One Way to Face Facts

Greg Restall*

Philosophy Department
The University of Melbourne
restall@unimelb.edu.au

version 1.01
February 27, 2004

Abstract: Stephen Neale takes theories of facts, truthmakers, and non-extensional connectives to be threatened by triviality in the face of powerful "slingshot" arguments. In this paper I rehearse the most powerful of these arguments, and then show that friends of facts have resources sufficient to not only resist slingshot arguments but also to be untroubled by them. If a fact theory is provided with a model, then the fact theorist can be sure that this theory is secure from triviality arguments.

Stephen Neale presents, in *Facing Facts* (Oxford: Clarendon Press, 2001), one convenient package containing his reasoned complaints against theories of facts and non-extensional connectives. The *slingshot* is a powerful argument (or better, it is a powerful *family* of arguments) which constrains theories of facts, propositions and non-extensional connectives by showing that some of these theories are rendered trivial. This book is the best place to find the state of the art on the slingshot and its implications for logic, language and metaphysics. It provides a useful starting point for anyone who has wondered what all of the fuss about the slingshot amounts to. Neale shows that the fuss does amount to something, and that theories of facts must "face facts" and present an adequate response to the slingshot.

However, Neale's *evaluation* of the state of play for theories of facts is too pessimistic. As the book draws near to a close, Neale writes:

> As I have stressed, Russell's Theory of Facts, according to which facts have properties as components, is safe. It is certainly tempting to draw the moral that if one wants non-collapsing facts one needs properties as components of facts. I have not attempted to prove this here, but I suspect it will be proved in due course. (page 210)

Neale concludes that while theories which take facts to be structured entities are safe from slingshot arguments, and he suspects that this is the only kind of fact theory safe from slingshot-style collapse. If this were the case, then theories such as situation theories or accounts of truthmakers may well be threatened. However, Neale's suspicion is ill-founded, as I shall soon show. Not only do Russellian theories of facts survive the slingshot unscathed, but so can theories of facts which take them to be unstructured entities. Furthermore, the way that this may be not only argued for, but *proved* can provide a new weapon in the armoury of the theorist investigating fact theories.

*Many thanks to Allen Hazen, Graham Priest and Barry Taylor for helpful comments and encouragement, and to anonymous referees of this journal for useful critical discussion.

This criticism of Neale's understanding of the terrain does not mean that the book is not worth the time and effort required for a close reading. There is a great deal of good sense between its covers. The historical discussion of slingshot arguments is measured and accurate, and the chapter on descriptions is lucid, thorough and convincing, as one would expect from Neale. I recommend this chapter especially to anyone who wants a cogent argument to the conclusion that descriptions can be treated fruitfully as *quantifiers* rather than as referring expressions.

* * *

The core of the book is to be found in the formal presentation of various slingshot arguments. The highlight is a chapter on Gödel's slingshot, which, after years of development and simplification by Neale is a finely honed instrument. I will present it here, with a further simplification, to present the proof in one paragraph rather than over a couple of pages as Neale does. The core idea will be that, allowing some seemingly-innocuous inference principles involving descriptions, we will be able to make seemingly illicit substitutions inside seemingly non-extensional contexts, such as claims about the identity of facts or modal statements.

Neale's version of Gödel's proof requires two different kinds of substitution, one inside a *term* context, and another inside a *sentential* context. A term context $\Sigma(\ )$ (that is, a context $\Sigma(\ )$ such that $\Sigma(t)$ is a well-formed formula if t is a term) is said to be $+\iota$-SUBS if and only if it allows the following substitutivity principles for definite descriptions.

$$\frac{\iota x\phi = \iota x\psi \quad \Sigma(\iota x\phi)}{\Sigma(\iota x\psi)} \qquad \frac{\iota x\phi = a \quad \Sigma(\iota x\phi)}{\Sigma(a)} \qquad \frac{\iota x\phi = a \quad \Sigma(a)}{\Sigma(\iota x\phi)}$$

