume's Principle need to be taken to lie in the domain of its *objectual* quantifiers. For more on the multiplicity of notions of impredicativity relevant to the abstractionist programme, see Linnebo (forthcoming).

[15] Dummett (1998)

Still, there is a further outstanding obstacle to the idea that a priori knowledge of Hume's Principle might be accomplished on the back of its success as an implicit definition and our general reliability in processes of successful implicit definition. It may be brought out by comparing the stipulation of Hume's Principle with the following stipulation:[16]

Jack the Ripper is the perpetrator of all these gruesome slayings of prostitutes.

Probably no one back in the 1890s heyday of the Ripper's notoriety ever actually explicitly said that. Still a practice emerged in the British news media of the day that accorded with it, and constituted, in effect, an implicit definitional equivalent of it. That practice assigned canonical grounds to claims about Jack the Ripper —to warrant such a claim, one would have to warrant a corresponding claim about the unique perpetrator of the crimes. The assigned logical consequences of such claims were then those proper to any such singular statement, with an appropriate further range of pragmatic and contextual consequences assigned by the association of the name 'Jack the Ripper' with the reference-fixing description, 'the perpetrator ..'. There seems no reason to question that this practice with the name was effective in the sense of establishing a common understanding of it. Clearly, however, the admission of this effectiveness is less than the admission that the practice in question was *well founded in its assumptions*. Key among those assumptions was that the crimes in question were indeed the work of a single individual. That is something that we do not know even to this day. The critical point is thus that even if the Ripper stipulation effectively conveys an understanding of how the name is to be understood, it would be inadmissible unless its presupposition —the existence of a single unique perpetrator of the slayings —is true. If that presupposition fails, the truth of the vehicle of the definition stated above fails with it.

The moral is that an implicit definition's success in fixing a concept is one thing and the holding of the presuppositions of the truth of its vehicle may be something else.[17] In effect, we must look askance at the key assumption of the "temptingly straightforward"

---

16 Arguably not an *implicit* definition, but the point I am about to use it to make is not compromised by that.

17 This important point is well made in Ebert (2005).

transition from successful implicit definition to a priori knowledge of its vehicle: the assumption, as I expressed it above, that when an implicit definition succeeds in introducing a concept in good standing, its vehicle *will be true*. The critic of abstractionism may contend that Hume's Principle is in like case with the explicit definition of "Jack the Ripper"; that the infinity of the series of natural numbers —whose derivation from Hume's Principle is supposed to be the crowning glory of the abstractionist reconstruction of arithmetic —is merely a presupposition of the admissibility of the stipulative statement of Hume's Principle. Our argued entitlement to the good standing of the principle as an explanation of the concept of (finite) cardinal number thus falls short of an entitlement to take it that it is true; a fortiori it falls short of an entitlement to take it that its consequences, and in particular the Dedekind-Peano axioms, are true, —let alone that they can be known by deriving them from it.

§11. Before attempting to address that concern, I want to improve on the implicit definitional conception of the epistemological status of Hume' Principle—indeed, from one perspective, to surpass it.

It will assist us if we first consider another objection. A derivation of the Peano axioms from Hume's Principle cannot — surely? — even in the best circumstances, bestow upon them an epistemic status in any way superior to that enjoyed by Hume's Principle itself. Let us allow, an objector may say, that a case can be made that there is an entitlement to accept Hume's Principle as an effective introduction of the concept of (finite) cardinal number. The objection just noted resists any attempt to translate that directly into an entitlement to accept Hume's Principle as a truth. But suppose it is accepted that there is such a further entitlement. Still, whatever type of epistemic credentials we have for Hume's Principle, those same credentials —it appears —will be the best we get for the Dedekind-Peano axioms on the back of Frege's Theorem. However the content of an entitlement of cognitive project is not, just on that account, an item of *knowledge*. If all I have right now is an entitlement to take it for granted that my eyes are working properly, then that they are may not strictly be regarded as *known* by me. So if what we have for Hume's Principle is, roughly, merely a rational permission to trust in its truth, then rational trust in the truth of the Dedekind-Peano axioms is, surely, is the most we can milk out of Frege's Theorem. The

result is accordingly not knowledge of the laws of arithmetic but a mere license, as it were, to proceed on the assumption that they hold. Such a result would fall disappointingly short of the foundationalist spirit that abstractionism claims to inherit from Frege. But worse, it would amount to the failure of the abstractionist programme for arithmetic, whose target was to demonstrate the possibility of a priori knowledge of Peano arithmetic by supplying a model of how it could be accomplished.

