Synthese (2009) 170:349–369
DOI 10.1007/s11229-007-9264-8

Hume's Big Brother: counting concepts and the bad company objection

Roy T. Cook

Accepted: 31 August 2007 / Published online: 22 November 2007

Abstract A number of formal constraints on acceptable abstraction principles have been proposed, including conservativeness and irenicity. Hume's Principle, of course, satisfies these constraints. Here, variants of Hume's Principle that allow us to count concepts instead of objects are examined. It is argued that, prima facie, these principles ought to be no more problematic than HP itself. But, as is shown here, these principles only enjoy the formal properties that have been suggested as indicative of acceptability if certain constraints on the size of the continuum hold. As a result, whether or not these higher-order versions of Hume's Principle are acceptable seems to be independent of standard (ZFC) set theory. This places the abstractionist in an uncomfortable dilemma: Either there is some inherent difference between counting objects and counting concepts, or new criteria for acceptability will need to be found. It is argued that neither horn looks promising.

Keywords Frege · Neo-logicism · Abstraction · Arithmetic · Higher-order logic · Bad company objection

1 Logicism, abstractionism, and bad company

Some logical preliminaries: An abstraction principle is any formula of the form:

$$(\forall\alpha)(\forall\beta)(@(\alpha) = @(\beta) \leftrightarrow \mathrm{E}(\alpha, \beta))$$

where "@" denotes a unary function mapping entities of the type ranged over by α (usually concepts, objects, or sequences of such) to objects, and "E(,)" is an equiva-

R. T. Cook (✉)
Villanova University, 831 Heller Hall, 271, 19th Ave. S, Minneapolis, MN 55455-0310, USA
e-mail: cookx432@umn.edu

lence relation on those same entities. Abstraction principles allow us to introduce new terms (and thus presumably gain privileged epistemological access to the referents of those terms) by defining the identity conditions for those objects using linguistic resources that are already understood (i.e. those resources occurring in the equivalence relation "E(,)"—in the cases of interest here "E(,)"will be a purely logical formula, although in general it need not be).[1] Thus, an abstraction principle is meant to act as an implicit definition of sorts, providing (so the story goes) an account of the meaning of novel terms of the form "@(a)" (see Hale and Wright (2000) for details).

Perhaps the first notable occurrence of an abstraction principle occurs in Frege's logicist reconstruction of arithmetic. Frege (all but) notes, in the *Grundlagen* [1974], that the standard (higher-order) Peano axioms for arithmetic follow from the abstraction principle now known as *Hume's Principle* (the explicit derivation of the Peano axioms from Hume's Principle was "extrapolated" from Frege's comments by Wright (1983), Boolos (1990a), Heck (1993), and Boolos and Heck (1998), among others).

Hume's Principle is a formalization of the thought that, given two arbitrary (first-level) concepts X and Y, the number of X's is identical to the number of Y's if and only if the X's and the Y's can be put in a one-to-one correspondence. More formally, we have:

$$\text{HP: } (\forall X)(\forall Y)(\#(X) = \#(Y) \leftrightarrow (X \approx Y))$$

where $X \approx Y$ abbreviates the second-order claim that X and Y are equinumerous, i.e. that there is a one–one onto function from the X's to the Y's.

We can formulate rather natural definitions of arithmetical notions such 'natural number', 'successor' and 'addition' in terms of the numerical operator "#". The fact that, given these definitions, the second-order Peano axioms for arithmetic follow from *Hume's Principle* is quite notable as a purely mathematical result, and the result has come to be called *Frege's Theorem* (for a detailed examination of this result, and various streamlined versions of it, see Heck (1997a)). Interest in Hume's Principle as an implicit definition of number has been rekindled by the publication of Wright's *Frege's Conception of Numbers as Objects* [1983].

One major thread of criticism within the literature on abstractionism has come to be called the Bad Company Objection. As is often the case, the 'Bad Company Objection' is not, actually, a single objection to abstractionism, but is rather a cluster of worries, all of which take something like the following form:[2]

Bad Company: Simple Version:

There are abstraction principles similar to HP but which have unattractive formal or philosophical properties.

[1] For example, Stewart Shapiro's (2000) reconstruction of the real numbers proceeds through a series of abstraction principles, each one introducing new objects in terms of an equivalence relation on the (non-logical) objects introduced by the previous principle.

[2] In the present paper I make no attempt to decide exactly which such concerns do and do not deserve the label "Bad Company", and shall instead consider a series of such worries in this general spirit.

The worry here is that, if there are abstraction principles that have unattractive properties that rule them out as legitimate definitions of mathematical concepts, then the general method of abstraction cannot be defended. But if abstraction principles are not acceptable across the board, then what reason do we have for thinking that Hume's Principle is okay? The question regarding which principles are acceptable is a critical one if the abstractionist project is to be extended from its current status as a promising foundation for arithmetic to an adequate treatment of larger tracts of mathematics such as set theory and analysis.

Russell's Paradox provides us with the first, and simplest, version of the Bad Company Objection. As is well known, not all abstraction principles are consistent—Basic Law V:

$$\text{BLV: } (\forall X)(\forall Y)(\S(X) = \S(Y) \leftrightarrow (\forall z)(X(z) \leftrightarrow Y(z)))$$

for example, allows us to derive a contradiction. So:

Bad Company 1:

There are abstraction principles that are inconsistent.

The response to this worry is rather obvious—we need merely restrict our attention to consistent abstraction principles. In other words, consistency is a necessary condition for the acceptability of an abstraction principle.

Even if we set aside worries over the consistency of Hume's Principle (see, e.g. Boolos (1997)), and accept that it is consistent (since it is equiconsistent with second-order Peano Arithmetic), it is easy to see that restricting the abstractionist account to consistent abstraction principles is not enough. The first to notice this was (as usual) Boolos, who pointed out that there are abstraction principles that are consistent but incompatible with each other.[3] Thus:

Bad Company 2:

There are consistent abstraction principles that are incompatible with Hume's Principle.

Thus, if Hume's Principle is acceptable, then mere consistency, while necessary, is not sufficient for an abstraction principle to be acceptable.

The most well-known example of such a principle is due, actually, to a defender of the abstractionist project. Wright (1997) pointed out that the following Nuisance Principle:[4]

$$\text{NP: } (\forall X)(\forall Y)[\text{NUI}(X) = \text{NUI}(Y) \leftrightarrow \text{FSD}(X, Y)]$$

can be satisfied on domains of any finite cardinality, but on no domains of infinite cardinality. As a result, NP is satisfiable (and thus consistent), but the theory obtained

[3] More generally, he noted that there are many pairs of abstraction principles where each formula in the pair is satisfiable, yet the theory obtained by conjoining them is unsatisfiable. Thus, the problem is independent of the acceptability of Hume's Principle itself, although it is most easily framed in these terms.

[4] Here FSD(X, Y) abbreviates the second-order formula asserting that the symmetric difference of X and Y, that is, the collection of objects that are either X-and-not-Y or are Y-and-not-X, is finite.

by conjoining NP and HP is unsatisfiable (since HP implies the existence of the natural numbers, and thus is only satisfiable on infinite domains).

Abstractionists have a response at this point, however. The problem with principles like the Nuisance Principle above, so the response goes, is that these principles are non-conservative: They entail statements about non-abstracts that are not entailed without the abstraction principle in question.