The context $\Sigma(\ )$ is $-\iota$-SUBS if and only if it is not $+\iota$-SUBS. The other substitution required is in sentential or formula position: A formula context $C[\ ]$ (that is, if ϕ is a formula, so is $C[\phi]$) is $+\iota$-CONV if and only if the inference

$$\frac{C\big[a = \iota x\big(x = a \wedge \Sigma(x)\big)\big]}{C[\Sigma(a)]}$$

is allowed in both directions. The context $C[\ ]$ will be said to be $-\iota$-CONV when it is not $+\iota$-CONV. This inference will feature repeatedly in what follows, so it will pay to introduce a shorthand for the kind of substitution used. Given the formula $\Sigma(x)$ with x free, and a name a, let $\Sigma(\underline{a})$ be shorthand for the term $\iota x\big(x = a \wedge \Sigma(x)\big)$. Then the inference is

$$\frac{C\big[a = \Sigma(\underline{a})\big]}{C[\Sigma(a)]}$$

Note that the "empty context" is $+\iota$-CONV on any reasonable theory of descriptions. If $\Sigma(a)$ is true, then $a = \Sigma(\underline{a})$ is also true, for a is indeed the unique object which is identical to a and is such that $\Sigma(a)$ is true. The converse inference is also clearly valid.

Given these two inferences, we can present a powerful slingshot argument.

$$\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\flat Fa}{\flat(a = F\underline{a})}\,{+\iota\text{-}\textsc{conv}} \qquad \dfrac{\dfrac{Fa}{a = F\underline{a}} \quad \dfrac{Rab}{a = R\underline{a}b}}{F\underline{a} = R\underline{a}b}}{\flat(a = R\underline{a}b)}\,{+\iota\text{-}\textsc{subs}}}{\flat Rab}\,{+\iota\text{-}\textsc{conv}}}{\flat(b = Ra\underline{b})}\,{+\iota\text{-}\textsc{conv}} \qquad \dfrac{\dfrac{Gb}{b = G\underline{b}} \quad \dfrac{Rab}{b = Ra\underline{b}}}{G\underline{b} = Ra\underline{b}}}{\flat(b = G\underline{b})}\,{+\iota\text{-}\textsc{subs}}}{\flat Gb}\,{+\iota\text{-}\textsc{conv}}$$

This shows that if the connective $\flat$ is $+\iota$-CONV and if the contexts $\flat(a = [\ ])$ and $\flat(b = [\ ])$ are $+\iota$-SUBS, then $Fa, Gb, Rab, \flat Fa \vdash \flat Gb$. Since we can take Rab to be each of $a = b$ and $a \neq b$ in turn, we have $\phi, a = b \vdash \psi$ and $\phi, a \neq b \vdash \psi$, and it follows that $\phi \vdash \psi$. It follows that the assumption of Rab is unnecessary, and we have $Fa, Gb, \flat Fa \vdash \flat Gb$, a terrible collapse of the context $\flat$. If $\flat\phi$ is to be read "the fact that Fa = the fact that ϕ" then if Fa and Gb are true, the premises of our reasoning all seem true (as the fact that Fa = the fact that Fa), but the conclusion would be a terrible consequence for a putative theory of facts: The fact that Fa should not be the fact that Gb simply because they both happen to be true.

* * *

What are we to make of this? It's clearly not a happy result for the fact-theorist, as the troublesome inferences seem plausible at first blush. Neale does a good job of showing where fact-theorists have been committed to them. However, the arguments do not mean that we must abandon all talk of facts. It is open for the fact-theorist to reject the collapsing argument. Neale recognises this.

> There is no knock-down argument against facts in this; but it is now abundantly clear that *unless a theory of facts is presented with an accompanying theory of descriptions and an accompanying logic of* [*fact identity conditions*], there is every reason to treat it with caution. The task for the friend of facts is to put together a theory according to which facts are not so fine-grained that they are unhelpfully individuated in terms of true sentences, and not so coarse-grained that they collapse into one. (page 223, emphasis mine)

This diagnosis of the situation overstates the case, and it will be illuminating to consider why. It is quite possible for a friend of facts to present a theory giving *no regard* to a theory of descriptions. Nonetheless, they may still be sure that their theory is consistent and resistant to collapse. This is because the fact-theorist may present and defend a theory of facts by providing a *model* of that theory. Models have many virtues. They provide consistency proofs, and they provide a systematisation of logical consequence. Neale does not consider this route to a theory of facts at all, and his account suffers as a result of it.

Let me illustrate this point by giving a simple back-of-the-envelope model for a theory of facts. The result will be a demonstrably non-collapsing theory of facts, which does not take facts to be individuated by constituent properties but allows them to be totally unstructured entities, and for which the facts "corresponding to" the sentences ϕ and ψ differ if it is possible for ϕ and ψ to differ in truth value.