This second challenge —in effect, what I have elsewhere called the *leaching* problem[18]—takes us into some very deep issues in epistemology. Here I will merely offer an indication of what I think should be the shape of the correct abstractionist response, and of the framework in which further debate had best be pursued.

§12. We have recognised three species of entitlement of cognitive project: to the proper functioning, on an occasion, of relevant cognitive faculties, to the conduciveness of the prevailing circumstances to the successful operation of those faculties, and to the good standing of the concepts essentially exercised in formulating the cognitive project concerned and carrying it through. But there is clearly a fourth kind of case. We are also entitled to rely on the soundness of the basic inferential machinery, if any, involved in the execution of the project. Not that, if a rule of inference is challenged, we may not often be able to address the concern. But addressing it is going to involve inference, and, familiarly, very often a seemingly unavoidable reliance upon a principle of inference in a meta-language of the very same pattern as the rule under scrutiny. Since any legitimate concern about the original rule should, manifestly, not be assuaged by meta-theoretic reasoning of the very same pattern, it follows that, at least in cases where we have no antecedent reasons for misgivings about the rule, reliance upon it should be regarded as an entitlement.

There are subtleties here about which exactly are the rules of inference to which we may regard ourselves as entitled in this way, and to what extent a principled demarcation can be made between them and rules the right to use which has to be cognitively earned. But

---

[18] Wright (2004) and (forthcoming).

however the discussion of those matters may go in detail, it's plausible to assume that the rules of *modus ponens* and *conditional proof* as represented by the schematic transitions:

$$\frac{\{A_1.....A_n\} \vdash A \quad ; \quad \{B_1.....B_n\} \vdash \text{if } A\text{, then } B}{\{A_1.....A_n, B_1....B_n\} \vdash B}$$

and

$$\frac{\{A_1.....A_n\} \vdash B}{\{A_1.....A_{n-1}\} \vdash \text{ if } A_n\text{, then } B}$$

will rank as basic entitlements of the intended kind. Notice, crucially, that in saying this, I am implicitly rejecting one historically quite common conception of the epistemological ground of such basic rules. In classifying the acceptance of such rules as entitlements of cognitive project, we affirm (i) their presuppositional status in a given (very large) range of cognitive projects, together with the considerations (ii) that we are possessed of no reason to call their soundness into question and (iii) that, were an attempt to justify them to be made, it would necessarily involve reliance on an inferential apparatus of no more secure an antecedent standing —(actually, in this case, a reliance on the very same inferential apparatus.) But reflect that the last point is simply wrong if there is another, *non-inferential* way whereby the soundness of the rules in question might be recognised. Just that possibility is embraced by the venerable thought that the validity of our most basic rules of inference is given to us by a kind of *rational insight* or *intuition*. I make so bold as to suggest that, for our present purposes, the venerable thought is merely pretentious. Let me explain why.

I am not suggesting that there is no role for some form of non-inferential a priori insight *anywhere* in a satisfactory epistemology of logic and mathematics. But any faculty that enables an agent to recognise truths needs to operate in a context where the truths in question are antecedently understood—you need to understand the proposition that your keys are on the bedside table, for instance, or that thirty-seven is a prime, before you can bring your perceptual, or arithmetical capacities to bear on their verification. So too, in the present instance, a faculty of intuition apt for the recognition of the validity of basic rules of inference, however exactly conceived, would need to go to work in a context where a thinker *antecedently* fully understood the conditional, e.g., but was so far *open-minded* about the

status of, say, modus ponens. Is there any such possible state for a rational thinker to occupy?

