2022 TheoLogica
An International Journal for Philosophy of Religion and Philosophical Theology
S. I. ANALYTIC THEOLOGY AND THE TRI-PERSONAL GOD
DOI: https://doi.org/10.14428/thl.v6i2.63893

Richard Cartwright on Logic and the Trinity

PETER VAN INWAGEN
The University of Notre Dame
vaninwagen.1@nd.edu

Abstract: It is often maintained that the doctrine of the Holy Trinity implies a contradiction. It is sometimes maintained that if the doctrine is formulated in a way consonant with the thesis that "identity is always relative to a sortal term," it can be shown that it does not involve a contradiction. Richard Cartwright has contended that an appeal to "the relativity of identity" cannot change the fact that the doctrine of the Trinity is inconsistent with a principle that is "evident to the natural light of reason," namely *If every A is a B, then there cannot be fewer B's than A's.* The purpose of this essay is to examine and evaluate that contention.

Keywords: Peter Geach, Richard Cartwright, Trinity, Relativity of identity, Formal contradiction

1. Introductory Remarks

Many philosophers believe that any statement of the doctrine of the Holy Trinity must be (if truly orthodox) self-contradictory. And a few philosophers following Peter Geach, have maintained that that a certain radical position in philosophical logic—that "identity is always relative to a sortal term"—implies that the appearance of contradiction on which the belief of the many is founded is illusory.

Richard Cartwright has contended that the principle

> If every A is a B, then there cannot be fewer B's than A's,

(a) is "evident to the natural light of reason," and (b) when conjoined with various statements integral to the doctrine of the Trinity implies tritheism—the doctrine that that there are three Gods.[1] And tritheism is not only heretical, but

[1] Richard Cartwright (1987, 187–200). See particularly pages 196 and 197.

contradicts the monotheism that is an essential component of the doctrine of the Trinity. If, therefore, "Cartwright's Principle" is true, and if Cartwright's deduction of 'There are three Gods' contains no logical error (and in fact it does not contain any logical error), the doctrine of the Holy Trinity is self-contradictory.

My topic in this essay is Cartwright's deduction. I shall pay special attention to a claim he makes for this deduction, namely that it is valid and sound even if Cartwright's Principle is true. Most of the essay, however—the part comprising sections 2, 3, and 4—is devoted to an exposition of the logic and philosophy of the idea of relative identity and its theological applications. The content of sections 2, 3, and 4 is applied in an analysis of Cartwright's Principle ('If every A is a B, then there cannot be fewer B's than A's') in the final section, Section 5. It is there argued that Cartwright's argument fails, since Cartwright's Principle presupposes "absolute counting," which is not possible if identity is always relative to a sortal term.

2. The Question of the Logical Consistency of the Doctrine of the Holy Trinity: Identity and Relative Identity

The following three statements would seem to be essential components of the doctrine of the Holy Trinity:

1. The Father is God
2. The Son is God
3. The Father is not the Son.

But these three statements appear to imply a logical contradiction. They *certainly* imply a contradiction if to say that x is y is to say that x is numerically identical with y. (In this essay, 'identity' is always to be understood as 'numerical identity'.) That is, the three statements

1′. The Father = God
2′. The Son = God
3′. The Father ≠ the Son.

imply a contradiction—for, owing to logical rules grounded in the fact that

identity is both symmetrical[2] and transitive, the contradictory of (3′) is an immediate logical consequence of (1′) and (2′).

I hope it goes without saying that one cannot refuse to deal with this (or any) alleged proof that the doctrine of the Trinity implies a contradiction on the ground that God is "above logic." (If it does not go without saying, I hereby say it.) Professor Geach has said of medieval "two-name theorists" that they

> . . . are in general liable to say about some abstract logical rule, "Haec regula habet instantiam . . . in mysterio [Sanctissimae] Trinitatis . . .". [This rule has a counterexample in the mystery of the Most Holy Trinity] To my mind, the need to say such a thing about the rule simply shows that the rule is wrongly formulated and that we must try for an unexceptionable formulation of it.[3] (Geach 1972, 297)

And this excellent point applies not only to two-name theorists but to any philosopher or theologian who says of a supposed logical rule that, although it is valid, has a counterexample in the mystery of the Trinity—for logic applies to statements about everything whatever, and, if the Trinitarian God exists, anything that applies to statements about everything whatever applies to statements about him. This is not to say that logic is a superior force to which God must bow. It is rather to say that the requirement that it apply to statements about everything whatever—God included—is a superior force to which logic must bow. If one is convinced that the propositions 'The Father = God' and 'God = the Son' are true and the proposition 'The Father = the Son' is false, one should not conclude that the Principle of the Transitivity of Identity, although a valid rule of logical inference, does not apply to statements about the Persons of the Trinity (which fall outside the domain of statements to which the rules of logic apply). One should conclude, rather, that the Principle of the Transitivity of Identity is invalid.

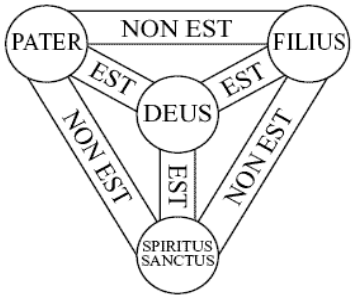
---

[2] The symmetry of identity raises an interesting question: If (1′) and (2′) are correct representations of, respectively, (1) and (2), why is it (at the very least) more natural to say "The Father is God" and "The Son is God" than it is to say, "God is the Father" and "God is the Son"?

[3] As far as I am able to determine, the label "the two-name theory" is Geach's invention and is not in general use. He defines it this way: it is the theory "that in an affirmative predication, the subject is a name and so is the predicate, and the predication is true if and only if the subject-name and the predicate-name stand for the same thing or things" and identifies Occam as one of its principal defenders.

By essentially the same token, one who accepts the doctrine of the Trinity cannot say, "All right, since the Principle of the Transitivity of Identity is logically valid, and since propositions (1′), (2′) and (3′) are essential to the doctrine of the Trinity, that doctrine is logically inconsistent. But so what? *Credo quia impossibile est*."[4] For to say that a statement is impossible is to say that it cannot possibly be true, and a statement that cannot possibly be true is a statement that *is* not true. And no one (I fervently hope) will say, "*Credo quia non verum est*."

