INTUITION AND RUSSELL'S PARADOX

MARGARET CUONZO

Long Island University, Brooklyn Campus

Abstract

In this essay I will examine the role that intuition plays in Russell's paradox, showing how different approaches to intuition will license different treatments of the paradox. In addition, I will argue for a specific approach to the paradox, one that follows from the most plausible account of intuition. On this account, intuitions, though fallible, have epistemic import. In addition, the intuitions involved in paradoxes point to something wrong with concept that leads to paradox. In the case of Russell's paradox, this is an ambiguity in the notion of a class.

1. Definitions of "Paradox"

Intuition, which has been variously defined as "seeming truth", "spontaneous mental judgment", "what we would say" in a given situation, and "non-inferential belief", has a prominent place in each of the two standard definitions of the term *paradox*. On one definition (cf. Sainsbury), a *paradox* is a logical argument with *seemingly* true premises, employing *seemingly* correct reasoning with an *obviously* false or contradictory conclusion. The *seeming* truth of the premises, *apparent* correctness of the reasoning, and *obvious* falsity of the conclusion distinguish a mere argument from a paradox. On another definition, a *paradox* is a set of mutually inconsistent propositions, each of which, taken individually, *seems* true (cf. Schiffer). A set of mutually inconsistent propositions is not a paradox unless each proposition, taken individually, has strong intuitive force. On both definitions, it is a necessary condition for being a paradox that we have strong intuitions about the truth-values of the parts. In fact,

it would not be too far off the mark to define a *paradox,* as "a set of mutually inconsistent intuitions in which each intuition is individually very strong". On this definition, it is neither the argument nor the set of mutually inconsistent propositions that is emphasized, but rather the intuitions. Just as there are many sets of intuitions that are not paradoxes, there are many arguments and sets of mutually inconsistent propositions that are not paradoxes, either. Indeed, this third definition is more faithful to the etymological roots of the term *paradox,* which comes to from the Greek terms for "against" or "beyond" (*para*) and "expectation" or "opinion" (*doxa*). The Greek terms emphasize the counterintuitive nature of paradoxes, and not that they are arguments or sets of propositions. Since on each of the three definitions, a *paradox* is defined in terms of the intuitive plausibility of its components a successful account of a paradox will give an accurate analysis of these intuitions.

2. Overview of Russell's Paradox

Russell's paradox, for example, contains each of the three elements mentioned as definitive of a paradox. (It has premises that seem true, a conclusion that is obviously false, and seeming validity) (cf. Sainsbury):

Let R be the class of all classes that are not members of themselves.

1. For any object x, x is a member of R if, and only if it is not the case that x is a member of itself.
2. R is a member of R, if, and only if it is not the case that R is a member of itself.
3. Therefore, R is a member of R if, and only if it is not the case that R is a member of R.

Or, more simply:

1. For any class x, $x \in R$ iff $\neg x \in x$.
2. $R \in R$ iff $\neg R \in R$.

In these arguments we have intuitively plausible premises, apparently correct reasoning and a contradictory conclusion. Although the condition of R may be somewhat hard to read, there is no *prima facie* problem with a class of classes that do not contain themselves as members. There are classes, it seems, that do contain themselves as members. The class of classes is, it seems, a member of itself. In addition, there are many classes that do not contain themselves as members. For example, the class of flowers is not a flower and is therefore not a member of itself. So why not a class of classes that are not do not contain themselves as members? The condition for R is licensed by Cantorian, "naïve" set theory's principle of abstraction. The principle of abstraction holds that "A formula P(x) defines a set A by the convention that the members of A are exactly those objects a such that P(a) is a true statement". That is, a formula is the defining property of a set if, and only if, all and only members of that set satisfy the formula. As a result, every property determines a set. For properties such as being a round square, the set is empty. And so the property of being a non-self membered class, it follows, determines a class as well. Next, R replaces the x in (1), leading to the contradictory conclusion. So, that's the paradox.