Well, not in the view of a certain hard-line kind of inferentialism. On the hard-line inferentialist view, it is *constitutive* of an understanding of the conditional to acknowledge the rule of modus ponens. So an understanding of the conditional cannot coherently be supposed to precede an intuitive recognition that modus ponens is valid. If it could, there ought to be such a thing as understanding the conditional perfectly yet—because of a failure of one's intuitive faculty rather than of understanding—failing to be arrested by the validity of the rule. But, the hard-line inferentialist will contend, there is no such possibility, for no one will *count* as grasping the conditional unless they acknowledge the validity of modus ponens. And that means that here there is here no work for intuition to do—no epistemic space for it to work in. Call this the *squeezing argument*.

Now, Timothy Williamson has recently argued that no such inferentialism can be right.[19] Manifestly, someone *can* grasp the conditional and yet clear-headedly refuse to acknowledge modus ponens. As Williamson observes, Vann McGee is living proof.[20] Indeed, Williamson argues that for any pattern of inference, however obvious-seeming and elementary, sense can be made of the possibility that someone fully understand it yet not be disposed to acknowledge its validity.

I think this line of objection is successful, but only against a needlessly crudely conceived version of inferentialism. The inferentialist should not deny that doubts may rationally be entertained about the validity of what are in fact meaning-constituting inferential rules; and should not deny that commonly explicitly accepted inferential rules may be out of kilter with each other,[21] or with those rules that actually constitute the meaning of a

---

[19] Williamson (2003), (2006) and (2007), chs. 3 and 4.

[20] McGee (1985)

[21] This, incidentally, is the way the McGee cases actually present themselves: they are not pure paradoxes for modus ponens but cases where modus ponens — the elimination rule for the conditional— presents as inconsistent with an unspecified introduction rule somewhat to the effect that a conditional is assertible only if the addition to one's information of the truth of the antecedent justifies the affirmation of the consequent. It is because they violate the latter that the conditionals in which the McGee arguments conclude impress as unacceptable.

targeted logical operator. That a given group of rules of inference are valid is a theoretical claim, and may be called in question by bad though well-motivated theory, or indeed—should they indeed contain some hitherto unnoticed flaw—by good theory. What inferentialism cannot accept is that someone might grasp the conditional and yet have no inclination, even in the absence of any theoretical doubt about their validity or coherence, to practice in accordance with what are *in fact* the rules of inference that individuate that operation.

Now, this qualified inferentialism does not, to be sure, directly support the squeezing argument against intuition outlined above, since it does not so immediately collapse the gap between understanding a targeted operation and explicit acceptance of its constitutive rules. The connection it makes is between understanding the conditional, e.g., and inferential *practice* in accordance with modus ponens —(allowing that subjection to modus ponens is indeed, pace McGee, constitutive of the conditional.) That is not, at least not immediately, to make any connection between understanding the conditional and acceptance of the proposition that modus ponens is valid. A fortiori, it is not to make a connection close enough to squeeze out any role for intuitional recognition of the truth of that proposition.

What does happen, though, is that another objection to the intuitional epistemology now surfaces into view. Consider a chess player who, for some reason, has never explicitly encountered the rule controlling the movement of the Bishop but has, as we say, picked it up by immersion in the practice of the game. Suppose that now, for the first time, she considers a formulation of the rule—say, "From the square it occupies, the Bishop may move diagonally, backwards or forwards, through any number of consecutively unoccupied squares, and may only so move." Her assent to this proposition is to be expected. Indeed she may find the so-formulated rule *obviously* correct. But what she assents to is a proposition whose normative force concerns acceptable practice—that this is how the game is played. However the proposition that corresponds to practice in accordance with modus ponens, modulo performance error, in the way that that proposition corresponds to performance in accordance with the Bishop's rule, is not the proposition that modus ponens is valid but rather the proposition, roughly, that here is how to infer from a conditional, that this is how

conditional inference goes. A reflective assent to that proposition, based on reflection on one's practice, is nothing remarkable. It is of a piece with the general capacity of knowledge we all have of our intentions. Knowing that this is how I play is not the same thing as knowing that how I play is correct. In the case of the Bishop's rule, there is of course no further issue about correctness. But in the case of modus ponens, crucially, there is — it is exactly the further issue of validity.