Very well: if we are to accept the doctrine of the Trinity we must believe that its constituent propositions are true and are therefore logically consistent. If those propositions can be logically consistent only if the Principle of the Transitivity of Identity is invalid, might we Trinitarians not therefore simply reject the validity of that venerable principle? The famous diagram—if diagram is the word I want—called the Shield of the Trinity (*Scutum Trinitatis*) or the Shield of Faith (*Scutum Fidei*)



might, and with some plausibility, be thought to display graphically that the Principle of the Transitivity of Identity *habet instantiam in mysterio Sanctissimae Trinitatis*.) I will remind anyone who is seriously considering this option that the validity of the Principle of the Transitivity of Identity is a consequence of the validity of Leibniz's Law—roughly, the principle that if x is identical with y, then everything that is true of x is true of y and everything that is true of y is true of x. (Two rules of inference constitute the whole foundation of the logic of identity: Leibniz's Law and the Principle of the Reflexivity of Identity—that is, 'Everything is identical with itself'.) Suppose, for example, that God and the Father are identical and God and the Son are identical. Leibniz's Law tells us that if God and the Father are identical, then

[4] "I believe it because it is impossible"—a saying often attributed, but without textual support, to Tertullian. The Latin phrase in the next sentence but one means, "I believe it because it isn't true."

> Everything that is true of God is true of the Father.

And if God and the Son are identical, then

> One of the things that is true of God is that he is identical with the Son.

It follows that

> One of the things that is true of the Father is that he is identical with the Son.

That is to say:

> The Father is identical with the Son.

And, therefore, anyone who accepts (1′) and (2′) and (3′) and proposes to save the logical consistency of the doctrine of the Trinity by rejecting the Principle of the Transitivity of Identity must answer the following objection—an objection that is, to my mind, unanswerable:

> Look, if the Father is identical with God, then the Father and the God are *one*—one thing, one entity, one being, "one anything," so to speak. That is what *numerical* identity means. And *of course* that thing, entity, or being is identical with the Son only if it is identical with the Son. After all, if the Father and God are one, and if God and the Son are one, then the Father and the Son are one.

It therefore seems the doctrine of the Trinity must be rejected if it implies propositions (1′), (2′), and (3′). But obviously the doctrine does imply propositions (1), (2), and (3). And how shall we understand (1), (2), and (3) if not as (1′), (2′), and (3′)?

It was professor Geach who first proposed to answer (essentially) this question in terms of the idea that "identity is always relative to a sortal term."[5] I will state this idea in a form that I prefer to Geach's; his statement of the idea was in terms of the syntax of assertions of identity; my statement is rather in terms of the ontology that underlies assertions of identity. And my statement is:

[5] Geach and Anscombe (1963, 118–20); Geach (1977, 72–81).

> There is no such relation as identity.[6] That is, there is no relation *R* that is both *universally reflexive* (everything—everything whatever—bears *R* to itself) and *forces indiscernibility* (if a thing *x* and a thing *y* stand in *R*, then whatever is true of *x* is true of *y* and whatever is true of *y* is true of *x*). There are, rather, a *multiplicity* of identity-*like* relations, so to call them, relations expressed by open sentences of the form '*x* and *y* are the same N' where the dummy letter 'N' indicates the position of a sortal term.[7]

(In this essay, I will use the form of words '*x* and *y* are the same N' rather than the usual '*x* is the same N as *y*'.) This is one way of stating the doctrine or position or idea, or whatever the word I want is, that came to be called the doctrine of "the relativity of identity" or simply the doctrine of "relative identity." Having introduced the doctrine of relative identity, Geach went on to propose that the propositions of Trinitarian theology be formulated in terms of relative identities. He proposed that we say not "The Father is not identical with the Son" but rather "The Father is not the same person as the Son" (or, in my usage, 'The Father and the Son are not the same person'). And he proposed that we should also say, "The Father *is* the same God as the Son" (= 'The Father and the Son are the same God').

In this essay, I will depart in certain ways from Geach's terminology. I will not use the word 'God' as a sortal term. In my view, this word—'God' spelt with a capital 'G'—is not a count-noun and therefore is not a sortal term. Instead of 'The Father and the Son are the same God' I will say 'The Father and the Son are the same being'. For I shall take the noun 'God', that is, 'God' spelt with a capital 'G', to mean 'the divine being',[8] and I shall take 'The Father and the Son are the same divine being' to mean 'The Father is divine and the Son is divine and the Father and the Son are the same being'. And, finally, I shall take 'The Father, if he exists, is divine' and 'The Son, if he exists, is divine' to be analytical sentences—

[6] This sentence was chosen for dramatic effect. It would of course be technically true that there was no such relation as identity if there were two or more universally reflective relations that forced indiscernibility. (Suppose there were two such relations, *R1* and *R2*. That fact would have no consequences for first-order logic with identity: every sentence formed by flanking the identity sign with terms would have the same satisfaction conditions/truth-value on the interpretation 'The identity sign expresses *R1*' as it has on the interpretation 'The identity sign expresses *R2*'.)

[7] All sortal terms are count-nouns, at least in my usage. But Geach allows the 'N' position to be occupied by mass terms as well as count-nouns.

[8] I use the adjective 'divine' to express the divine essence or the essence of God. Spell it out as you will ('something a greater than which cannot be conceived' or 'necessarily existent being who possesses all perfections essentially' or . . .) or wisely decline to spell it out at all.

conceptual truths. (I will discuss the meanings of the sortals 'person' and 'being' presently.)

Following, then, the spirit, if not the precise letter, of Geach's proposal, let us explore the idea that (1), (2), and (3) be understood (respectively) as

*RI*1. God and the Father are the same being
*RI*2. God and the Son are the same being
*RI*3. The Father is a person and the Son is a person and the Father and the Son are not the same person.

Obviously, a contradiction cannot be derived from these three statements in the simple, straightforward way in which a contradiction can be derived from (1′), (2′), and (3′). But can a contradiction be derived from (*RI*1), (*RI*2), and (*RI*3) in some other, less straightforward way? If we are to answer this question, we must have a logic of relative identities. Many years ago I proposed such a logic, a logic I called RI-logic.[9] RI-logic is first-order logic (without identity, and with no terms but variables) supplemented as follows:

> To the vocabulary of first-order logic is added a special class of two-place predicate letters, *underlined* predicate letters—'$\underline{F}$', '$\underline{G}$' and so on.