3. Problems Raised by Russell's Paradox

As Charles Chihara notes in his article, "Semantic Paradoxes", there are two main problems that would-be solutions to paradoxes typically address. The first, called the "diagnostic problem", concerns the notion that leads to paradox. To solve the diagnostic problem a would-be solution would expose what it is about the relevant notion, in our case, the notion of a class, which leads to Russell's paradox. The second problem is the "preventative" problem of constructing a logical system in which the paradox does not arise. For example, Russell's solution to the "diagnostic" problem of the paradox was his Vicious Circle Principle (VCP) and his solution to the "preventative" problem posed by paradox was the theory of types. VCP claims that no totalilty, such as a class or a statement, can contain members that are fully specifiable in terms of itself. Russell's class R is specified as

the class of all classes that are not members of themselves. That is, for any class x, $x \in R$ iff $\neg x \in x$. This specification concerns a totality: all classes. Since it holds of any class, it should hold of R as well. But in order to specify R, this would have to be done by saying that R is a member of R if, and only if it is not the case that R is a member of R. But this is specifying R in terms of the totality (i.e., R). R cannot belong to R because this would violate the VCP. R is specified as a totality, but if R were to be a member of itself, it would have to be fully specified in terms of that totality. And this violates VCP.

Russell's theory of types uses VCP to justify why objects ought to be arranged in hierarchies according to different types. The first type, or lowest level, is that of individuals. Next, is classes of individuals, then classes of classes of individuals, and so on. On the theory of types, classes should not be formed with members from levels that are higher than or the same as the class itself. For example, a class of classes cannot have a class of classes of classes as a member. Now take R:

$$\text{For any class } x,\ x \in R \text{ iff } \neg x \in x.$$

R, being a class of classes, is on a higher level than x, which describes a class. To plug R into this and derive the contradiction:

$$R \in R \quad \text{iff} \quad \neg R \in R$$

we have violated the hierarchy of levels. The above is not an intelligible condition because the R is on the same level as itself. On the theory of types, since the condition for R contains "$\neg x \in x$", the condition for R is unintelligible as well. Since x denotes an individual, it is a violation of the theory of types to talk of individuals being members of individuals.

4. Happy-face and Unhappy-Face Solutions to Russell's Paradox

VCP/theory of types' solution to the Liar attempts to show that the first premise of the paradox is unintelligible. It then attempts to show how our intuitions regarding premise one are mistaken. Although it

is licensed by an unrestricted use of the principle of abstraction, and it *prima facie* looks perfectly intelligible, the premise is unintelligible given the VCP and theory of types. Such a solution is what Stephen Schiffer has called a *happy-face solution*. According to Schiffer:

> A *happy-face* solution to a paradox would do two things: first, it would identify the odd-guy-out, the seemingly true proposition that isn't really true; and second, it would remove from this proposition the air of seeming truth so that we could clearly see it as the untruth it is (20, italics Schiffer's).

Although premise one is licensed by an unrestricted use of the principle of abstraction, the intuitiveness of the principle is removed by the theory of types. Thus, our intuitions about the principle of abstraction, and what follows from it, were mistaken.

Other types of solutions to paradoxes, what Schiffer has called "unhappy-face" solutions, give diagnoses of the cause of the paradox, but do not attempt to provide a logical system on which the paradox can be avoided. For example, such a solution might claim that the paradox proves that the relevant notion is incoherent and no new logical system can sidestep the paradox. An example of this kind of solution is Michael Dummett's solution to the sorites paradox. Such solutions diagnose the cause of paradox and show how no successful solution to the preventative problem of sidestepping the paradox can be given. Since the VCP/theory of types solution attempts to solve both the diagnostic and preventative problems raised by the paradox, it is best thought of as a happy-face solution to the paradox.

But here two deeper questions emerge: What justifies VCP and its accompanying theory of types, other than the fact that an unrestricted use of the principle leads to a paradox? And, more importantly, does Russell's paradox admit of a happy-face solution at all?

The VCP/theory of types solution has often been criticized as being *ad hoc*, having its motivation derived solely from their ability to sidestep Russell's and other paradoxes. Indeed, Russell himself does not give much of an argument for the Vicious-Circle Principle.[1] However, I believe that the VCP/theory of types solution to the paradox, although flawed, is not *ad hoc*. The solution is licensed by an account of intuition that treats certain types of intuition as having

little or no philosophical import: Although premise one is *prima fa-cie* plausible, it leads to contradiction. Thus something must have gone wrong somewhere in the argument. It is only natural, given this specific approach to intuition, that premise one is flawed. On this approach, because premise one leads to contradiction, our intuitions regarding it must be mistaken. Intuition, when faced with contradiction must be abandoned. And although I have been focusing on Russell's solution to the paradox up until now, a similar analysis can be given for the formalist, Zermelo-Frankel, and even intuitionist treatments of the paradox. The formalist, in restricting logic to the rules of inference that are absolutely certain and to finite, well-defined and constructible objects, rejects premise one. And the intuitionists who claimed that one cannot assert the existence of mathematical objects without also indicating how these objects can be constructed reject premise one as well. Even Zermelo's axiomatization of set theory restricts the principle of abstraction so that premise one is again un-intelligible. Although each response to the paradox is distinct, each rejects premise one. The grounds for this rejection vary, but when the point is pushed far enough, the ultimate justification is that to take another approach would lead to paradox. When faced with a contradiction, the system must be revised in a way that avoids the contradiction. This is not an outrageous idea, to say the least. And what happens to R and its intuitive plausibility? Our intuitions were mistaken.