In sum: when the inferentialist view ceases to be the crude hard-line view—when it holds instead merely that the understanding of the conditional, e.g. is given by mastery of its distinctive inferential role— a gap does indeed open between the understanding and the acceptance of any particular proposition encoding that role. But the new objection to the intuitional account of the epistemology of basic inference is now that there is no prospect whatever of justifying the description of the process that leads to such an explicit acceptance as implicating a recognition of validity rather than merely a becoming explicitly conscious of the nature of one's own inferential practice. No good theoretical motive has been provided for describing the matter in the former way.

I don't expect these remarks to silence the friends of intuition (I doubt if anything will), though I do think they present them with a very significant challenge. What is of interest here, however, is the dialectical situation *if* we now discard the intuitional view. There was already a strong case, prefigured earlier, for saying that an acceptance of the validity of modus ponens is at least an entitlement of cognitive project, one operative indeed wherever conditional inference is part of a cognitive project. But now there is also the makings of a case for saying that modus ponens is a rule of inference to rely on which we have *only* an entitlement of cognitive project —that no superior form of cognitive achievement is here possible.

There are various possible failings —inconsistency, epistemically irresponsible forms of non-conservativeness, etc. —which, in general, an inferential practice may prove to exhibit. And there is, I acknowledge, a strong inclination to say that we *know* that our practice with the conditional is, insofar as we are concerned with features just attributable to the role of the conditional, innocent of such failings. If we get into trouble, we are sure it

won't be modus ponens' fault. But it's hard to see what possible reason we could give ourselves for thinking so that would not variously rely on conditional reasoning. Yet that kind of boot-strapping justification is available for *any* inferential rules. To be sure, for a theorist who wants to construe knowledge in some brutely externalist way, it can still be true that we do, properly speaking, *know* that our rules of inference for the conditional are valid merely in accepting them and in being, in fact, reliable in tracking validity in our basic logical acceptances. But if the question is our right to *claim* such knowledge, then it's hard to see that we are in position to do so; it is hard to see that we are better placed than to claim an entitlement to take it that they are valid. I do not think that we are.[22]

§13 Actually, I do not think this is a terribly surprising conclusion. That basic rules of inference enjoy this status —that of, so to say, mere entitlements, beneath cognitive achievement —is, I think, something which many would have found antecedently quite plausible. What is crucial, though, is that it is not a conclusion that should disturb our right to claim *knowledge on the basis of deductions in accordance with such rules*. That is, even if the consequences of *premises* which are mere entitlements cannot, just on account of their being consequences, enjoy an any more robust form of cognitive status than that, it does not seem that the same limitation should apply to the conclusions of inferences from known premises drawn in accordance with *rules* which we are merely entitled to suppose to be sound. If that were not so, then inference in general would not be a means of extending

---

[22] In a number of papers—see e.g. his (2001)— Paul Boghossian has argued that we need to recognize that warrant-productive inference can, and in certain basic cases, must be *blind*: that reasoning can produce warrant for a thinker for conclusions in cases where it is uninformed by any beliefs she has about its validity, indeed in cases where she may have no developed concept of valid inference at all. I regard this observation as correct and important. But granting that is consistent with rejecting the application of the idea that Boghossian makes to underwrite the suggestion that *rule-circular* inference can be warrant-productive—that someone might, for example, use modus ponens in a blind but warrant-productive proof of its own validity. One problem with this idea is that it is very doubtful whether the range of cases where blind inference can produce warrant includes inferences to *schematically general* conclusions: the sophistication involved in grasping such a conclusion requires, plausibly, a correspondingly general conception of the validity of the reasoning that leads to it and hence explicit beliefs about its validity. Cf Dogramaci (2010). But in any case our interest is in our right to *claim* knowledge of the validity of modus ponens. And where what is at issue is not just the inferential acquisition of knowledge of the validity of an inference pattern but the justification of a claim to knowledge of it, there has to be a presupposition that one is in a position to claim that the inferential machinery deployed is sound. Whether or not knowledge can, *claims* to knowledge cannot be underwritten by blind inference. (For elaboration, see Wright (2004b) at pp. 158-9.)