In applications of RI-logic, underlined predicate letters may be replaced by and only by two-place predicates of natural languages of the forms '**1** and **2** are the same N' and '**2** and **1** are the same N' where 'N' represents a sortal term—that is, a count-noun (or a count-noun phrase like 'lump of clay') that *sorts*, a count-noun that is not, by definition, of unrestricted generality. (If we so use 'object' that everything, everything *whatever*, is an "object," then the count-noun 'object' does not "sort" and is thus not a sortal term.) Predicates of this kind we call *RI-predicates*. (In applications of RI-logic, we shall use the following abbreviation: for any variable α, and any sortal *N*, ⌜ α is a(n) *N* ⌝ abbreviates ⌜ α and α are the same *N* ⌝. So, for example, '*y* is a horse' abbreviates '*y* and *y* are the same horse'.[10])

The rules of inference of RI-logic are those of first-order logic supplemented by two additional rules:

[9] See: Inwagen (1988, 241–278).

[10] Or '*y* is a horse' might abbreviate '∃*x* *y* and *x* are the same horse'. '∃*x* *y* and *x* are the same horse' and '*y* and *y* are the same horse' (for example) are provably equivalent in RI-logic.

Symmetry

> For any variables α and β and any underlined predicate letter $\underline{F}$, if ⌜ $\underline{F}$ αβ ⌝ has occurred as a line in a derivation, then ⌜ $\underline{F}$ βα ⌝ may occur as a later line.

Transitivity

> For any variables α and β and γ and any underlined predicate letter $\underline{F}$, if ⌜ $\underline{F}$ αβ ⌝ and ⌜ $\underline{F}$ βγ ⌝ have occurred as lines in a derivation, then ⌜ $\underline{F}$ αγ ⌝ may occur as a later line.

How do the rules of RI-logic compare with the rules of first-order logic with identity? The latter are the rules of first-order logic supplemented by two rules that govern numerical identity:

Reflexivity

> For any variable α ⌜ α = α ⌝ may occur as a line at any point in a derivation.

Leibniz's Law

> For any variables α and β, and any sentences ***p*** and ***q***, if β does not occur in ***p*** and ***q*** is like ***p*** except for containing occurrences of β at some or all places at which ***p*** has free occurrences of α, then ⌜ α = β → (***p*** ↔ ***q***) ⌝ may occur as a line at any point in a derivation.

(For example, '$x = y \rightarrow$ (Fx ↔ Fy)' may occur as a line at any point in a derivation, since 'y' does not occur in 'Fx' and 'Fy' is like 'Fx' except for containing occurrences of 'y' at some or all places at which 'Fx' has free occurrences of 'x'. A more elaborate example: '$y = z \rightarrow$ ($\exists w$ G$yyxw$. ↔ $\exists w$ G$yzxw$)' may occur as a line at any point in a derivation, since 'z' does not occur in '$\exists w$ G$yyxw$' and '$\exists w$ G$yzxw$' is like '$\exists w$ G$yyxw$' except for containing occurrences of 'z' at some or all places at which '$\exists w$ G$yyxw$' has free occurrences of 'y'.)

Leibniz's Law is a very powerful rule—although, oddly enough, it does not suffice to prove the "trivial" sentences '$x = x$', '$z = z$', and so on (which of course is the reason why first-order logic with identity requires the rule Reflexivity in

addition to Leibniz's Law). Leibniz's Law and Reflexivity can, however, be used to demonstrate the validity of two "derived" rules:

Symmetry

> For any variables α and β, if ⌜ α = β ⌝ occurs as a line in a derivation, then ⌜ β = α ⌝ may occur as a later line.

Transitivity

> For any variables α, β, and γ, if ⌜ α = β ⌝ and ⌜ β = γ ⌝ occur as lines in a derivation, then ⌜ α = γ ⌝ may occur as a later line.

Notice that the rules of RI-logic include no rule that in any way corresponds to or is analogous to Leibniz's Law: they include no rule such as

> For any variables α and β and any underlined predicate letter $\underline{\text{F}}$, and any sentences ***p*** and ***q***, if β does not occur in ***p*** and ***q*** is like ***p*** except for containing occurrences of β at some or all places at which ***p*** has free occurrences of α, then ⌜ $\underline{\text{F}}$ αβ → (***p*** ↔ ***q***) ⌝ may occur as a line at any point in a derivation.

For if there were such a rule, then, for example,

> *x* and *y* are the same horse

would have the same logical properties as

> *x* is a horse and *y* is a horse and *x* = *y*,[11]

and RI-logic would be of no interest. It is of course open to anyone who employs RI-logic to assume that certain RI-predicates *dominate* various other predicates. To affirm that, e.g., '**1** and **2** are the same horse' dominates '**1** is quadrupedal' is to affirm that

[11] Or the same logical properties as '*x* is a horse and *x* = *y*' or '*y* is a horse and *x* = *y*': 'F*x* & *x* = *y*' and 'F*y* & *x* = *y*' and 'F*x* & F*y* & *x* = *y*' are logically equivalent.

> $\forall x \forall y$ (x and y are the same horse & x is quadrupedal **.** $\rightarrow$ y is quadrupedal)

is a conceptual truth. An RI-predicate can, moreover, be said to dominate other RI-predicates. For example, '**1** and **2** are the same horse' dominates '**1** and **2** are the same animal' if it is a conceptual truth that

> $\forall x \forall y$ (x and y are the same horse & x and x are the same animal **.** $\rightarrow$ x and y are the same animal)

—that is:

> $\forall x \forall y$ (x and y are the same horse & x is an animal **.** $\rightarrow$ x and y are the same animal).

If one is applying RI-logic to some philosophical problem, and if one believes that, e.g., 'same horse' dominates 'same animal' one may, of course, if one wishes, introduce the sentence '$\forall x \forall y$ (x and y are the same horse & x is an animal **.** $\rightarrow$ x and y are the same animal)' into one's reasoning as a premise. But that is what one must do: RI-logic does not *endorse* that sentence.