The formulation of conservativeness[5] requires the notion of relativizing a formula P to an open formula A(x). Let A(x) be a formula with x the only free variable. Then we can define the relativization of a formula relative to A(x) as follows:

(1) $P^A = P$ [where P atomic]
(2) $(\neg\Phi)^A = \neg(\Phi^A)$
$(\Phi \wedge \Psi)^A = (\Phi^A) \wedge (\Psi^A)$
$(\Phi \vee \Psi)^A = (\Phi^A) \vee (\Psi^A)$
$(\Phi \rightarrow \Psi)^A = (\Phi^A) \rightarrow (\Psi^A)$
$(\Phi \leftrightarrow \Psi)^A = (\Phi^A) \leftrightarrow (\Psi^A)$

(3) $(\forall y)(\Phi)^A = (\forall y)(A(y) \rightarrow \Phi^A)$
$(\exists Y)(\Phi)^A = (\exists y)(A(y) \wedge \Phi^A)$
$(\forall Y)(\Phi)^A = (\forall Y)((\forall x)(Y(x) \rightarrow A(x)) \rightarrow \Phi^A)$
$(\exists Y)(\Phi)^A = (\exists Y)((\forall x)(Y(x) \rightarrow A(x)) \wedge \Phi^A)$

In what follows we are interested in whether or not a particular abstraction principle AP is conservative over T, where **T** is a theory in a language L not containing the abstraction operator @ introduced by AP, and L+ is the language obtained by expanding L through the addition of @.

An abstraction principle AP is conservative over a theory T if and only if:[6]

If $T^{\neg(\exists Y)(x=@(Y))}$, AP $\Rightarrow C^{\neg(\exists Y)(x=@(Y))}$, then T $\Rightarrow$ C.

In other words, if the principles of theory T, restricted to the non-abstracts, plus AP prove some claim C that is restricted to the non-abstracts, then T alone (and unrestricted) proves C (unrestricted).

Given a particular background logic (e.g. standard second-order logic), there is still an issue to be settled: Do we mean for "⇒" to represent deductive or semantic (model-theoretic) consequence in the above definition? In what follows we shall (like Weir (2003)) restrict our attention, for the most part, to semantic consequence. There are two reasons for this—the first related to the Gödelian incompleteness phenomenon, and the second more practical.

First off, there is some reason for thinking that, in evaluating various formal claims made by the abstractionist, the semantic consequence relation is the appropriate target notion. If the abstractionist meant for the entailment relation in question to be the

[5] Here we utilize a formulation of conservativeness that Weir (2003) calls "Field Conservative". Weir also considers an alternative notion in his paper: "Caesar-neutral conservative". All results in this paper hold for this second notion as well.

[6] Here, and in what follows, "⇒" will represent whatever consequence relation is appropriate. As shall be explained below, in most cases this will be the relation given by standard semantics for second- and higher-order logics. For details on such systems, see Shapiro (1991).

deductive one, then Gödel's incompleteness results (or, more carefully, the application of these results to show that second-order logic is incomplete) would entail that there are, for example, truths of arithmetic that do not, in fact, follow from Hume's Principle. But Hume's Principle is meant to be a foundation for arithmetic. So the semantic relation must be the one at issue.

Second, and more practical, is the difficulty of demonstrating deductive conservativeness results. For example, it is known that Hume's Principle is conservative if the consequence relation is taken to be the semantic one (see, e.g. Weir (2003) for details). It is, however (as far as the present author can determine) an open question whether Hume's Principle is conservative, in either sense, on the deductive reading of these definitions. Thus, while conservativeness results relative to the deductive consequence relation would be interesting and important (at least mathematically), for the remainder of this paper we shall restrict our attention (for the most part) to results that are a bit more attainable (i.e. semantics conservativeness/non-conservativeness results).

In restricting our attention to the semantic (i.e. model-theoretic) notions, we shall have to assume ZFC explicitly in the metatheory, in order to provide a substantial theory of sets upon which to base our set-theoretic semantics. As a result, the approach taken here is an 'external' one, insofar as we are examining the prospect for abstractionism from the perspective of a traditional set theorist. Ideally, the abstractionist position should eventually be developed and defended 'internally', that is, from the perspective of an abstractionistically acceptable version of set theory. Since there is little agreement regarding what such an abstractionist set theory might look like, however, and some doubt regarding the very possibility of such a theory (see Uzquiano's contribution to this volume), we retain the external perspective here.

Now, as already noted, Hume's Principle is not only satisfiable, it is conservative. In addition, the Nuisance Principle fails to be conservative since it entails that the non-abstracts are finite in number. So we can, in good conscience, now claim that both satisfiability and conservativeness are necessary conditions for acceptable abstraction principles.

Unfortunately, Weir (2003) has demonstrated that conservativeness, although perhaps a necessary condition for the acceptability of an abstraction principle, cannot be sufficient. First, let us, following Weir, define an unbounded abstraction principle as follows:

AP is *unbounded* $\leftrightarrow$ For any cardinal κ, there is a cardinal $\lambda > \kappa$ such that AP is satisfiable on domains of size λ.

In other words, an abstraction principle is unbounded if there is no upper limit to the size of models of that principle. Weir then proves the following:

Theorem (Weir (2003)): *All unbounded abstraction principles are conservative.*

This result, while simplifying matters considerably, also allows us to formulate a third version of the Bad Company Objection (one which Weir calls 'Embarassment of Riches'):

Bad Company 3:

There are pairs of abstraction principles where each principle is, individually, both satisfiable and conservative, yet the pair is jointly unsatisfiable.

Examples are not hard to come by. For example, we can utilize a trick due to Heck, who noted that, for any formula Φ containing no abstraction operators, the abstraction principle:

$$AP_{\Phi}: (\forall X)(\forall Y)(@(X) = @(Y) \leftrightarrow (\Phi \vee (\forall z)(X(z) \leftrightarrow Y(z))))$$

is satisfiable on a (non-empty) domain of size κ if and only if Φ is. So, the following pair of principles will do the trick:

$$(\forall X)(\forall Y)(@_1(X) = @_1(Y) \leftrightarrow (\text{Succ} \vee (\forall z)(X(z) \leftrightarrow Y(z))))$$

$$(\forall X)(\forall Y)(@_2(X) = @_2(Y) \leftrightarrow (\text{Limit} \vee (\forall z)(X(z) \leftrightarrow Y(z))))$$

where "Succ" abbreviates the second-order formula asserting that the universe is the size of a successor cardinal, and "Limit" abbreviates the second-order formula asserting that the universe is the size of a limit cardinal. By Weir's theorem, both of these principles are conservative, yet they are quite patently incompatible.[7]

Weir (2003) pushes things further, suggesting a further condition that abstraction principles might be required to meet, namely that they be irenic:[8]

AP is *irenic* $\leftrightarrow$ AP is compatible with any conservative abstraction principle.

As he did earlier with conservativeness, Weir, in the same paper [2003], provides us with a simple model-theoretic test for irenicity. First, a definition:

AP is *stable* $\leftrightarrow$ there is a cardinal κ, such that that AP is satisfiable on domains of size $\gamma \geq \kappa$.