knowledge —not if all inference must in the end depend on the basic rules, and they are merely entitlements. To be tempted by the thought that inference in accordance with merely entitled rules must correspondingly downgrade the status of its conclusions is to be tempted by a false modesty. If we are entitled to claim that a principle of inference in sound, then we are also entitled to claim knowledge of a statement which we have recognised to follow from known premises by inference in accordance with it. We are not restricted to a mere entitlement to such conclusions.

The point is actually quite general. In order to acquire knowledge, we do not need to know that the cognitive apparatus utilised can and does deliver knowledge. It is enough that it can and does do so. This goes for all presumed knowledge-acquisitive faculties: perception, memory, and reasoning of all kinds. To be sure, ascending a level, to *claim* to have acquired knowledge will require the claim that the cognitive apparatus concerned *has* delivered. But this claim can be entered as an entitlement. It does not need to be known in turn (and it cannot always be known in turn; though it is a fine thing when it can.)

The leaching worry was: if all we have, epistemically, is an entitlement to take it that Hume's Principle is true, deductions from it cannot generate any superior form of epistemic warrant than that. I have outlined two considerations that may be used to address that worry. First, it is arguable that even where basic but (for the majority) utterly uncontroversial principles of logical inference are concerned, all we have, epistemically, is an entitlement to take it that that they are valid. Second, this admission does not disable them from service in the generation and transmission of knowledge. Rather, being entitled to claim that they are valid, we are thereby entitled to claim that they are knowledge-productive.

If this, however, is how it is for modus ponens, it is clearly inappropriate to ask more of Hume's Principle. From a purely proof-theoretic perspective, of course, it makes no difference whether we take Hume's Principle as an axiom in a suitable second-order logic, or whether we take it as a pair of additional rules controlling the introduction and elimination of the cardinality operator. But from the perspective of the epistemology of logic, it makes a big difference. One needs to have a justified claim to know an axiom before derivations from it can justify claims to knowledge of their conclusions. One does not need to have a justified

claim to know that a rule of inference is valid before derivations using it can justify claims to knowledge of their conclusions; it is enough, ceteris paribus, that one is entitled to take it that the rule is valid. This entitlement is then inherited by those claims.

My recommendation, in summary, is that the epistemology of good abstraction principles should be assimilated to that of basic principles of logical inference, and that this involves recognising (i) that their validity is *beneath* knowledge, at least if knowledge in such cases is taken to require some form of reflectively certifiable intellectual processing; (ii) that this limitation is consistent with a rational entitlement to take it that such rules are valid; and (iii) that there is a consequential rational entitlement to take it that they are at the service of knowledge production and extension by inference. If this is accepted, then the proof of Frege's Theorem in a system of second-order logic augmented by rules corresponding to the two halves of Hume's Principle can issue in an entitlement to claim knowledge of the Dedekind-Peano axioms. Moreover since the latter entitlement is generated purely inferentially, deploying only basic rules of deductive inference, the knowledge we are thereby entitled to claim is, if we indeed have it, a priori knowledge.