Take a formal language appropriate for quantified modal logic, with some stock of primitive predicates, terms (constants, function symbols and variables) and with the connectives $\supset$ (material implication), $\sim$ (negation) and $\Box$ (necessitation), and the universal quantifier $\forall x$ as primitive (for each variable x). We leave the other connectives of the language to be defined in the usual fashion. We extend this language with a new "connective" $\triangleright$ to be read as 'is a fact that.' This "connective" is a hybrid between a true connective and a predicate, because it takes a term to its left and a formula to its right to make a new formula. The formula '$a \triangleright \phi$' is read 'a is a fact that ϕ.' We will express our theory of facts in this language.

To interpret this language we will use models for the constant domain quantified modal logic S5. I choose this model theory not because I think that constant domain quantified S5 gets things right, but because it is simple, and because it has been taken seriously by others as providing the logic of necessity and quantification. For our purposes we need treat the model theory merely as an uninterpreted algebra. The consistency proof will stand even if we totally reject the philosophical significance of possible worlds models as a *semantics*. As a purely algebraic construction with no interpretational significance whatsoever, a possible worlds model will still act as a guarantee that the resulting theory is safe from collapse and from slingshot argument. (Nonetheless, any interpretation we might find for the model theory has application as an interpretation of the language we use it to model.)

A model for the language we have chosen involves a nonempty set W of worlds and a nonempty domain D of objects. Each formula will be true at some set of worlds and false at the complement of that set. We aim to have different *facts* for formulas which differ in truth value at some worlds. That is, if ϕ and ψ are not true in exactly the same worlds, then the fact that ϕ should differ from the fact that ψ. One easy way to manage this is for there to be a fact for every set of worlds, so we will do this. We require that the domain D includes the set $P(W)$ of all sets of worlds. It may well include other objects as well, but we do not require this.

Each n-place predicate F is interpreted as function $[\![F]\!]$ from W to subsets of D^n. Given an assignment α of values to the variables (a function from the set V of variables to the domain D, so $\alpha(x) \in D$ is the value of the variable x on the assignment α), we assign the denotation $d_{\alpha,w}(t)$ of each term t in each world w in the usual recursive fashion. This allows for terms to vary in denotation from world to world, but not to have no denotation in a world, as is usual in constant domain modal logics. An x-variant of an assignment α is another assignment of variables α' which assigns the same values to every variable except possibly x. $\alpha[x \leftarrow d]$ is the x-variant of α where the variable x is now assigned the value $d \in D$.

Given a recursive assignment of denotations for terms, we can then assign truth conditions to all of the formulas of the language, relative to assignments of values to variables and to worlds, as follows. ('$\alpha, w \Vdash \phi$' is to be read as 'relative to assignment α and at world w, ϕ is true').

(tc0) $\alpha, w \Vdash F(t_1, \ldots, t_n)$ iff $\langle d_{\alpha,w}(t_1), \ldots, d_{\alpha,w}(t_n) \rangle \in [\![F]\!](w)$.

(tc$\supset$) $\alpha, w \Vdash \phi \supset \psi$ iff $\alpha, w \nVdash \phi$ or $\alpha, w \Vdash \psi$.

(tc$\sim$) $\alpha, w \Vdash {\sim}\phi$ iff $\alpha, w \nVdash \phi$.

(tc$\forall$) $\alpha, w \Vdash \forall x \phi$ iff $\alpha', w \Vdash \phi$ for all x-variants α' of α.

(tc$\Box$) $\alpha, w \Vdash \Box\phi$ iff $\alpha, v \Vdash \phi$ for all $v \in W$.

(tc$\triangleright$) $\alpha, w \Vdash t \triangleright \phi$ iff $d_{\alpha,w}(t) = \{w : \alpha, w \Vdash \phi\}$ and $\alpha, w \Vdash \phi$.

The only clause which is at all innovative is (tc$\triangleright$). The formula $t \triangleright \phi$ is true at a world w if and only if the term t denotes the set of worlds at which ϕ is true, and w is one of those worlds. So $t \triangleright \phi$ is true at w only when ϕ is true at w.