Weir proves the following:

Theorem (Weir (2003)): *An abstraction principle is irenic if and only if it is stable.*

Thus, the irenic abstraction principles are exactly those that are satisfiable on all domains above some given cardinality.

At this point the reader would be forgiven for thinking that another version of the Bad Company Objection is forthcoming, i.e. that there will again be principles that are individually irenic, but somehow incompatible. But such a result is in a certain sense impossible, as the following result shows:

Theorem 1 *Any set of irenic principles is satisfiable.*[9]

7 It is worth pointing out that Weir's own example of such a pair of conservative, yet incompatible principles is a bit more complicated than this, due to his focusing his attention on abstraction principles of a certain form—what he calls distraction principles.

8 Here we will ignore Wright's further thought that abstraction principles should not be 'paradox-exploitative', since this approach seems technically unfeasible. See Wright (1999) for his informal explication of the idea, and Weir (2003) for a convincing argument that there is no formal way to flesh out the constraint.

9 I leave the interesting issue regarding whether, given a rich enough language, there might be proper classes of irenic principles with no models to ambitious readers. The version proven here, in terms of sets of principles, is surely strong enough for the abstractionist's purposes.

Proof Let X be a set of irenic abstraction principles, and $S = \{\kappa$: there is an AP $\in$ X such that κ is the least cardinal such that AP is satisfiable on all cardinals $\geq \kappa\}$. Then X is satisfiable on any domain whose cardinality is at least the supremum of S. □

Thus (assuming that ZFC provides a good guide to what principles are acceptable, but see below!), there does not seem to be any reason to doubt the acceptability of the collection of irenic abstraction principles, since objections of the sort we have been examining here will not be forthcoming.

So: Conservativeness is necessary for acceptability of abstraction principles, and irenicity is sufficient. Moreover, the examples above show that conservativeness is not necessary for acceptability (assuming, of course, that the collection of acceptable abstraction principles must be jointly satisfiable). The question remains, however: Is the collection of irenic abstraction principles exactly the collection of acceptable abstraction principles (i.e., is irenicity necessary as well as sufficient), or is there some broader class of principles all of whose members are acceptable?

There is at least some prima facie reason for thinking that irenicity is both necessary and sufficient for acceptability. Assume that it were not: Then there must be some conservative but non-irenic abstraction principle AP_1 which is acceptable. But, since AP_1 is non-irenic, there will be another conservative, but non-irenic principle AP_2, such that AP_1 and AP_2 are incompatible. So (again, assuming that the collection of acceptable abstraction principles must be consistent), AP_2 cannot be acceptable. But why is it AP_1, and not AP_2, that is acceptable? After all, AP_2 (just like AP_1) is consistent with all of the irenic principles. The acceptability of AP_1, and not AP_2, looks a bit ad hoc.

Of course, the comments of the previous paragraph are merely suggestive. After all, there might be some as-of-yet undiscovered formal property, weaker than irenicity yet stronger than conservativeness, which holds of exactly the acceptable principles, and which explains the acceptability of AP_1 but not AP_2.

Nevertheless, the recent literature on the Bad Company objection appears to present irenicity as both necessary and sufficient for acceptability (see, e.g., Weir (2003), although he is not completely explicit with regard to this). As a result, we shall, in the remaining sections, carefully spell out the status of the various principles under scrutiny both from the perspective of irenicity as merely a sufficient condition for acceptability, and from the perspective of irenicity as both necessary and sufficient.

2 Counting concepts

Hume's Principle, as we have already seen, provides a mapping from concepts to their associated number. In other words, HP provides us with the resources for counting objects, since to count a collection of objects we need (at least, in some technical sense) merely to determine the cardinal number that HP associates with the concept that holds of exactly the objects in that collection. But of course, objects are not the only 'thing's that we can count.

Consider the following statement:

The number of cats is 8.

The obvious way to formalize this formula, utilizing the number operator provided by Hume's Principle, is:

$\#(\text{x is a cat}) = 8.$

where "8" is some canonical name for the number 8. Our problem arises when we try to generalize this account to other instances of straightforward number talk. In particular, consider the following concept-counting claim:

The number of concepts that hold of no more than a, b, and c is 8.

As before, the obvious way to formalize this along Fregean or abstractionist lines would be something like:

$$\#((\forall y)(X(y) \rightarrow (y = a \vee y = b \vee y = c))) = 8$$

where "8" is (again) some canonical name of the number eight. The problem, in the present context, is that Hume's Principle only assigns numbers to first-level concepts (i.e. concepts that hold of objects). The predicate:

$$(\forall y)(X(y) \rightarrow (y = a \vee y = b \vee y = c))$$

however, has only the second-order variable "X" free. Thus (assuming the full third-order comprehension schema) this predicate 'designates' a second-level concept (i.e., one that holds of first-level concepts).

Before further examining why the abstractionist needs to worry about such higher-level number talk, it is worth noting that Frege himself was not plagued by such problems. The reason for this is that Frege had Basic Law V, and as a result, any first- (or second-, or higher-) level concept had a unique extension. As a result, he could use extensions as objects that would serve as proxies for the concepts in question, and thus he could rest content with only defining the number operator for first-level concepts.

For example (ignoring, for the moment, the inconsistency of Basic Law V), Frege could have formalized the sentence above as something like:

$$\#((\exists X)(z = \S(X) \wedge (\forall y)(X(y) \rightarrow (y = a \vee y = b \vee y = c)))) = 8$$

Thanks to Russell's Paradox, however, Basic Law V is unavailable to the abstractionist, as is the idea that each concept (of any order) can have a unique object as its extension. As a result, if the abstractionist wishes to be allowed to assign any predicate (of any order) a number (i.e. if the abstractionist wishes to be able to count, not just objects, but concepts), then some other resource for doing so must be found.[10]

Of course, one might initially suspect that the abstractionist might not need to count concepts. What reasons do we have for thinking that the numbers obtained by Hume's

[10] Readers familiar with the literature might at this point think that there is a simple solution to the problem, namely, to symbolize this sentence using the resources provided by Hume's Principle plus some restricted (and thus consistent) version of Basic Law V, such as Boolos (1989) NewV, which provides a unique extension for every concept that holds of fewer objects than there are in the universe. While this will handle the present example, it will not deal with cases where we are counting 'Big' (or, more generally, 'Bad') concepts, such as:

The number of concepts whose complements hold of no more than a, b, and c is 8.

Similar examples can be constructed for other restricted variants of Basic Law V.

Principle alone are not enough to do all the arithmetic we need? After all, Hume's Principle alone provides us with all of the natural numbers.

There are a number of (interconnected) reasons why the abstractionist needs to be able to count not just objects but concepts. The first and most practical reason is just this: We do seem to be able to count concepts in everyday language—there does not seem to be anything wrong with the sentence:

The number of concepts that hold of no more than a, b, and c is 8.

If the abstractionist account is meant to provide an account of our arithmetical knowledge—that is, all of our arithmetical knowledge—then prima facie it should provide an account of the meaning of, and our knowledge of, such apparent instances of counting concepts.