§14 The preceding perspective, though, is hostage to the unfinished business I left at the conclusion of §10. We have shifted from a deployment of the notion of entitlement of cognitive project that aimed to justify the acceptance of Hume's Principle as a legitimate implicit definition to one that aims to justify its acceptance as a pair of complementary basic inferential rules governing the cardinality operator. (On an inferentialist conception of the conceptual and semantic significance of basic inferential rules, the latter proposal is, of course, perfectly consonant with the former.) However whereas the former ploy was open to the objection that there is in general a potential gap between acceptance of a principle as an effective implicit definition and acceptance of the presuppositions of its truth, the latter ploy is surely no less open to the corresponding objection that there is in general no entitlement to accept new rules of inference as valid whose introductory components implicate novel and untested substantive claims—rules, for example, whose validity carries additional ontological purport not already validated in the language to which they are added. But that is just what an

introduction-rule corresponding to the right-to-left direction of an abstraction principle will do.

Though the epistemological context—that of the conditions for an entitlement to take it that proposed basic inferential rules are valid—is new, this is, of course, an issue that has divided discussants of the abstractionist programme from the start and has already generated a great deal of debate. I do not expect to be able to change entrenched contrary opinion here, but I will give some indication of the way I think the case for the defence should be conducted.

The worry, localised to the case of Hume's Principle, is whether there are, or with what right we take it that there are, any objects to serve as the referents of the numerical singular terms that, exploiting its right-to-left direction, we can enlist Hume's Principle to introduce. In the present dialectical setting, we can set to one side irrelevant general nominalist qualms about admitting abstract objects into one's ontology at all and focus on a doubter whose misgivings specifically concern the apparent ontological presumptuousness of abstraction principles. With what right do we take it that identity contexts of the kinds introduced by abstraction principles can soundly sustain quotidian first-order existential generalisation?

In response, it is notable that there is, in the case of the conditional rules, no terribly impressive corresponding misgiving to be had about the analogous existential generalisations —that is, about the Ramsey-sentences, if you like, obtained by existential generalisation on the places occupied by an expression for the conditional in suitable statements of the validity of the inference patterns licensed by the two rules. Rather, we are, most of us, unshakeably convinced that the acceptance of the soundness of the conditional rules is perfectly warranted; and warrant to take it that the patterns of inference licensed by modus ponens and conditional proof are sound is *eo ipso* a warrant to take that there is indeed an appropriate such function validating those rules.[23] The cautious view — that modus ponens and conditional proof are good *provided* there is any such function — seems merely neurotic. Is

[23] Prescinding, again, from irrelevant nominalist concerns about whether there are any functions at all.

there any basis for such confidence in the case of the conditional rules that goes missing with Hume's Principle, or other good abstraction principles?

We can manufacture a context for a non-neurotic doubt. To suppose that the acceptability of the patterns of inference concerned should await some independent reason to allow that there is such a function, — rather as the acceptability of the practices implicitly defining 'Jack the Ripper' might await independent reason to suppose there was a unique perpetrator of the relevant crimes, —would be sensible if the context were one where, for some reason, an acceptable validation required finding a candidate from within some pre-selected domain of functions to discharge the described inferential role. For instance, the context might restrict us to the repertoire of binary truth-functions that feature in the standard semantics for classical sentential logic; then we might propose to identify the conditional introduced by modus ponens and conditional roof with the most eligible of these (the material conditional, of course.) But in general one's confidence that there is a function that answers to modus ponens and conditional proof does not rest on the possibility of such independent identification of it from within such a preselected, constrained domain. Rather, one would like to say, the rules concerned themselves exhibit the function—they *show* the conditional.

That, roughly, is that I want to say about good abstraction principles and the referents of the singular terms they introduce. They are not to be received in a *reference-fixing* spirit—a spirit whereby Hume's Principle, for instance, is viewed simply as introducing a condition of identity that numbers, if indeed there are any, are required to satisfy.[24] Rather they introduce us to the abstracts concerned in the manner in which—at least on an *abundant* conception of properties— an explanation of the satisfaction-conditions of a predicate introduces us to the property it expresses. There are conceptions of properties— various forms of so-called 'sparse' conception— according to which a doubt can still be entertained whether a predicate with well-defined satisfaction conditions actually latches on to any real property, a doubt whether the world co-operates in putting up a real natural distinction that the satisfaction-conditions in question track. But those conceptions contrast with that on