We may say further that an RI-predicate that dominates *all* predicates is *dominant*. (If, e.g., '**1** and **2** are the same horse' is dominant, 'x and y are the same horse' has the same logical properties as 'x is a horse & y is a horse & $x = y$'.) RI-logic will be of interest to those and only to those who believe that some RI-predicates are not dominant[12].

Finally, there is no rule that corresponds to the principle of the reflexivity of identity. If there were such a rule (it would have to look something like this:

> For any variable α and any underlined predicate letter $\underline{F}$, ⌜ $\underline{F}$ $\alpha\alpha$ ⌝ may occur as a line at any point in a derivation)

'$\forall x$ $\underline{F}xx$' would be a theorem of RI-logic—which would imply that

[12] In "And Yet They are Not Three Gods but One God," (1988) I considered a variant on the "relative identity approach" to the problem of the logical consistency of the doctrine of the Trinity according to which there *is* such a relation as identity *and* '**1** and **2** are the same being' and '**1** and **2** are the same person' fail to dominate certain predicates—and are thus not equivalent to, respectively, '**1** is a being & **2** is a being & **1 = 2**' and '**1** is a person & **2** is a person & **1 = 2**'. I shall not consider that option in the present paper.

$\forall x$ x and x are the same horse,

that is,

$\forall x$ x is a horse,

was an instance of a theorem of logic, a logical truth. (We should note, however, that

$\forall x\ (\exists y\ \underline{\mathrm{F}}xy\ \mathbf{.}\rightarrow \underline{\mathrm{F}}xx)$

is a theorem of RI-logic: If Bucephalus and something are the same horse, then Bucephalus and Bucephalus are the same horse; if the neutron star Kesteven 79 and something are the same horse, then Kesteven 79 and Kesteven 79 are the same horse; if the Riemann Zeta Function and something are the same horse, then the Riemann Zeta Function and the Riemann Zeta Function are the same horse . . . .)

3. The Sortals "Being" and "Person"

I use the count-noun 'being' to express the idea or concept that, in traditional metaphysics, was expressed by the words '*ousia*' and '*substantia*' and 'substance'. (And I take 'are the *same* being' to express the relation that in creedal statements is expressed by the formulas 'is consubstantial with' and 'is *homoousios* with'.) I believe—but this statement is bound to be controversial—that this is precisely the sense the word 'being' has in non-technical expressions like 'a human being', 'a rational being', 'an extraterrestrial being', 'a created being', and 'an omnipotent being'.

The word 'person', as every schoolchild knows, comes from the Latin word *persona*, and, as every schoolchild no doubt also knows, *persona* originally meant a mask of the sort worn by the actors in Greek and Roman dramas. *Persona* came to mean by extension a character in one of those dramas (a *dramatis persona*). Tertullian was responsible for the word's acquiring a technical meaning in theology: as a word for "what there are three of in the Trinity." Such a term no doubt answered to a theological need, although the use of *that* word to meet that need possibly lent unearned support to the heretical doctrine called modalism.

In the sequel, I will use the English word 'person' to represent indifferently the Latin word *persona* and the words of modern European languages other than English that are directly derived from *persona*. The theological technical term 'person', borrowed from everyday speech, at some point re-entered everyday speech with another meaning than 'mask'—whether in late Antiquity or in the Middle Ages I do not know—, but what that new meaning was is not entirely clear and is disputed. (Its meaning in theology is likewise not entirely clear and is likewise disputed, but the dispute about its meaning in theology and the dispute about its meaning in everyday speech are different disputes. And there is a third, albeit closely related, dispute: the dispute concerning the relation between its meaning in theology and its meaning in everyday speech.) It is my position that the word 'person' in Trinitarian theology—or at any rate in its Latin creedal summaries—means exactly what it means in ordinary speech. More exactly, I have a certain view about what 'person' means in everyday speech, and it is my position that it has *that* meaning in Trinitarian theology. If I were to be convinced that I had been mistaken about what 'person' meant in everyday speech, I would cease to affirm that 'person' in Trinitarian theology means what it means in everyday speech. But what *does* 'person' mean in everyday speech? This is, as I have said, a *quaestio disputata*. I will propose an answer—the answer I alluded to when I said I had an opinion about what the word meant in everyday speech.

It is clear enough that in the ordinary sense of the word human beings are "persons" (unless perhaps they are in some way radically cognitively defective: whether there could be a human being that was not a "person" is a question that belongs to the dispute about the meaning of 'person' in everyday speech), but it is doubtful whether the ordinary word 'person' *means* 'human being'. Rather, the ordinary meaning of 'person' seems to be something like 'rational being'—"individual substance of a rational nature," as Boethius puts it. Thus many things that are not human beings—angels, for example—are all persons in the ordinary sense of the word. Pagan gods, moreover, and Tolkien's elves and orcs and Wells's Martians and Selenites would be persons if they existed. But if I were asked to provide an explicit definition of 'person' (in its everyday sense) and I responded by saying 'A person is a rational being', I should no doubt quite properly be asked to say what I took the word 'rational' to mean, and that is a task I beg to be excused from. I will, accordingly, offer a definition of 'person' that does not contain the problematical word 'rational' but which is nevertheless, in my view, not entirely unrelated to 'A person is a rational being'. And my definition is: A person is a being whose nature allows it to be *addressed*. Or, to express the same idea rather more informally, a person is a *thou*—a being that is, as one might say, an "I" to

itself and a (potential) "thou" to such other beings as are capable of speech. But I must ask you, when you consider this definition, to take 'address' and 'thou' in their most literal, their most fundamental, their *central*, senses. For there are extended senses of those words in which practically anything can be addressed and in which practically everything is a potential "thou"—"London, thou art the flour of Cities all," "Wee, sleekit, cowrin, tim'rous beastie/O, what panic's in thy breastie!", "Devouring Time, blunt thou the lion's paw," and so on. The natures of cities and mice (much less that of Time) do not permit them to be addressed, not literally *addressed*, for their natures are incompatible with the possibility of their *responding* to a statement or a question or a command that (perhaps as a result of some operatic misapprehension on the part of the speaker) has been directed at them. And if a speaker knows that a certain being (or thing or object or item) cannot, by its very nature, respond to a proffered speech-act, that speaker cannot—not in the strictest and most literal sense—*address* that being. But the fact that speakers sometimes address—address in inverted commas, one might say—things they know cannot be *secundum litteram* addressed (and are therefore not, by my account, persons) is, as I once said, the exception that proves the rule, for we call that trope personification.