Put another way: Abstractionism depends on the legitimacy of second- (and perhaps higher-) order logic. Whatever the details of the eventual defense of higher-order logic look like, it will, in some sense or other, amount to some version of the claim that quantification over such entities is no more problematic than quantification over first-order things (i.e. objects). But if quantification over concepts (of whatever order) is no more problematic than quantification over objects, why should counting concepts be any more problematic than counting objects (since quantification, after all, is merely a very simple way of counting—"all", "none", "at least one", etc.)?

The second, more principled reason for thinking that a complete abstractionist account of mathematics should include resources for counting, not just objects, but concepts, is a version of the Bad Company Objection. If it turns out that the abstraction principles that we need in order to count concepts (i.e. those principles that provide cardinal numbers for second- and higher-level concepts) are unacceptable (e.g. they turn out to be non-irenic), then this would seem to throw some doubt onto the innocent status of the cardinal numbers provided by Hume's Principle, since presumably there is nothing going on that is *conceptually* new when we move from counting objects to counting concepts. In other words, there does not seem to be anything additional of mathematical significance that one needs to learn in order to count concepts, if one is already competent at counting objects (This issue shall be revisited at the end of the paper, however).[11]

At any rate, it is tempting to think that if Hume's Principle is an acceptable principle, then similar principles allowing us to assign cardinal numbers to concepts ought to be acceptable as well. If they were not, then the abstractionist would be faced with the rather difficult task of explaining why these higher-order principles were not relevantly similar to Hume's Principle. Pointing out some technical difference in their model theory is surely not enough—what must be done, if such higher-order variants of Hume are to be ruled out, is to provide some philosophical explanation regarding why counting concepts is, in principle, different from counting concepts. Although

[11] There is a final possibility here. One might argue that we can count objects, but not concepts, because counting requires having identity conditions for the 'things' being counted. Since we are treating concepts here along the standard second-order logical lines, and thus individuating them extensionally, however, we do, at least from a technical perspective, have identity conditions of a sort for them.

such an argument might well exist, its general shape is not obvious. Thus, our default assumption should be that an account of such 'higher-order' numbers is desirable.

If we therefore adopt a default assumption that counting concepts (i.e. assigning numbers to second- and higher-level concepts) should be a part of the abstractionist account, the next step is to formulate principles that will allow us to obtain such numbers along abstractionist lines. As we shall see in the next section, formulating such principles is not difficult, although, as the section after that demonstrates, these higher-order analogues of Hume's Principle are a good bit less well-behaved than Hume itself.

3 Upper Hume and Hume's family

In order to formulate higher-order analogues of Hume's Principle, that is, principles that assign cardinal numbers to second- and higher- level concepts, we will need to introduce a bit of terminology. First, we will need to have variables for each finite order. For simplicity's sake we will restrict our attention to monadic variables, since these will be of primary interest (analogous numbering conventions will be understood to hold of n-ary predicates, for $n > 1$). Superscripted natural numbers will be used to designate the 'level' of a variable, so "X^1" is a first-level variable (i.e. one that takes terms as argument), X^2 will be a second-level variable (i.e. one that takes first-level variables as argument), and, in general, "X^i" is an i^{th}-level variable (i.e. one that takes $i\text{-}1^{th}$-level variables as argument).[12] Along these lines, we can reformulate Hume's principle as:

$$\text{HP: } (\forall X^1)(\forall Y^1)(\#(X^1) = \#(Y^1) \leftrightarrow (X^1 \approx_1 Y^1))$$

where $X^1 \approx_1 Y^1$ abbreviates the (second-order) claim that X^1 and Y^1 are equinumerous, i.e. that there is a one-one onto function from the X^1's to the Y^1's.

With higher-order resources in place, we can also formulate a principle that will allow us to 'count' first-level objects, that is, a variant of Hume's Principle that assigns cardinal numbers to second-level concepts. We will call this principle Upper Hume:

$$\text{UH: } (\forall X^2)(\forall Y^2)(\#(X^2) = \#(Y^2) \leftrightarrow (X^2 \approx_2 Y^2))$$

where $X^2 \approx_2 Y^2$ abbreviates the (third-order) claim that X^2 and Y^2 are equinumerous, i.e. that there is a one-one onto function from the X^2's to the Y^2's.[13]

We can then, of course, formulate a principle that provides cardinal numbers for third-level concepts:

$$\text{HP}_3\text{: } (\forall X^3)(\forall Y^3)(\#(X^3) = \#(Y^3) \leftrightarrow (X^3 \approx_3 Y^3))$$

where $X^3 \approx_3 Y^3$ abbreviates the (fourth-order) claim that X^3 and Y^3 are equinumerous, i.e. that there is a one-one onto function from the X^3's to the Y^3's.

[12] Formation rules, etc., for this ω-order logic are as in Shapiro (1991).

[13] One can obtain "$(X^{i+1} \approx_{i+1} Y^{i+1})$" recursively from "$(X^i \approx_i Y^i)$" by just replacing each occurrence of an n^{th} level variable with a corresponding $n{+}1^{th}$ level one (for all n). (Strictly speaking one must also substitute co-extensionality claims of the appropriate level for first-order identity claims when carrying out the substitution, unless we assume that "=" is defined on terms from all levels).

More generally, we can obtain an ω-sequence of such principles according to the following schema:

$$HP_i: (\forall X^i)(\forall Y^i)(\#(X^i) = \#(Y^i) \leftrightarrow (X^i \approx_i Y^i))$$

where $X^i \approx_i Y^i$ abbreviates the (i+1th-order) claim that X^i and Y^i are equinumerous, i.e. that there is a one-one onto function from the X^i's to the Y^i's. Note that $HP = HP_1$ and $UP = HP_2$. We shall call the theory obtained by combining all instances of HP_i Hume's Family:

$$HF = \{HP_i : i \in \omega\}$$

We could, of course, have moved further up the ordinal hierarchy, formulating even higher-order versions of Hume's Principle (e.g. HP_ω, which assigns cardinal numbers to any ω-level concept—that is, one that holds only of objects or concepts of level less than ω). While doing so might be interesting, both technically and philosophically, the present, countably infinite collection of principles is already complex enough to keep us busy for the remainder of this paper. At any rate, the arguments sketched in the previous section, regarding the need for principles that number concepts of higher and higher order, do not obviously generalize into the tranfinite (since it is not clear that the abstractionist need countenance quantification over concepts of non-finite order).[14]

So, formulating abstraction principles that assign cardinal numbers to concepts of second- or higher-level is straightforward. As will be demonstrated in the next section, however, showing that these principles have the desired properties is another matter.

4 Counting concepts and the continuum hypothesis

In this section we shall examine whether or not Upper Hume, and the theory Hume's Family, have the meta-theoretical properties that Wright, Weir, and others have suggested are necessary for acceptable abstraction principles. In particular, we shall ask whether or not these principles are satisfiable, conservative, and irenic. As we shall see, however, the answers to these questions do not come in a simple "yes" or "no" form. First, we introduce the following definition:

$$C_{(1,\kappa)} = \kappa$$
$$C_{(\gamma+1,\kappa)} = 2^{C(\gamma,\kappa)}$$

The following result sums up the important model-theoretical behavior of our generalizations of Hume's Principle:

Theorem 2 *HP_i has a model of size κ if and only if κ is infinite and, where α is the ordinal such that:*

[14] Here is a quick inductive argument for the claim that the abstractionist ought to countenance quantification over nth-level concepts, for any level n: Base case: Without quantification over first-level concepts, Hume's Principle is not formulable, so abstractionism is a non-starter. Inductive step: Assume that quantification over nth-level concepts is acceptable. Then the abstractionist ought to be able to assign cardinal numbers (and other acceptable types of abstract) to nth-level concepts. But the abstraction operator involved in such principles is a particular instance of an n+1th level function. So the abstractionist countenances at least some n+1th level concepts (understanding n-ary functions here as n+1-ary concepts). So the abstractionist ought to be able to quantify over all such n+1th level concepts. Q.E.D.