---

[24] Writers who very explicitly assume that abstraction principles *are* to be taken in such a reference-fixing sprit include Field (1984), Boolos (1997), and Fine (1998)

which, once a predicate has been well-defined, there is no additional element of risk carried by acknowledging that there is a property— a way of being— which the predicate denotes. On the abundant conception, the satisfactory specification of the satisfaction-conditions of a predicate does not merely set a condition which any property denoted by the predicate must meet—viz. that it has to be a property whose instantiation by an object is ensured by its meeting the satisfaction-conditions in question. Rather, such a specification *displays* the property concerned. It leaves no space for an intelligent doubt about whether the world co-operates.[25]

What obstructs this kind of way of looking at the matter in the case of abstraction principles is the assumption that their ontological implications need to be redeemed by reference to some *independently given* population of objects. (Compare the kind of artificial context envisaged above for an intelligent doubt about the existence of the conditional.) That, however, is a gratuitous and, when generalised, quite misguided assumption. There is no requirement that the objects in question should yet be available to thought other than under the very concepts of them that, e.g., Hume's Principle serves to introduce. In order to recognise that there are indeed such objects, it is not required that we hit on some prior range of things, given to us in some other way and so comporting themselves that they are somehow fitted to qualify as the referents of the new numerical singular terms. Rather the sole means of satisfying oneself that there are indeed such objects can be by verification of statements involving reference to them. And indeed when we contemplate the conditions for justified singular thought in general, that *has* to be, ultimately, the pattern: on pain of regress, there has to be such a thing as justified thought involving a reference-demanding singular mode of presentation where no independent such mode of presentation is deployed in the justification. The thoughts concerning abstracts that abstraction principles introduce us to should be received as *basic* singular thoughts in this sense. The requirement that some independent assurance be given that terms introduced by abstraction principles refer misses this epistemologically fundamental point.

---

[25] These ideas are elaborated in Hale and Wright 2009

§15 A great deal more needs to be said, of course.[26] My intent here has merely been to outline some steps towards a more explicit account of the kind of epistemological stage setting which I believe abstractionism needs if its philosophical significance, and limitations, are to be properly understood.

One cautionary corollary of the foregoing discussion is perhaps worth a closing emphasis. We need to realise that the traditional conception of the basic a priori as a realm of apodictic certainty—a conception in which Frege was immersed up to his ears—is a great mistake. A priori knowledge, no less than empirical knowledge, is subject to the ultimate groundlessness that Wittgenstein stressed in his last philosophical writings. Here I have tried to outline how the abstractionist project looks when that point is taken on board — and how indeed the point can help it address certain of the criticisms to which it has been subjected.[27]

References:

Boghossian, Paul
— 2000 "Knowledge of Logic" in Paul Boghossian and Christopher Peacocke, eds., *New Essays on the A Priori* Oxford, the Clarendon Press), 229-254

Boolos, George
—1990 "The standard of equality of numbers" in G . Boolos, ed., *Meaning and Method: Essays in Honor of Hilary Putnam*, Cambridge: Cambridge University Press, 261-77; reprinted in his *Logic, Logic and Logic* Cambridge, Mass: Harvard University Press 1998, 202-219
—1997 'Is Hume's Principle Analytic?' in Richard G. Heck, Jr., ed. *Logic, Language, and Thought* Oxford: Oxford University Press 1997; reprinted in his *Logic, Logic and Logic* Cambridge, Mass: Harvard University Press 1998, 301-14

Burgess, John
— 2005 *Fixing Frege*, Princeton University Press

Demopoulos, William
— 1998 "The Philosophical Basis of Our Knowledge of Number" *Nous* 32, 481-503
— 2000 "On the Origin and Status of the Concept of Number", *Notre Dame Journal of Formal Logic* 41, 210-26

---

[26] Some of what needs to be said is broached Hale and Wright (2009).