I contend that the meaning of 'person' in Trinitarian theology is exactly this: something is a person just in the case that it is a *thou*, just in the case that it can (the proper qualifications being made) be *addressed*. Whether or not this is the sense the word 'person' has in statements like, 'For there is one Person of the Father, another of the Son, another of the Holy Spirit,' it is evident that each of the persons of the Trinity *is* a person in this sense, for each can be addressed.

This fact, the fact that each of the persons of the Trinity can be addressed, suggests the following definition of 'x and y are the same person', a definition that does not imply, or at least does not obviously imply, that '**1** and **2** are the same person' is a dominant RI-predicate: x and y are the same person just in the case that x can be addressed and y can be addressed and to address x is to address y. Thus, the Father and the Holy Spirit are, by this definition, not the same person: to say, 'Forgive them, Father, for they know not what they do' is to address the Father and is not address the Holy Spirit; to say, '*Veni, Creator Spiritus*' is to address the Holy Spirit and is not to address the Father. I confess myself unable to suggest a definition or analysis of '**1** and **2** are the same being' that would *not* obviously imply that this predicate was dominant. But this inability is perhaps not surprising, for the consubstantiality of the persons of the Holy Trinity is a mystery. And perhaps this is as good a place as any for me to repeat my oft-made declaration that although the ideology (in Quine's sense of

'ideology') and logic of relative identity provide us with a way of stating the doctrine of the Holy Trinity from which no *formal* contradiction can be validly deduced, they do nothing what ever to enable us to see how that doctrine could possibly be true. The logic of relative identity does not explain how the Father and the Son, who are the same being, and who are both persons, could possibly not be the *same* person. The Father and the Son and the Spirit dwell in inaccessible light, and the logic of relative identity does not function as, nor was it intended to function as, a pair of magical spectacles that would enable one to see in that Light.

I will make only one assumption concerning the relation between the concepts "being" and "person": It is a conceptual truth that every person is a being.

4. The Elimination of Singular Terms: 'God' and 'the Father' and 'the Son'

We have asked whether the three statements (*RI*1), (*RI*2), and (*RI*3) are formally consistent. But RI-logic appears to be incapable even of addressing this question, since its language contains no terms but variables, and these three sentences contain (and the question of their mutual consistency depends on) three singular terms, 'God', 'the Father', and 'the Son'. It is, moreover, no accident that the language of RI-logic contains no singular terms, for (as the word 'singular' perhaps implies) there are singular terms only if there is such a relation as numerical identity. A singular term is a term that denotes exactly one object. That is, the following statement is a conceptual truth:

$$\forall x \forall y \forall z \ (x \text{ is a singular term } \& \ x \text{ denotes } y \ \& \ x \text{ denotes } z \ . \rightarrow y = z).$$

But this problem has a solution (of a kind), for it is possible to provide paraphrases of (*RI*1) and (*RI*2) and (*RI*3) that do not contain singular terms by a method that might be described as the relative-identity analogue of Russell's famous technique for eliminating definite descriptions by paraphrase. Suppose we have at our disposal the predicate '**1** is divine'[13]. (We so understand this predicate that 'x is divine $\rightarrow$ x is a person' is a conceptual truth.) We may then paraphrase 'God created the heavens and the earth' (or 'The unique divine being created the heavens and the earth') as

[13] See note 8.

> $\exists x$ (x is divine & $\forall y$ (y is divine $\rightarrow$ y and x are the same being) & x created the heavens and the earth).

And suppose we also have at our disposal the "Trinitarian" predicate '**1** begets **2**'. (We so understand this predicate that 'x begets y $\rightarrow$ **.** x is divine & y is divine & ~ x and y are the same person' is a conceptual truth.)

Now two definitions:

> x begets $=_{df}$ $\exists y$ x begets y (that is, $\exists y$ **1** begets **2** xy)
>
> x is begotten $=_{df}$ $\exists y$ y begets x (that is, $\exists y$ **1** begets **2** yx).

We may now paraphrase 'The Son was made flesh' (or 'The only-begotten divine person was made flesh') as

> $\exists x$ (x is begotten & $\forall y$ (y is begotten $\rightarrow$ y and x are the same person) & x was made flesh).

Using these two examples as models, we paraphrase (*RI*1), (*RI*2), and (*RI*3) as follows:

*RI*1′. $\exists x \exists y$ (x is divine & $\forall z$ (z is divine $\rightarrow$ z and x are the same being) & y begets & $\forall z$ (z begets $\rightarrow$ z and y are the same person) & x and y are the same being)

*RI*2′. $\exists x \exists y$ (x is divine & $\forall z$ (z is divine $\rightarrow$ z and x are the same being) & y is begotten & $\forall z$ (z is begotten $\rightarrow$ z and y are the same person) & x and y are the same being)

*RI*3′. $\exists x \exists y$ (x begets & $\forall z$ (z begets $\rightarrow$ z and x are the same person) & y is begotten & $\forall z$ (z is begotten $\rightarrow$ z and y are the same person) & ~ x and y are the same person).

It can be shown by an elementary model-theoretic argument that no deduction valid in RI-logic leads from these three sentences to a contradiction.[14] This result is not surprising; it is a consequence of the fact that the rules of RI-logic are very

[14] See (Inwagen 1988, 264 and 268).

weak (owning principally to the fact that they contain nothing corresponding to Leibniz's Law).

We now—finally—have the materials we shall need to analyze and evaluate Cartwright's argument.

5. Cartwright's Argument

The conclusion of Cartwright's argument is that even relative-identity formulations of the doctrine of the Trinity imply tritheism—and tritheism is not only heretical, not only inconsistent with the affirmations set out in the Apostle's Creed, but a rejection of the monotheism that is a common element of the three Abrahamic religions.