$C_{(i,\kappa)} = \aleph_\alpha$

we have:

$|\alpha| \leq \kappa$.

Proof Assume that HP_i has a model of size κ. First, κ must be infinite, since any HP_i proves an analogue of Frege's Theorem. So, there are κ-many objects, 2^κ many first-level concepts, ... and $C_{(i,\kappa)}$ many i^{th} -level concepts. Thus, if $C_{(i,\kappa)} = \aleph_{\alpha}$, then there must be $|\omega + \alpha|$ many cardinal numbers (i.e. ω-many finite cardinals, plus the α-many alephs). So κ must contain at least $|\omega + \alpha|$ many objects to serve as these numbers. So $|\alpha| \leq \kappa$. The converse is similar. □

To see more intuitively what is going on here, let us consider Upper Hume. By Theorem 1, Upper Hume will be satisfiable on a cardinal κ if, and only if, where $\aleph_\alpha = 2^\kappa$, we have that $|\alpha| \leq \kappa$. Since κ is infinite, if $|\alpha| > \kappa$ then $|\alpha|$ is the number of cardinals occurring between κ and 2^κ. Thus, put loosely, if Upper Hume is to be satisfiable at a cardinal κ, then there cannot be more than κ many cardinals between κ and 2^κ.

To put this even more loosely: Upper Hume is satisfiable on a domain of cardinality κ if, and only if, the Generalized Continuum Hypothesis does not fail too badly at that cardinal. We do not, however, know if the Generalized Continuum Hypothesis holds at all, none, or some of the cardinals, however, and, moreover, we have no real idea of how badly the GCH can fail at those cardinals at which it might fail (more on this below).[15] Hence, the problems with counting concepts.

Before we look more closely at some of these problems, however, we should note that not everything goes bad. In particular, we can prove the following:

Theorem 3 *ZFC + GCH proves that Hume's Family has a model of size κ, for any infinite κ.*

Proof Assume that $\kappa = \aleph_\beta$ is infinite. Given GCH, $C_{(i,\kappa)} = \aleph_{\beta+i-1}$. So, for each HP_i in Hume's Family, HP_i will be satisfiable at κ if, and only if, $|\beta + i - 1| \leq \kappa$. So, Hume's Family will be satisfiable at κ if, and only if, $|\beta + \omega| \leq \kappa$. The last follows from the fact that $\beta \leq \kappa$ and κ is infinite. □

Assuming that our second-order deductive system is sound, we obtain:

Corollary 4 *Con(ZFC + GCH) entails Con(Hume's Family)*[16]

Since Godel's inner model method provides us with:

Con(ZFC) entails Con(ZFC + GCH)

We obtain, by hypothetical syllogism:

[15] It is worth noting that part of the folklore regarding set theory has it that most set theorists believe that the continuum hypothesis is false. I have no idea what sort of data might justify such claims, and I doubt that, even if true, such a sociological observation will be of much help to the abstractionist. At any rate, it seems unlikely that the majority of set theorists believe the continuum hypothesis, and its generalization, to fail as badly as is required to prevent the irenicity of the various HP_i's. Again, the question is how much, if any, help such an observation is to the abstractionist.

[16] Here, and below, I use the prefix "Con" to designate proof-theoretic consistency.

Corollary 5 *Con(ZFC) entails Con(Hume's Family)*

Thus, there is no reason to worry over the proof-theoretic inconsistency of any of the higher-order analogues of Hume's Principle (or, at least, no reason beyond those we might have regarding ZFC itself).

Theorem 3 also provides us with a result relevant to the irenicity of Upper Hume and the members of Hume's Family:

Corollary 6 *ZFC + GCH entails HP_i is irenic, for $i \geq 1$.*

In other words, if the Generalized Continuum Hypothesis holds, then each HP_i can be satisfied on any infinite cardinal κ, so each HP_i is stable, and thus satisfiable, conservative, and irenic.

So far, all of this looks good—we have proven that the abstraction principles needed to assign cardinal numbers to second- and higher-level concepts are proof-theoretically consistent, and that they are irenic if the GCH holds. Treating the second-order theory as a many-sorted first-order theory, the former result also guarantees that our class of number-defining abstraction principles will have a Henkin (i.e. possibly non-standard) model. But we do not, as yet, know whether there exists any standard second-order model of Hume's Family, or even of any single instance of HP_i where $i > 1$.

Thus, the question to ask is: Can we prove (without reliance on questionable principles such as the Generalized Continuum Hypothesis) that any instance of HP_i (other than Hume's Principle itself) has a standard model? The answer, unfortunately, is "no" To show this, we will need a result due to Easton (1970):

Easton Forcing: An Easton function is a function f from cardinals to cardinals such that:

For all regular cardinals κ, γ, where $\kappa < \gamma$, $f(\kappa) \leq f(\gamma)$.
For all regular cardinals κ, $\text{cf}(f(\kappa)) > \kappa$.

If f is an Easton function, then there is a model of ZFC where $2^\kappa = f(\kappa)$ for all regular cardinals κ.

Using Easton Forcing, we can obtain our central result. First, we construct the necessary Easton function:

Lemma 7 *Let:*

$f(\kappa)=\aleph_\gamma$ where γ is the least regular cardinal $> \kappa$.

Then f is an Easton function.

Proof Cardinal arithmetic, left to reader. □

Given f as defined above, we obtain:

Theorem 8 *It is consistent with ZFC that there is no model of HP_i for any $i > 2$.*

Proof It is sufficient to show that it is consistent with ZFC that, for every κ, $2^{(2\kappa)} = \aleph_\alpha$ where $|\alpha| > \kappa$. Let:

$f(\kappa) = \aleph_\gamma$ where γ is the least regular cardinal $> \kappa$.

So there is a model of ZFC where, for each regular cardinal κ, $2^\kappa = \aleph_\gamma$ where γ is the least regular cardinal $> \kappa$. Thus, for any cardinal κ, $2^\kappa \geq \kappa^+$, so $2^{(2\kappa)} \geq 2^{\kappa+}$, but since successor cardinals are regular, this gives us $2^{(2\kappa)} \geq \aleph_\gamma$ where γ is the least regular cardinal $> \kappa$, so $2^{(2\kappa)} = \aleph_\delta$ for some δ, where $\delta \geq$ the least regular cardinal $> \kappa+$. So $2^{(2\kappa)} = \aleph_\delta$ where $|\delta| > \kappa$. □

In other words, if the generalized continuum fails badly enough at every regular cardinal, then it is consistent with ZFC that (a) Upper Hume (HP_2) is satisfiable at singular cardinals, if it is satisfiable at all,[17] and (b) for i > 2, HP_i has no models at all.