[27] This paper descends remotely from a talk given at an Arché Abstraction workshop in 2002. Subsequent versions were presented at a philosophy of mathematics colloquium in Bristol in 2009 and in my graduate seminar on abstractionism at NYU in Spring 2010. I cannot remember all who have offered me helpful comments and criticisms, but they certainly include Roy Cook, Philip Ebert, James Ladyman, Øystein Linnebo, Agustin Rayo, and Ted Sider. Thanks too, as always, to Bob Hale.

— forthcoming "Generality and Objectivity in Frege's *Foundations Of Arithmetic*" in *Logic, Language and Mathematics: Essays for Crispin Wright*, Alex Miller (ed.): Oxford: Oxford University Press

Dogramaci, Sinan
— 2010 "Knowledge of Validity", *Noûs* 44 (3), 403-432

Dummett, Michael
—1998 "Neo-Fregeans: in Bad Company", in Matthias Schirn, ed., *The Philosophy of Mathematics Today*, Oxford: Clarendon Press, 386–7

Ebert, Philip
—2005 *The Context Principle and Implicit Definitions*, University of St. Andrews PhD thesis.

Enoch, David and Schechter, Joshua
— 2008 "How Are Basic Belief-Forming Methods Justified?" *Philosophy and Phenomenological Research*, 76(3), 547–579

Field, Hartry
— 1984 "Platonism for cheap? Crispin Wright on Frege's context principle" *Canadian Journal of Philosophy* 14: 637-62; reprinted in Field's *Realism, Mathematics & Modality*, Oxford: Blackwell, 1989, 147-70
Fine, Kit
— 1998 "The Limits of Abstraction" in Matthias Schirn, ed., *The Philosophy of Mathematics Today*, Oxford: the Clarendon Press, 503-629.

Hale, Bob and Wright, Crispin
—2000 "Implicit Definition and the A Priori", in Paul Boghossian and Christopher Peacocke, eds., *New Essays on the A Priori*, Oxford: the Clarendon Press 286-319
—2009 "The Metaontology of Abstraction" in David Chalmers et al, eds., *Metametaphysics*, Oxford: Oxford University Press, 178-212

Heck, Richard
— 1997 "Finitude and Hume's principle" Journal of Philosophical Logic 26, 589-617; reprinted with a new postscript in his (2011), 237-266
— 2011 *Frege's Theorem*, Oxford: Oxford University Press

Linnebo, Øystein
—forthcoming "Impredicativity in the Neo-Fregean Programme" in Philip Ebert and Marcus Rossberg (eds.), *Abstractionism in Mathematics: Status Belli*, Oxford: Oxford University Press
—2009 ed. *Synthese* 170 (3), special number on Bad Company

McGee, Vann
—1985 "A counterexample to modus ponens" *Journal of Philosophy* 82 (9), 462-471

Williamson, Timothy
—2003 ‘Understanding and inference’ (in symposium on ‘Blind Reasoning’), *The Aristotelian Society*, sup. vol. 77, 249-293.
—2006 “Conceptual truth”’. *The Aristotelian Society*, sup. vol. 80, 1-41
—2007 *The Philosophy of Philosophy* Oxford: Blackwell

Wright, Crispin
—2004 "Warrant for Nothing: On Epistemic Entitlement", *Aristotelian Society Supplementary Vol*. LXXVIII, 167-212
—2004(b) "Intuition, Entitlement and the Epistemology of Logical Laws", in *Dialectica* 58, 155-75
—2012 "Frege and Benacerraf's problem" in Robert DiSalle, Mélanie Frappier, and Derek Brown, eds., *Analysis and Interpretation in the Exact Sciences: Essays in Honour of William Demopoulos*, Springer, Dordrecht Heidelberg London New York, pp. 117-33
—*Forthcoming*; "Welfare State Epistemology (Entitlement II)", in Dylan Dodd and Elia Zardini, eds., *Contemporary Perspectives on Scepticism and Perceptual Justification*, Oxford: Oxford University Press