The argument opens with a statement of a premise that Cartwright says is a "trivial truth," namely, that every person is a being ((CA1)). Whether or not (CA1) is a trivial truth, I am willing to say that it is a truth—and, indeed, I did say this at the end of section 3. From (CA1) we infer the proposition that every divine person is a divine being ((CA2)). And, as any logic text will tell you, that inference is valid in first-order logic; and an inference valid in first-order is valid in RI-logic. We now introduce a second premise, a premise that, Cartwright contends, must be accepted by all Trinitarians, even those who formulate the doctrine of the Trinity in terms of relative identities. And this second premise ((CA3)) is 'There are at least three divine persons'. And, finally, from (CA2) and (CA3) we deduce the conclusion of the argument, 'There are at least three divine beings'. That this proposition follows from (CA2) and (CA3) is, Cartwright contends, an obvious consequence of the principle ("Cartwright's Principle"):

> If every A is a B, then there cannot be fewer B's than A's,

This principle, Cartwright maintains, is "evident to the natural light of reason" (p. 196). And it is certainly indisputable that Cartwright's Principle does have the consequence that if every divine person is a divine being, and if there are at least three divine persons, then there are at least three divine beings.

There are a number of questions that could be raised in connection with this argument. Let us begin with this one: *Is* Cartwright's Principle evident to the natural light of reason? Well, it is certainly mathematically demonstrable (and therefore, I suppose, evident to the natural light of reason) that each of the sentences in the infinite sequence

> If every A is a B, and if there is at least one A, then there is at least one B
> If every A is a B, and if there are at least two A's, then there are at least two B's
> If every A is a B, and if there are at least three A's, then there are at least three B's
>
> .
> .
> .

is—or at least corresponds to—a theorem of first-order logic with identity. Consider, for example, the second sentence in the sequence. This sentence "corresponds to" the sentence

$$\forall x\,(\mathrm{A}x \rightarrow \mathrm{B}x)\ \&\ \exists x \exists y\,(\mathrm{A}x\ \&\ \mathrm{A}y\ \&\ x \neq y) \rightarrow\ \exists x \exists y\,(\mathrm{B}x\ \&\ \mathrm{B}y\ \&\ x \neq y),$$

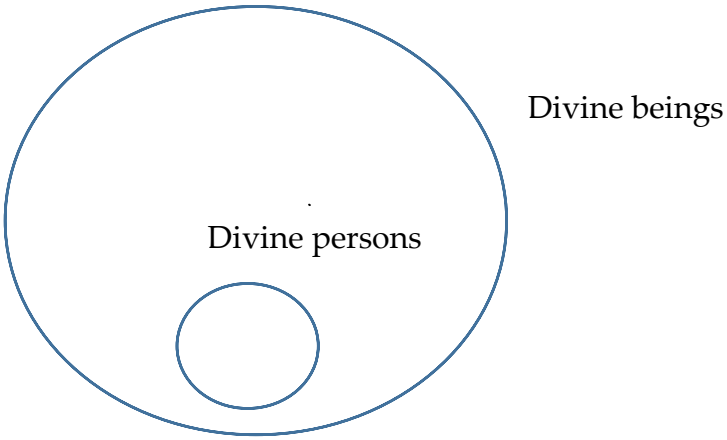
and this sentence is a theorem of first-order logic with identity—and an instance of the following theorem of first-order logic *simpliciter*

$$\forall x\,(\mathrm{A}x \rightarrow \mathrm{B}x)\ \&\ \exists x \exists y\,(\mathrm{A}x\ \&\ \mathrm{A}y\ \&\ \mathrm{F}xy) \rightarrow\ \exists x \exists y\,(\mathrm{B}x\ \&\ \mathrm{B}y\ \&\ \mathrm{F}xy)$$

provided that the predicate '**1 = 2**' is an allowable substituend for the two-place predicate-letter 'F^2'. These statements are uncontroversial statements of logical fact. But, of course, no one who holds that there is no such relation as identity will grant that they establish the validity of Cartwright's Principle, or, indeed, that they establish the validity of anything whatever. I can see no better ground for supposing that Cartwright's Principle is evident to the natural light of reason other than that every sentence in the above infinite sequence is a theorem of first-order logic with identity.[15]

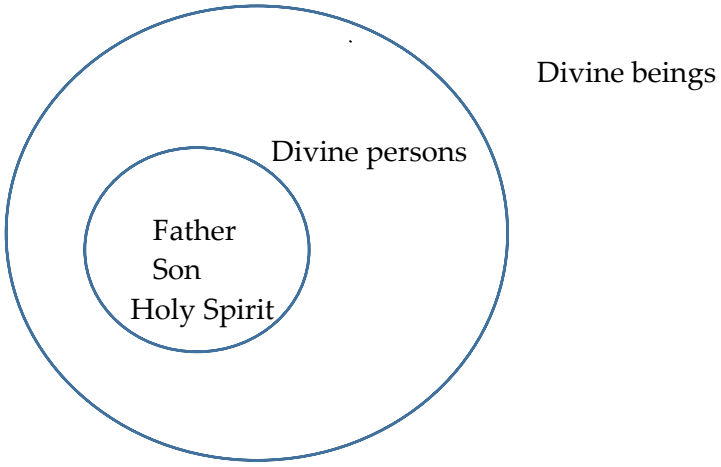
A reader for *TheoLogica* has suggested that a simple Venn diagram makes it clear why someone would suppose that Cartwright's Principle was evident to the natural light of reason. If all divine persons are divine beings then a Venn diagram that represents this relation will look like this

[15] On a very natural reading, Cartwright's Principle is an informal statement of a theorem of set theory. But I do not think that this is a better reason for supposing that it is evident to the natural light of reason than is the fact that every sentence in the above sequence is a theorem of first-order logic with identity. (Note, moreover, that the underlying logic of a first-order set theory—such as Zermelo-Fraenkel set theory—is first-order logic with identity.)



And, since it will be evident to the natural light of reason that anything within the *Divine persons* region will "automatically" be within the *Divine Beings* region, it will also be evident to the natural light of reason that there cannot be more divine persons than divine beings. And yet, according to the doctrine of the Holy Trinity, there are three divine persons and there is only one divine being.

Now since there is no divine being who is not a divine person, the filled in Venn diagram should look like this



And the argument will be:

> There are three items within the *Divine persons* region; the *Divine beings* region lies wholly within the *Divine beings* region. Therefore, there must be at least three items in the *Divine beings* region—the three in the *Divine persons* region if no others.