Theorem 8 plus Theorem 3 provide the following:

Corollary 9 *For all i > 2:*

"HP_i is conservative"

and:

"HP_i is satisfiable"

are independent of ZFC.

Proof For all i > 2, Theorem 8 provides a model of ZFC where HP_i has no models. □

Corollary 10 *For all* i > 1*:*

"HP_i is irenic"

is independent of ZFC.

Proof For all i > 1, Theorem 8 provides a model where HP_i is not stable (since HP_i either has no models, or has models at only singular cardinals (if any at all) in the case i $= 2$). □

Note that Corollary 10, but not Corollary 9, holds of Upper Hume.

To sum up these formal results: Whether or not the various HP_i's have the properties necessary in order to be legitimate abstractionist definitions is independent of ZFC.

Before moving on it is worth noting that these results provide a new version of the Bad Company objection, one with a new twist not present in the earlier sequence of worries and responses. The variants of Bad Company canvassed in Sect. 2 above were all concerned with determining where we ought to draw the line between acceptable and non-acceptable abstraction principles. At the end of that debate, we tentatively settled on irenicity as at least sufficient (and possibly necessary) for an abstraction principle to provide a legimate definition of a mathematical concept, and conservativeness

[17] At present there seems to be no forcing (or other) method whose application might settle whether or not Upper Hume can be satisfied at singular cardinals. The reason for this is that, although Easton Forcing allows us to vary the value of 2^κ for regular κ at will, constrained only by Konig's Lemma, similar methods which provide the same for singular cardinals have not been forthcoming. Thus, the exact status of Upper Hume (i.e., whether it is conservative) remains an interesting open problem.

as necessary. Nothing in the results above would seem to cast that taxonomical result into question—irenicity and conservativeness are still the criteria for acceptability. Instead, we have replaced a logical or metaphysical worry (what are the criteria for goodness?) with an epistemological one (how, and when, can we know whether an abstraction principle satisfies the criterion for goodness?).[18]

5 Consequences for abstractionism

The formal results of the previous section look, at first glance, to be mostly bad ones, but the critical question remains: How bad are they?

On one way of looking at things, the answer would seem to be "pretty bad". Although we have shown that Upper Hume and the members of Hume's Family are deductively consistent, we should probably not make too much of this—after all, the incompleteness phenomenon allows for the consistency of all sorts of demonstrably unsatisfiable (and thus false) theories—for example, the second-order Peano axioms plus the negation of the Gödel sentence relative to those axioms. As already suggested, if the acceptability of abstraction principles is to hinge (even partially) on consistency, then it is semantic consistency (i.e. satisfiability) that we are interested in.

If so, however, then significant problems loom. Presumably, in order for someone to be in a position to successfully lay down an abstraction principle as a legitimate definition of a mathematical concept, it is not enough merely that (unbeknownst to them) the principle has the right formal properties. In addition, there must be some sort of epistemic requirement to be met—i.e. the person doing the 'defining' must have some sort of assurance that the principle in question is acceptable. In fact, part of the point of the back-and-forth objection-response-new-objection dialectic falling under the 'Bad Company' heading is to determine exactly what those criteria are.

If this is the case, then the justificatory story must go something like this: One can successfully lay down an abstraction principle if, and only if, he has reason to believe that the principle in question displays certain philosophical or technical features. The discussion above has suggested that conservativeness and irenicity are exactly the right sort of features to require of acceptable abstraction principles, but we now need to think a bit more about what is meant by a "reason to believe" that a particular principle has one or both of these characteristics. Simply put, there are (at least) two important components tied up in this notion that need to be examined: First, how strong must our justification be? Do we require a proof, or will something weaker suffice?

[18] Weir's own reservations regarding irenicity as the criteria for acceptable abstraction principles, what he calls 'Embarassment of Riches II', also falls under this general description—in other words, it is not that we have doubts regarding whether or not irencity is necessary and sufficient, but rather, that we have doubts regarding our in principle ability to determine which principles are, in fact, irenic. Weir's worry is a distinct one from that presented here, however, insofar as he merely considers a number of (rather artificial) so-called Distraction Principles, and points out that there seems to be no (non-circular) way to determine whether or not they are irenic, since different background set theories will draw the line in different places. The present problem is even more serious, however, in that we have a number of principles which (at least seemingly) ought to turn out to be acceptable, but which cannot be shown to be so according to our best mathematical theories (even if we allow for circular justifications, i.e. we just flat out assume that ZFC is the correct background theory!).

Second, what is the appropriate background theory within which we can provide such a justification? ZFC? Something else? In other words, how should we fill in the variables in the following:[19]

Constraint Schema:

We have reason to think AP acceptable ↔ BT 'show's that AP is irenic/conservative/whatever.

where "BT" is a placeholder for our background theory and "show" is a placeholder for the appropriate epistemic relation.

It is now pretty much established that, perhaps outside certain atypical foundational areas of mathematics such as category theory or logic itself, ZFC (or, perhaps, ZFC plus appropriate large cardinal axioms)[20] is the de facto arbiter of mathematical legitimacy. In fact, throughout the earlier parts of the present paper (and especially in the last section) we have been using ZFC in exactly this way. So perhaps we should continue with tradition and evaluate the status of abstraction principles from this well-respected vantage point. If this is right (and we shall return to doubts regarding this assumption below), then there are at least two strategies that one might adopt:

The Strong ZFC Constraint:

We have reason to think AP acceptable ↔ iff ZFC proves that AP is irenic/conservative./whatever.

The Weak ZFC Constraint:

We have reason to think AP acceptable ↔ ZFC does not prove that AP is not irenic/conservative/whatever.

There are a number of problems with each of these approaches, however.

At first glance the weak constraint looks rather promising: All members of Hume's Family turn out acceptable on this formulation (since Hume's Principle is irenic, and the irenicity of higher-order variants is independent of ZFC). On closer examination, however, there are two serious problems with this way of fleshing out the required criterion of acceptability.

First off, the weak constraint seems somewhat unmotivated. It suggests that the required justification need not take the form of positive evidence that the principle in question is irenic—we merely need a lack of proof-theoretic evidence to the contrary. But this seems too weak, if not outright bizarre—especially in the face of the incompleteness phenomenon. One way to see this is to note that if we were, for whatever reason, to come to doubt ZFC and replace it with a weaker theory, the result would be that more abstraction principles would turn out to be acceptable, since fewer principles could be proven to be non-irenic.

[19] The discussion that follows merely outlines the complex epistemological issues at hand. For a fuller discussion of these issues, the reader is encouraged to consult Shapiro and Ebert's contribution to this volume.

[20] It is worth noting that large cardinal axioms typically have little or no consequences for structures 'below'—in other words, they tend not to impact the status of the continuum hypothesis, its generalized form, or other independent statements that do not involve large cardinals themselves.

Second, there seems to be a serious technical problem plaguing the weak constraint as formulated above. Notice that Theorem 1 demonstrates that any set of irenic principles is satisfiable, but it does not show that any set of principles that cannot be proven to be non-irenic is satisfiable. The latter claim is significantly stronger than the former (and I suspect it is false). As a result, it seems likely that the Weak ZFC Constraint might, on its present formulation, provide us with reasons to accept incompatible principles. And that is exactly the type of situation that got us in the Bad Company mess in the first place.