5. Intuition and Russell's Paradox

In order to show how this approach to intuition leads to solutions like Russell's and the others, it will be helpful to get a sense of the kind of intuition that premise one (R: for any class x, $x \in R$ iff $\neg x \in x$) engenders.

Intuition has usually been divided into two classes corresponding to what Hume calls "relations of ideas" and "matters of fact", or what Hintikka calls "intuitions concerning empirical truths and those concerning conceptual (including linguistic) ones" (143). Examples of empirical intuitions are my intuitions regarding my being in New

York, that New York is in the U.S., and perhaps that it is wrong to murder an innocent child. Examples of conceptual intuitions are intuitions about the correct way to use a language and intuitions about logical truths.

Premise one of Russell's paradox determines the condition for membership in R. There is no appeal to matters of fact. Instead, premise one's plausibility is due to it being a stipulation of the condition for membership in a class and therefore, premise one falls into the "relations of ideas" or "conceptual" category. Of the types of intuition this is the one taken most seriously, and often thought of as least likely to be false. Yet these categories may be somewhat artificial, and even conceptual intuitions are ones that are subject to revision.

Recently, cognitive psychologists and philosophers have called the reliability and philosophical relevance of each form of intuition into question. Relatively recent studies in cognitive psychology, such as Nisbett and Ross (1980) suggest that with regards to inductive inference, the rules that best capture intuitive judgments that people make are unacceptable (Ramsey and DePaul, 1998). Studies on the rationality of betting behavior suggest a similar conclusion. The critique from cognitive psychology applies especially to reflective equilibrium, a notion made explicit by Nelson Goodman and John Rawls. According to Goodman and Rawls, philosophers start with basic beliefs that they are forced to accept as a starting point. In using the method of reflective equilibrium, philosophers attempt to mold their intuitive judgments into a coherent whole. In some cases, philosophers will revise systems and theories in light of conflicts with intuitive judgments. Many famous cases of this kind of revision come from Plato's dialogues. For example, consider Cephalus' definition of "justice" in Plato's *Republic* as "telling the truth and paying your debts". A counterexample is proposed, the friend who has lent a weapon when sane, and now insane, wants it back so that he can kill someone. In this case, the intuition that it would not be right to return the weapon or speak truly to the friend is offered as support that the definition of "justice" as "telling the truth and paying your debts" is wrong (cf. DePaul and Ramsey, vii).

In addition, in a recent essay titled "The Emperor's New Intuitions", Jaakko Hintikka argues forcefully against (a) treating intu-

itions as support for a philosophical thesis, and (b) treating the ex-
planations of our intuitions as a worthy philosophical enterprise. He
writes, "The most amazing fact about the current fashion of appeal-
ing to intuitions is the same as the proverbial dog's walking on two
feet: not that it is done well but that it is done at all. For what is
supposed to be the justification of such appeals to intuition? One
searches the literature in vain for a serious attempt to provide such
a justification... This blind faith is below the intellectual dignity of
philosophers for whom unexamined intuitions are not worth intuit-
ing" (*JOP*, March 1999). Harsh words. Hintikka sees the reliance of
contemporary analytic philosophers such as Saul Kripke on intuition
as having its most recent roots in Chomsky's linguistics. Chomsky,
however, was an unabashed Cartesian. In being a Cartesian, innate
ideas, inner reflections, etc. were infallible. Those who do not sub-
scribe to the view that there are such ideas must provide some kind
of justification of the use of intuition.