Our formal "theory of facts" is the collection of all formulas true at every world in every model. This theory includes some oft-claimed truisms about facts. Here are two examples:

$$\phi \equiv (\exists x)(x \triangleright \phi)$$

— ϕ is true if and only if there is a fact that ϕ.

$$(a \triangleright \phi) \wedge (a \triangleright \psi) \supset \Box(\phi \equiv \psi)$$

— a fact can be a fact that ϕ and a fact that ψ only if ϕ and ψ necessarily stand or fall together. Many other such claims may be verified by reasoning about all models in the usual fashion. We get a simple "theory," extending the familiar modal logic of constant domain quantified s5 with a new operator $\triangleright$.

This "theory of facts" may be verified to be non-collapsing by providing a simple model. Take a model featuring two worlds w_1 and w_2, with a statement Fa true at both w_1 and w_2 but Gb true only at w_1. Then $w_1 \Vdash Fa$ and $w_1 \Vdash Gb$ but $\alpha, w_1 \Vdash t \triangleright Fa$ only when $d_{\alpha,w_1}(t) = \{w_1, w_2\}$ ($t \triangleright Fa$ is true only when t denotes the set of *all* worlds in the model, as Fa is true everywhere) and $\alpha, w_1 \Vdash t \triangleright Gb$ only when $d_{\alpha,w_1}(t) = \{w_1\}$ ($t \triangleright Gb$ is true only when t denotes the set $\{w_1\}$), so this model is a guard against collapse of the theory of facts.

So, commitment to any of the inferences validated by these models will never result in a trivial theory of facts. Of this we can be completely sure. If any collapse threatens, it must come from *outside* this theory. This theory demonstrably does not collapse, while at the same time, this theory demonstrably does not take facts to be composed of properties or any other such thing. The *theory* is silent about the composition of facts. The *models* of the theory take them to be sets of worlds, but that is not a part of the theory. The model theory is a technique to provide a tool for separating valid and invalid inferences in the formal language, and it does this job even if we take the model theory to be merely an algebraic construction devoid of other semantic significance.

This 'theory' is a toy, and I do not mean to propose it as a serious theory of facts. Nonetheless, it has a very serious consequence for any genuine theory of facts. The class of models we have seen assures us that for any *genuine* theory of facts, which restricts itself to claims endorsed in these models, is demonstrably non-collapsing. So, perhaps we do not endorse all of constant domain quantified s5. Perhaps we do not endorse all of the properties this theory takes $\triangleright$ to have. If our genuine theory of facts is properly *weaker* than this theory, it is still demonstrably non-collapsing. Slingshot arguments can have no effect against any such theory.

Notice what we have done: We have proved that this account of facts is non-collapsing, even though we have said nothing about descriptions. This means that our theory is incomplete and it needs supplementation if it is to tell us what to say about the validity of arguments involving descriptions, but it does not mean that this theory, in its incomplete state, is under any suspicion of collapse. We can be completely confident that if we add a theory of descriptions which can be interpreted using the models we have presented then this extension of the theory is secure against slingshot arguments. I will demonstrate this by providing two different interpretations for descriptions.

The first interpretation of descriptions is Russellian. The phrase 'the fact that ϕ' can be translated in a Russellian fashion, taking $[\text{the}_x : \phi(x)]\psi(x)$ to be shorthand for the pre-existing formula in our language: $(\exists x)(\phi(x) \wedge (\forall y)(\phi(y) \supset y = x) \wedge \psi(x))$. This interprets descriptions in without extending the original language in any way at all, and we are no nearer collapse than we were before.

The second interpretation of descriptions is referential. In this case, we add a denotation clause for the new term $\iota x\phi$ as follows:

(dι) $d_{\alpha,w}(\iota x\phi)$ is the unique $d \in D$ where $\alpha[x \leftarrow d], w \Vdash \phi$ if there is such a d, or it is $\emptyset$ otherwise.

This takes definite description terms to refer to the unique object satisfying them, if there is any such object, or in cases where the standard reference fails, it takes them to refer to the object $\emptyset$ in $P(W)$. (This seems like a natural choice because $\emptyset$, the empty set of worlds, will never be the denotation of a succesful attribution of fact-hood, as $t \rhd \phi$ is true only when ϕ is actually *true* at some world.) This choice for definite descriptions makes them genuinely referential, and it is a proper extension of the language. It also introduces no new threat of collapse, because the rest of the language is interpreted as before, and our counterexamples still stand.