So, if we want to adopt something like the Weak ZFC Constraint, then we need to go back and reformulate our definition of acceptable principle. Replacing irenicity with a stronger notion, such as *Super-irenic*:

AP is *super-irenic* $\leftrightarrow$ AP is compatible with all abstraction principles that cannot be proven to be non-conservative.

might do the trick. I leave it to the reader to verify that this, or something like it, provides a notion of acceptability such that the collection of principles that we cannot prove to be unacceptable has a model. Even if such revisions work, however, it is clear that at this point we would be merely attaching epicycles to an initial approach that was flawed in the first place. As a result, we should abandon the weak approach, and consider other alternatives.

In the present context, the other main contender is the strong approach—we have reason to believe that an abstraction principle is acceptable if and only if we can prove that it is. This option does not suffer from the problems of the weak approach, since it is well-motivated and technically sound. The problem, however, is that it rules out the acceptability of Upper Hume and, in fact, all of Hume's Family.

Perhaps this is not the problem that it initially appears to be, however. ZFC, as we have already observed, is currently the litmus test by which the legitimacy of most mathematical theories is judged, but this does not automatically mean that it needs to be the theory by which we judge the acceptability of abstraction principles. In fact, the assumption that ZFC is the background against which we judge the acceptability of abstraction principles seems to rule out the possibility that there might be mathematical truths (such as the claim that certain abstraction principles are okay, or that certain esoteric facts hold regarding the 'behavior' of cardinals) which are themselves knowable through abstraction but which outstrip, or at least differ from, the resources of ZFC itself. Abstractionism, as usually formulated, is a relatively conservative position insofar as it seeks primarily to recapture existing mathematical theories, and not to discover new mathematical truths. Nevertheless, part of the abstractionist story involves the idea that the abstractionist development of a particular mathematical theory is superior to other axiomatic treatments of that same theory insofar as the abstraction principles are definitions of a sort, capturing and clearly displaying what is essential to the mathematical theory in question. As a result, it should not be surprising if occasionally such elegant reformulations lead to new mathematical insights, even ones which are independent of standard set theory.

If this is right—if there is at least a possibility that certain acceptable abstraction principles might allow us to prove set-theoretic claims that cannot be proven in ZFC, then perhaps ZFC is not the right theory within which to settle matters of acceptability (i.e. to prove conservativeness and irenicity results). What, then, is the appropriate theory within which we ought to be examining abstraction principles? The correct answer is both obvious and problematic: abstractionist set theory.

At first glance such a strategy looks promising. Might it not turn out to be the case that our abstractionist reconstruction of set theory, unlike ZFC itself, proves that there is an upper bound on how badly the Generalized Continuum Hypothesis might fail, and thus proves that the HP_i's are irenic? After all, ZFC (and the iterative conception of set that supposedly underlies it) is really, in a certain sense, just a particular sort of reaction to the set theoretic paradoxes of Russell, Cantor, and Burali-Forti—one motivated by a 'constructive', stage-theoretic notion of set formation and an outright ban on non-well-foundedness. Abstractionism, on the other hand, is a different sort of reaction to these very same problems. Isn't it at least conceivable that, as a result, these two different answers regarding how to avoid the paradoxes might provide different stories regarding other aspects of the realm of sets—in particular, different stories regarding the relationship between the size of a set and the size of its powerset?

Perhaps. But there are two serious problems here—one practical, and the other a matter of principle. On the practical side of the equation, we have the well-known failure of extant attempts to reconstruct a theory deserving of the honorific "set theory" within the abstractionist framework. Much work in this area has been done (see Boolos (1989), Shapiro and Weir (1999), Hale (2000), Fine (2002), Shapiro (2003), Cook (2003), and Uzquiano and Jane (2004)), but a suitably straightforward and powerful enough formal theory of abstractionist sets has yet to be forthcoming.

As a matter of principle, however, it is unclear, given the discussion above, how we would know that we had an acceptable set theory in the first place, even if one were laid at our feet.[21] Presumably, the final abstractionist account of set theory will be complex, and thus, will likely have rather subtle model-theoretic traits. As a result, we need some assurance that such a theory is, in fact, acceptable—in other words, we need to prove that the set theory in question is itself irenic. But this looks impossible, since the very abstractionist set theory that we are attempting to prove acceptable is simultaneously the background theory against which acceptability claims are adjudicated. Thus, we need to know that such a theory is okay before we can determine that any abstractionist theory (including the set theory itself) is okay. A vicious circle looms.[22]

Thus, we seem stuck between a rock and a hard place. On the one hand, philosophical considerations seem to suggest that higher-level analogues of Hume's Principle—Upper Hume and Hume's Family—ought to be legitimate if Hume's Principle itself is—in other words, if we can count objects, then we can count concepts. On the other hand, we have what seems to be a legitimate test for the acceptability of an abstraction principle (irenicity), yet the only reasonable set theory we have to hand (ZFC)—and,

21 It is for these reasons that throughout the paper, an 'external' approach was adopted, where we examine the abstractionist project from the perspective of working mathematics, that is, set theory. Ideally, abstractionist theories should be evaluated from an 'internal' perspective, that is, from the view of whatever set theory eventually turns out to be acceptable from within the abstractionist project. Given the vicious circle described above, however, it seems like the external perspective is all we have.

22 This circle seems more problematic than similar-looking circles that plague other positions, since additional considerations of elegance, fruitfulness, and the like seem unavailable to a foundationalist project such as abstractionism.

in fact, perhaps the only reasonable set theory that we *can* have to hand —tells us that, in principle, the test in question does not settle the matter of whether the higher-level versions of Hume are legitimate definitions of number concepts. Further, there seems to be little hope that we can break free of the circle outlined above in order to replace ZFC with some other background theory, in order to (it is hoped) secure the irenicity (or at least conservativeness) of the principles in question. The question, of course, is whether there is any reasonable route out of this mess.

The mathematical results of the previous section are, of course, not in question, and the prospects of replacing ZFC with a theory more amenable to the higher-level abstraction principles, as we have seen, does not seem promising. As a result, the only option seems to be to argue that, contrary to appearances, there is something wrong with the higher-level versions of Hume's Principle, and thus we cannot, contrary to intuition, count concepts (at least, not in the same way that we can count objects by utilizing the resources of Hume's Principle).

I will not attempt to develop in detail what such a maneuver might look like, but will merely mention two strategies that one might take in order to develop this idea. Both involve non-standard understandings of the second-order quantifiers occurring in the abstraction principles in question.

On the one hand, one might adopt Boolos' (1984), (1985) plural reading of the second-order quantifiers. Very roughly, on this reading of the second-order quantifiers, we would interpret:

$(\exists X)\ \Phi(X)$

as saying, not:

There is a concept X such that Φ holds of X.

But instead as:

There are objects X such that Φ holds of the X's.

In other words, the second-order quantifiers, on this reading, are not quantifiers that range over new things (i.e. the concepts), not covered by the first-order quantifiers. Instead, on the plural interpretation, the second-order quantifiers range over pluralities of objects—second-order quantifiers are just a different way of generalizing over the very same things ranged over by first-order quantification.