But suppose that there is no such relation as numerical identity—and thus no answer to questions like 'Is the Son identical with the Holy Spirit?' And suppose the following rather complex statement is true. (In this statement, '$\underline{\text{B}}$' abbreviates 'is the same being as' and '$\underline{\text{P}}$' abbreviates 'is the same person as'.)

> $\exists x \exists y \exists z$ (Dx & Dy & Dz & ~$\underline{\text{P}}xy$ & ~$\underline{\text{P}}yz$ & ~$\underline{\text{P}}xz$ & $\underline{\text{B}}xy$ & $\underline{\text{B}}yz$ & $\underline{\text{B}}xz$ & $\forall w$ (D$w \rightarrow$. $\underline{\text{B}}wx$ & $\underline{\text{B}}wy$ & $\underline{\text{B}}wz$ & ($\underline{\text{P}}wx \vee \underline{\text{P}}wy \vee \underline{\text{P}}wz$))).

In what sense, then, are there three "items" in the *Divine persons* region (and hence at least three "items" in the *Divine beings* region)? It cannot be for *this* reason:

> The Father, the Son, and the Holy Spirit are in the *Divine persons* region & the Father ≠ the Son, & the Son ≠ the Holy Spirit, & the Holy Spirit ≠ the Father, & $\forall x$ (x is in the *Divine persons* region $\rightarrow$. x = the Father $\vee$ x = the Son $\vee$ x = the Holy Spirit).

It is true that within the *Divine persons* region there are a being x, a being y, and a being z such that x and y are not the same person, and y and z are not the same person, and x and z are not the same person—but x and y are the same being, and y and z are the same being, and x and z are the same being. Are there then three "items" in the *Divine persons* region, or is there only one? This question has no answer. But there is more to be said about the Persons of the Trinity and counting, and we shall come back to this topic.

If Cartwright's Principle is not (from the point of view of the friends of the relativity of identity) evident to the natural light of reason, might it nevertheless (even from their point of view) be *true*? Have they any reason to suppose that it is not only not evident to the natural light of reason but *false*? In my view, there is no need for "Relative Trinitarians" (so to call those who accept a statement of the doctrine of the Holy Trinity in terms of relative identities) to address these questions. In my view, Relative Trinitarians should rather turn their attention to the second premise of the argument, (CA3)—that is, 'There are at least three divine persons'. And they should ask *this* question:

> *Does* the doctrine of the Trinity imply that there are at least three divine persons?

From this point on, I will, simply as a matter of expositional convenience, pretend

to be a Relative Trinitarian. (I say "pretend" because I am by no means a convinced Relative Trinitarian.) This pretense will enable me to avoid having repeatedly to insert the parenthesis "(from the Relative Trinitarian point of view)" into my argument.

My answer to this question is No. (A firm but qualified No. I shall in due course state the needed qualification.) My answer is No because, I maintain, 'There are at least three divine persons' is meaningless. When I say that the sentence 'There are at least three divine persons' is meaningless, I am not of course saying that it is like 'Gubble gubble gubble'—or even very much like 'Quadruplicity drinks procrastination' and '*Das Nichts nichtet*'. I mean only that neither that sentence nor its negation is true—that is, that neither that sentence nor its negation expresses a true proposition.

To see why I say that 'There are at least three divine persons' is neither true nor false, let us consider the general case. If one or the other of two sentences of the forms

There are at least *n* so-and-sos

It is not the case that there at least *n* so-and-sos

is true, then there must be an answer to the question

Are there at least *n* so-and-sos?

And if there is no such relation as identity—if there are only sortal-relative relations expressed by sentences of the form '*x* and *y* are the same N'—it is not guaranteed that there will be an answer to this question. Consider for example the question

Q. Are there at least two clay statues?

Suppose that there is no such relation as identity, but that there are many sortal-relative identity-like relations, among them the relations "are the same statue" and "are the same lump of clay." We define '*x* is a statue' as '*x* and *x* are the same statue'[16]. We define '*x* is a lump of clay' as '*x* and *x* are the same lump of clay'. We define '*x* is a clay statue' as '*x* is a statue & *x* is a lump of clay'—and thus

[16] Or '$\exists y$ *x* and *y* are the same statue'. (See note 10.)

SL1. $\forall x$ (x is a clay statue $\rightarrow$ x is a lump of clay).

is true by definition. And suppose that the two theses

SL2. $\forall x \forall y$ (x is a lump of clay & y is a lump of clay. $\rightarrow$ x and y are the same lump of clay)

SL3. $\exists x \exists y$ (x is a clay statue & y is a clay statue & ~ x and y are the same statue)

are also true. That is, I ask you to suppose that, to speak with the vulgar, there is only one lump of clay and that there are (timelessly speaking) at least two clay statues—for someone has at one time shaped that lump into a clay statue—and has later squeezed the lump into a shapeless mass and then formed it into a different statue. (No formal contradiction can be derived from (SL1), (SL2), and (SL3) in RI-Logic.) (SL1), we have said, is true by definition. If (SL2) and (SL3) are also true, then (given our assumptions about identity) the question Q has no answer. If (SL2) and (SL3) are true, we can say no more than the following in response to Q:

> Clay statues fall under at least two sortals, "statue" and "lump of clay." Counting clay statues by the sortal "statue," there are at least two clay statues; counting clay statues by the sortal "lump of clay," there is exactly one clay statue.

And we can say no more than the following in response to the question 'Are there at least two lumps of clay?'

> Lumps of clay fall under at least two sortals, "statue" and "lump of clay." Counting lumps of clay by the sortal "statue," there are at least two lumps of clay; counting lumps of clay by the sortal "lump of clay," there is exactly one lump of clay.

The sentence 'Counting clay statues by the sortal "lump of clay," there is exactly one clay statue' may be understood as follows:

> $\exists x$ (x is a clay statue) & $\forall x \forall y$ (x is a clay statue & y is a clay statue. $\rightarrow$ x and y are the same lump of clay).