In making these definitions for descriptions, it was not as if I had to struggle to find accounts for descriptions which would break either ι-CONV or ι-SUBS. I simply took two pre-existing accounts of definitions "off the shelf" and applied them. The models themselves dictated that one of ι-CONV and ι-SUBS would fail. In these cases it is +ι-SUBS. *All* contexts definable in our language are +ι-CONV provided that $\iota x\big(x = a \wedge \Sigma(x)\big)$ and $\Sigma(a)$ are true at the same worlds (they agree in *intension* as well as *extension* on our models). This holds for both of our accounts of descriptions, so ι-CONV will hold for every context expressible in our language.

The mistake in the collapse inference, according to our theory, is therefore ι-SUBS and a straightforward counterexample is not difficult to find. Consider the putative inference

$$\frac{\iota x Fx = a \quad t \rhd (a = \iota x Fx)}{t \rhd (a = a)}$$

which is an ι-SUBS inference for the context $t \rhd (a = [\ ])$. This inference fails in our theory (on either account of descriptions, whether they are referential or not) because $\iota x Fx = a$ may be true in some worlds (where the denotation of the name a is the unique object satisfying Fx) and not others (where the denotation of the name a is no longer the only object satisfying Fx). So, if $t \rhd (a = \iota x Fx)$ is true, then the denotation of t is the set of all worlds where $a = \iota x Fx$ is true. But this is *not* necessarily the set of all worlds whatsoever, which is what is required if $t \rhd (a = a)$ is to be true. So, it is straightforward to show that the context $t \rhd (a = [\ ])$ is −ι-SUBS, irrespective of whether descriptions are treated referentially or in a Russellian manner.

This model theory is not to be taken *too* seriously as a genuine contender for a model theory for a genuine theory of facts. I leave it to the fact theorist to show

how more comprehensive and interesting theories may be shown to be consistent.[1] It is merely a single example of a general technique. It shows to prove that many plausible inference principles involving facts and descriptions are unproblematic and secure against slingshots, no matter how they are refined. Neale seems either ignorant or unimpressed with this reasoning: he contents himself with sharpening up the slingshot arguments, and he leaves any non-triviality proofs to others. This leaves the uncommitted reader—who is agnostic on the matter of the triviality of fact theories in general and who wishes to gain an understanding of what works and what doesn't—feeling distinctly unsatisfied after reading the book. After all, logic is not just *proof* theory, it is also *model* theory. One can use ever-more sophisticated slingshot arguments to approach the boundary between the fact theories which work and the fact theories which don't from one side, but no matter how far you advance, this will not tell you as much as if you also advance to that boundary from the other side. A more balanced work on the topic would have approached this boundary from *both* sides. Neale's *Facing Facts* reminds you of the boxer who fights with only one fist. It is capable as far as it goes, and it is remarkable how well he does with the tools he has allowed himself. Nonetheless, it is ungainly. A more deft work, at the one time more measured and judicious, yet more interesting and definitive, would have resulted had Neale availed himself of the other fist.

* * *

In the abstract I claimed that friends of facts need not be troubled by slingshot arguments. Here is how I can make good this claim: Provided that a theory of facts has a model, it is resistant to collapse arguments. Any slingshot argument to a repugnant conclusion not true in the model must appeal to a principle not endorsed in that model. Models, then, can provide a guide to the options available to the friend of facts in resisting slingshot arguments. To make the point using a different metaphor, triviality arguments on the one hand, and models on the other, mark out different kinds of territory on the map of theories of facts. Slingshot arguments show that certain places depicted on that map are uninhabitable. They show that particular combinations of principles are incoherent. Models, on the other hand, show that other places on the landscape are safe. Given a model, the principles endorsed in that model are coherent and non-collapsing.

Of course, the coherence or consistency of a collection of principles is one thing, and its truth is another. Models for fact theories give us all the assurance we need that those theories are consistent, that life on that patch of land is possible. It is another thing altogether to decide that we should take up residence there. We will need more than just a slingshot to make that decision.

[1]This is not merely a *request* for future development of fact theories. There are extant theories of possibilities, facts, truthmakers and the like, developed with *models* which are given a treatment such as this. Neale never addresses this material. Not once does Neale broach the use of models as a technique for explaining where a slingshot argument might go wrong.