6. Three Accounts of Intuition and Their Corresponding Approaches to Paradoxes

Although Hintikka does not claim that all intuitions are altogether
unreliable (he has suspicions about empirical ones only), the section
quoted above, coupled with contemporary questions about the dis-
tinction between conceptual and empirical truths, could lead to a
more encompassing critique of the use of intuition in all its forms.
I'll call this position "radical anti-intuitionism". As Gary Gutting in
introduction to *Rethinking Intuition* explains, "Philosophers claiming
to have special access to a body of analytic truths have been con-
fronted with Quine's critique of the analytic-synthetic distinction;
those proposing to logically construct knowledge from basic sensory
givens have encountered Sellar's critique of the theory-observation
distinction; those hoping to make philosophy an investigation of the
a priori conceptual schemes through which we experience the world
have met Davidson's rejection of the scheme-content distinction.
These critiques have ... made philosophers far more uneasy about
the intellectual tools they have used and have led philosophers to

see their discipline as much more closely tied (if not assimilable to) empirical scientific inquiry" (6). So perhaps conceptual intuitions like premise one are unreliable as well.

An account of philosophical paradoxes, and indeed an approach to philosophy more generally follows from radical anti-intuitionism. Since a *paradox* is defined in terms of intuition, and the task of solving a paradox involves giving some account of our intuitions, paradoxes do not carry with them much philosophical import. On this account, explaining our intuitions regarding the components of a paradox is not a worthwhile endeavor. If intuition provides absolutely no reason to believe that a statement is true (or false), then a paradox merely points to the fact that sometimes our intuitions are in conflict.

One consequence of radical anti-intuitionism is that most, if not all, of the famous philosophical problems turn out to be unworthy of philosophical attention, and the standard philosophical treatments of these problems are not only unnecessary, but ineffectual. From Plato's thought experiments in the *Republic* to Putnam's Twin Earth and beyond, the thought experiment is a central tool of philosophy. To critique intuition in so radical a way is really to question the value of philosophical inquiry. This is neither to say neither that a critique of appeals to intuition is unneeded nor that it is unnecessary to provide an adequate justification of appeals to intuition. I am simply pointing out that it is a consequence of such a position that almost all forms of intellectual inquiry are baseless. Because Russell and the other philosophers of mathematics treated the paradox as posing a serious threat to set theory and the foundations of mathematics, they did not hold this position.

Instead Russell and the others took the second approach to be discussed here, a hierarchical approach which to intuitions that privileges some forms of intuitions over others. An example of such a solution is the epistemic solution to the sorites paradox. The paradox can be phrased in the following way:

1) A person with 0 hairs is bald.
2) For any number n, if a person with n hairs is bald, then a person with $(n+1)$ hairs is bald.
3) A person with 1,000,000 hairs is bald.

In the above paradox, the first premise points to the paradigm case of baldness. The conclusion, on the other hand, points to the paradigm of non-baldness. The epistemic theorist criticizes the second premise, commonly called the "sorites premise". The sorites premise claims that the difference of one hair is too small to warrant a change in classifying anyone as bald or non-bald. Since we cannot even detect such a difference we cannot say of two people, one with n hairs and the other with $(n+1)$ hairs, that one is bald and the other not. The sorites premise, according to the epistemicist, merely *seems* true. We take the premise to be true because we are conflating our inability to discern the sharp cut-off between baldness and non-baldness with the lack of such a cut-off. Why must the sorites premise be rejected? According to epistemic theorist Timothy Williamson, it is because it conflicts with the principle of bivalence. On this principle, roughly, every utterance of a declarative sentence is either true or false. In a series of utterances of the form "A person with n hairs is bald" in which the first sentence substitutes 0 for n, the second sentence substitutes 1 for n, and so on up until $1,000,000$, there will be a point in which the sentences shift from being true to being false. It follows from bivalence that there must be a sharp cut-off between baldness and non-baldness, otherwise there will be declarative sentences that are neither true nor false. Williamson claims that because of this conflict the sorites premise is false. And he explains why the sorites premise is intuitively plausible: since we are essentially ignorant of the cut-off between baldness and non-baldness, we confuse our inability to know the cut-off number with there being no such number.

Russell's, and the other treatments of the paradox is similar to the epistemic solution to the sorites paradox in that they give happy-face solutions to a paradox and involve giving up an intuitively plausible premise in order to avoid conflict with a more basic rule/principle. For Williamson, the principle of bivalence must not be violated, while for Russell, violating the principle of noncontradiction is to be avoided. In answer to the first question I posed a few minutes ago about what motivates Russell's VCP/theory of types solution to the paradox independent of avoiding the paradox, the answer is that it the solution is licensed by a specific approach to intuition, one which holds

that certain intuitions are to be forsaken in the light of other, more foundational ones.