And the sentence 'Counting clay statues by the sortal "statue," there are at least

two clay statues' may be understood as follows:

> $\exists x$ (x is a clay statue) & $\sim\forall x\forall y$ (x is a clay statue & y is a clay statue. $\rightarrow$ x and y are the same statue).

(The latter sentence is equivalent to

> $\exists x\exists y$ (x is a clay statue & y is a clay statue & $\sim x$ and y are the same statue).)

If, therefore, someone asks (in one of those lump-of-clay-impoverished possible worlds in which (SL2) and (SL3) are true), "How many clay statues are there?", there are two equally good answers: 'One, counting them by lumps of clay' and 'At least two, counting them by statues'. That is to say, these two answers are equally good from the point of view of logic and semantics. They are not, however, equally good from the point of view of pragmatics, for someone who asked that question—someone who asked a question framed in *those* words—would (surely?) find the "counting by statues" answer more interesting, find it a more relevant response to the question asked, than the "counting by lumps of clay" answer. In a certain sense, therefore, the answer 'At least two' may be regarded as the "correct" answer to the question, 'How many clay statues are there?' And, for the same reason, 'Exactly one' may—again, in a certain sense—be regarded as the correct answer to the question, 'How many lumps of clay are there?' (Of course, the question 'How many clay statues are there?' does have a semantically determinate answer if, for every sortal N under which clay statues fall, the RI-predicate ⌜ **1** and **2** are the same N ⌝ is dominant.[17])

The case is exactly parallel as regards divine persons and divine beings. For suppose that these three theses are true:

> $\forall x$ ($\underline{\mathrm{P}}xx \rightarrow \underline{\mathrm{B}}xx$)
>
> $\forall x\forall y$ (Dx & Dy **.** $\rightarrow$ $\underline{\mathrm{B}}xy$)
>
> $\exists x\exists y\exists z$ (Dx & Dy & Dz & $\sim\underline{\mathrm{P}}xy$ & $\sim\underline{\mathrm{P}}yz$ & $\sim\underline{\mathrm{P}}xz$ & $\underline{\mathrm{B}}xy$ & $\underline{\mathrm{B}}yz$ & $\underline{\mathrm{B}}xz$ & $\forall w$ (D$w \rightarrow$ **.** $\underline{\mathrm{B}}wx$ & $\underline{\mathrm{B}}wy$ & $\underline{\mathrm{B}}wz$ & ($\underline{\mathrm{P}}wx \vee \underline{\mathrm{P}}wy \vee \underline{\mathrm{P}}wz$))).

[17] Assuming that there are no borderline cases of "statues"—for example, a lump of clay that has been produced by deforming a clay statue just to the extent that it is indeterminate whether that lump is now a statue.

(No formal contradiction can be derived from these three theses in RI-Logic.) Then the question 'How many divine beings are there?' has no semantically determinate answer. Or put the matter this way: if *identity* is always relative to a sortal term, then *counting* is always relative to a sortal term; the doctrine of the relativity of identity implies the doctrine of the relativity of numbering. If I, a Relative Trinitarian, am asked the question, 'How many divine beings are there?' I can (from a purely semantical point of view—leaving all pragmatic considerations aside) say no more than

> Divine beings fall under at least two sortals, "person" and "being." Counting divine beings by the sortal "person," there are exactly three divine beings; counting divine beings by the sortal "being," there is exactly one divine being.

And similarly for 'How many divine persons are there?':

> Divine persons fall under at least two sortals, "person" and "being." Counting divine persons by the sortal "person," there are at exactly three divine persons; counting divine persons by the sortal "being," there is exactly one divine person.

But, again, we may say that if someone asks me "How many divine persons are there?", it makes pragmatic sense for me to answer, "Exactly three." (This is the qualification of the statement 'The answer to the question, "Does the doctrine of the Trinity imply that there are at least three divine persons?" is No' that I promised.) And, of course, this is also true, mutatis mutandis, for the question, 'How many divine beings are there?' and the answer 'Exactly one'.

Relative Trinitarianism therefore implies that the proposition that, according to Cartwright's second premise, the Relative Trinitarian must accept—that there are at least three divine persons—is meaningless. (Or, in Wolfgang Pauli's words, not even wrong.)

Interestingly enough, Relative Trinitarianism countenances Cartwright's Principle, or at least countenances its sortal-relative analogue—something like

> For every plural sortal *Ns*, ⌜ if every A is a B, then, counting A's and B's by *Ns*, there cannot be fewer B's than A's ⌝ is true.

The two conditional statements

> If every divine person is a divine being, then, counting divine persons and divine beings by persons, there cannot be fewer divine beings than divine persons

and

> If every divine person is a divine being, then, counting divine persons and divine beings by beings, there cannot be fewer divine beings than divine persons

are both consequences of this principle and Relative Trinitarianism. But the Relative Trinitarian will happily accept the consequents of both conditionals.

I am no great friend of the practice of labeling philosophical arguments question-begging. But I think it is not entirely inappropriate to apply this dyslogistic label to Cartwright's argument that, because there are three divine persons, even Relative Trinitarians must accept the proposition that there are at least three divine beings—for Relative Trinitarianism implies that (despite the fact that there is a place, even an essential place, for the sentences 'There are exactly three divine persons' and 'There is exactly one divine being' in our theological discourse) there is no such thing as the number of divine persons and there is no such thing as the number of divine beings.

Bibliography

Cartwright, Richard. 1987. "On the Logical Problem of the Trinity," in Richard Cartwright, *Philosophical Essays*, 187–200. Cambridge, Mass.: MIT Press.

Geach P. T. and G.E.M Anscombe, 1963. *Three Philosophers*. Oxford: Blackwell.

Geach, P. T. 1972. *Logic Matters*. Oxford: Basil Blackwell.

Geach, P. T. 1977. *The Virtues*. Cambridge: Cambridge University Press.

Inwagen, Peter van. 1988. "And Yet They Are Not Three Gods But One God." In *Philosophy and the Christian Faith*, edited by Thomas V. Morris. South Bend, IN: The University of Notre Dame Press, 1988.

Published Online: April 26, 2022

ERRATUM

The offset paragraph at the bottom of page 129 should read:

There are three items within the Divine persons region; the Divine persons region lies wholly within the Divine beings region. Therefore, there must be at least three items in the Divine beings region—the three in the Divine persons region if no others.