The reason this move might be advantageous in the present situation is that, if this is the proper way to interpret the second-order quantifiers occurring in abstraction principles, then we need not worry about Upper Hume and the higher-level variants found in Hume's Family. There are serious doubts regarding whether, and how, we can generalize the plural interpretation in order to account for third- and higher-order logic.[23] As a result, on this reading Upper Hume and Hume's Family would not, contrary to the points made in Sect. 3, be legitimate, merely because they would not be expressible in terms of abstractionist vocabulary.

Along somewhat similar lines, the abstractionist could adopt something like the view presented in Rayo and Yablo (2001), whereby second- and higher-order quantifiers do not range over anything. The idea that higher-order quantification is no more

[23] These doubts are not universal, however. Agustin Rayo (2006), for example, attempts to develop a treatment of third- and higher-order quantification within the plural quantification framework.

ontologically committing than the predicates in whose position the higher-order variables occur has come to be called the Yablo-Rayo principle, and is usefully summed up in the following passage:

> "Quine is right, let's agree, that 'there are red houses, roses, and sunsets' is not committed to anything beyond houses, roses, and sunsets, and that one cannot infer that 'there is a property of redness that they all share.' But why should 'they have something in common'—or better, 'there is something that they all are'—be seen as therefore misleading? If predicates are non noncommittal, one might think, the quantifiers *binding* predicative positions are not committal either." ([2001], p. 79)

In other words, first-order quantifiers commit us to objects, since the terms that first-order variables replace are already treated as committing us to objects. Since, on the other hand, the predicates which are replaced by second- and higher-order variables do not (at least, according to Quine) commit us to any additional entities, neither then does quantification where the variables in question occur in predicate position.

On this reading, the problems with Upper Hume and Hume's Principle might well disappear as smoothly as they do on the plural approach. Here, however, instead of Upper Hume and Hume's Family being inexpressible (as they might be viewed on the plural approach) they are expressible but are not legitimate definitions of mathematical concepts. In short: Upper Hume, on the Rayo-Yablo approach, might not be an acceptable principle because it allows us the resources to 'count' concepts, when, in fact, there are no such concepts to be counted.[24]

Of course, either of these approaches, in order to be successful, will require much more detailed development and defense, something I will not pursue here. In addition, there are other ways one might attempt to avoid these problems by altering the background logic—for instance, by restricting the comprehension principle that guarantes the existence of concepts (for an example of this approach, see Oystein Linnebo's contribution to this volume). Instead of further developing such alternative approaches, however, I will close by noting that any such defense will require a detailed account of how second-order logic is meant to contribute to the abstractionist project. Such an account, independent of the new problems presented in this paper, has long been a glaring lacuna in the literature on abstractionism (although see Shapiro and Weir (2000)). Hopefully the worries sketched above will provide further motivation for more closely examining the role of logic in the abstractionist project.

Acknowledgements Versions of this paper were presented at the *Colorado State University, North Carolina State University, The University of Minnesota, Wayne State University*, and *Arché: The AHRC Centre for the Philosophy of Logic, Language, Mathematics and Mind* at the *University of St Andrews*. The paper benefited considerably from the resulting discussion. In addition, special thanks are owed to Geoffrey Hellman, Douglas Jesseph, Oystein Linnebo, Stewart Shapiro, Gabriel Uzquiano, Philip Welch, Crispin Wright, and an anonymous referee for helpful comments and criticism. Any remaining mistakes are, of course, my own.

[24] I am by no means suggesting that either Yablo or Rayo endorse, or would endorse if asked, any of the views regarding number suggested here. The points above, instead, are merely intended to sketch a view of counting that seems compatible with their account of higher-order quantification, as outlined in Rayo and Yablo (2001).

References

Boolos, G. (1984). To be is to be the value of a variable (or to be some values of some variables). *Journal of Philosophy, 81*, 430–449.

Boolos, G. (1985). Nominalist platonism. *Philosophical Review, 94*, 327–344.

Boolos, G. (1989). Iteration again. *Philosophical Topics, 17*, 5–21.

Boolos, G. (1990). The standard of equality of numbers. In G. Boolos (Ed.), *Meaning and method: Essays in honor of Hilary Putnam* (pp. 3–20). Cambridge: Cambridge University Press.

Boolos, G. (1997). Is Hume's principle analytic? In R. Heck (Ed.), *Language, thought, and logic* (pp. 245–261). Oxford: Oxford University Press.

Boolos, G., & Heck, R. (1998). *Die Grundlagen der Arithmetik* §82–83. In G. Boolos (Ed.), *Logic, logic, and logic* (pp. 315–338). Cambridge, MA: Harvard University Press.

Cook, R. (2003). Iteration one more time. *Notre Dame Journal of Formal Logic, 44*, 63–92.

Easton, W. (1970). Power of singular cardinals. *Annals of Mathematical Logic, 1*, 139–178.

Fine, K. (2002). *The limits of abstraction*. Oxford: Clarendon Press.

Frege, G. (1974). *Die Grundlagen Der Arithmetic* (trans.: Austin, J. L.). Oxford: Basil Blackwell.

Hale, R. (2000). Abstraction and set theory. *Notre Dame Journal of Formal Logic, 41*, 379–398.

Hale, B., & Wright, C. (2000). Implicit definition and the a priori. In P. Boghossian & C. Peacocke (Eds.), *New essays on the apriori* (pp. 286–320). Oxford: Oxford University Press.

Heck, R. (1993). The development of arithmetic in Frege's *Grundgesetze der Arithmetik*. *Journal of Symbolic Logic, 10*, 153–174.

Heck, R. (1997a). Finitude and Hume's principle. *Journal of Philosophical Logic, 26*, 589–617.

Rayo, A. (2006). Beyond plurals. In A. Rayo & G. Uzquiano (Eds.), *Absolute generality*. Oxford: Oxford University Press.

Rayo, A., & Yablo, S. (2001). Nominalism through de-nominalization. *Nous, 35*, 74–92.

Shapiro, S. (1991). *Foundations without foundationalism: The case for second-order logic*. Oxford: Oxford University Press.

Shapiro, S. (2000). Frege meets Dedekind: A neologicist treatment of real analysis. *Notre Dame Journal of Formal Logic, 41*, 335–364.

Shapiro, S. (2003). Prolegomenon to any future neo-logicist set theory: Abstraction and indefinite extensibility. *British Journal for the Philosophy of Science, 54*, 59–91.

Shapiro, S., & Weir, A. (1999). NewV, ZF and abstraction. *Philosophia Mathematica, 7*, 293–321.

Shapiro, S., & Weir, A. (2000). Neo-logicist logic is not epistemically innocent. *Philosophia Mathematica, 8*.

Uzquiano, G., & Jané, I. (2004). Well- and non-well-founded extensions. *Journal of Philosophical Logic, 33*, 437–465.

Weir, A. (2003). Neo-Fregeanism: An embarassment of riches. *Notre Dame Journal of Formal Logic, 44*, 13–48.

Wright, C. (1983). *Frege's conception of numbers as objects*. Aberdeen: Aberdeen University Press.

Wright, C. (1997). On the philosophical significance of Frege's Theorem. In R. Heck (Ed.), *Language, thought, and logic*. Oxford: Oxford University Press.

Wright, C. (1999). Is Hume's principle analytic? *Notre Dame Journal of Formal Logic, 40*, 6–30.