However, in the case of solutions to paradoxes, the approach that is licensed by this account of intuitions is not very useful. Typically, paradoxes are not successfully solved this way. Exposing one intuition as "the odd guy out" usually doesn't address the main concerns regarding the notion that leads to paradox. In the long history of philosophical paradoxes, there has yet to be a successful solution to a genuine philosophical paradox that involved exposing some seeming truth for an untruth. On inductive grounds alone, such a solution is not plausible.

Thus, the answer to second question about whether Russell's paradox has a happy face solution is, I believe, no. Instead, it is better to claim that the paradox exposes some problem with the intuitive notion of a class, and then simply show how there can be no solution to the paradox that keeps the intuitive notion intact. To replace the intuitive notion with one that does not lead to paradox does not successfully treat the phenomenon that leads to the paradox. Consider the idea of a class. A class, intuitively, is a totality. Yet totalities can contain other totalities as members. Thus a class can be both a totality and a member of some other totality. The paradox arises when a class is treated as both a totality and a member of that same totality. The intuitive notion licenses both, but this leads to paradox. This may seem similar to the VCP, but there is a crucial difference; the VCP is a rule for defining conditions for membership in a class, while this treatment merely shows what leads to the paradox. The VCP says, "Don't do that". There is no "ought" in the present treatment. The intuitive notion of a class is ambiguous. And while the theory of types attempts to provide a "solution" to the preventative problem raised by the paradox, a better treatment would explain why there can be no such preventative measures that both (a) sidesteps the paradox; and (b) leaves the original notion basically intact.

A third and , I believe, best approach claims that all forms of intuition are fallible, but also claims that such fallible intuitions are epistemically useful nevertheless. Treating intuition in this way licenses what I take to be the best approach to philosophical paradox. Al-

though he does not discuss philosophical paradoxes Hilary Kornblith in "The Role of Intuition in Philosophical Inquiry: An Account with No Unnatural Ingredients". (DePaul and Ramsey, 129–41) provides an interesting justification of the use of strong, shared intuitions. According to Kornblith, wholly idiosyncratic intuitions are not useful philosophically. If no one shares my strong intuition, not only is my intuition not persuasive to others, but this lack of agreement ought to force me to question whether I am engaged in a discussion of the same phenomenon. Why privilege the intuitions of the majority over the one lonely but strong intuition? As Kornblith explains, "the extent of agreement among subjects on intuitive judgments is to be explained by common knowledge, or at least common belief, and the ways in which such background belief will inevitably influence intuitive judgment, although unavailable to introspection, are nontheless quite real" (134). If I have a strong but idiosyncratic intuition about something, there will be some background belief that I don't share with those that do not share my intuitions. Thus, intuitions are important at least to the degree that they point to background beliefs that may (or may not) be shared.

On the solution to philosophical paradoxes licensed by this account, the intuitions in a paradox, which are each very strong and held by most people, expose that there is something wrong with the background beliefs, namely, a flaw in the notion that leads to paradox. For Russell's paradox, this is the notion of a class.

7. Conclusion

In conclusion, Russell's and other treatments of the paradox reject the condition for membership in R. They then attempt to provide a logical system on which the paradox does not arise. I have argued that this preventative project is doomed to failure and that Russell's paradox does not admit of such treatment. A better way to treat the paradox is to show what leads to the paradox, and then to show how there can be no solution to the preventative problem raised by the paradox.

References

Chihara, Charles. "Semantic Paradoxes: A Diagnostic Investigation". *Philosophical Review* **88**: 590–618.

Depaul and Ramsey (eds.) *Rethinking Intuition.* New York: Rowman and Littlefield, 1998.

Hintikka, Jaakko. "The Emperor's New Intuitions". *Journal of Philosophy.* March, 1999.

Sainsbury, P. M. *Paradoxes.* New York: Cambridge UP, 1995.

Schiffer, Stephen. "Two Issues of Vagueness". *The Monist,* Spring 1999.

Williamson, Timothy. *Vagueness.* New York: Routledge, 1996.

Keywords

Margaret Cuonzo
Philosophy Department
Long Island University, Brooklyn Campus
1 University Plaza
Brooklyn, NY 11201 USA
Mcuonzo@aol.com
mcuonzo@liu.edu

Note

[1] It has been suggested to me that the reason that Russell did not present evidence for the truth of VCP was that he believed Poincare' to have provided adequate justification of the principle.