Logique & Analyse **143-144** (1993), 195-218

THE SLINGSHOT ARGUMENT

Stephen READ

There is an argument which has been used to devastating effect in several areas of philosophy in the past fifty years. It was originally presented by Church in his review of Carnap's *Introduction to Semantics*.[1] The argument was put forward by Church in support of Frege's claim that sentences designate truth-values and against Carnap's view that they designate propositions. For this reason it was for a time called "Frege's argument", but in fact, as we will see, it goes far beyond any argument Frege himself proposed. Latterly, the name has been dropped and it has come to be known as the Slingshot Argument. David beat Goliath with a single stone from his sling; and this argument truly appears to be a giant-killer. Many writers have used it believing it to be a sound refutation of the coherence of intensional notions. Many others have rejected it as unsound.[2] My position is that it constitutes a fallacy and a paradox. Its applications have certainly been unsound. But the source of its invalidity is revealing. If we take complex singular terms as contextually defined in Russell's manner, the principle of substitutivity of identicals (SI) cannot be applied directly to them. Doing so is what leads to the fallacy. On the other hand, if we take them as Fregean and treat them as primitive and *de re*, that is, as rigid designators, the logical equivalence which lies at the heart of the argument must fail to hold, and so we are not entitled to apply the substitutivity of logical equivalents (SE). The same is true if one uses a contextual definition equivalent to Frege's treatment. Finally, if they are treated as primitive but *de dicto* (that is, as flexible designators) the step of (SI) again fails. In other words, either the premises are true and the argument (taking the terms one way or the

[1] Church [2], Carnap [1]. The argument is attributed by Wallace [1] to Gödel, citing his paper, "Russell's mathematical logic", Gödel [1]. However, according to Parsons, Gödel submitted his paper for the Schilpp volume only on 17 May 1943 (see Gödel [2: p.102]), whereas Church's review of Carnap had already been published in *Philosophical Review* for May 1943. At the end of his paper, Gödel thanks Church for assistance with English expression.

[2] See Barwise and Perry [1, 2]; Cummins and Gottlieb [1]; Davies [1: pp. 210-1]; Hochberg [1]; Lycan [1]; Mackie [1]; Morton [1]; Neale [2, 3]; Sainsbury [1]; Sharvy [1]; Taylor [1]; and Wallace [1].

other) is invalid; or, taking the terms in another way, the argument is indeed valid, but one of the premises is false.

1. *A paradox.* The argument, as all good and philosophically interesting arguments, appears in a number of forms. Some may dispute whether these forms really have a unity. But to see its paradoxical nature, we need to see that the following is the core idea: it appears to show that any two sentences with the same truth-value are logically equivalent.

Take, without loss of generality, any two contingently true sentences, p and q. (We will see that the argument can be adapted without difficulty to the case where p and q are both false.) The argument works by constructing singular terms from p and q. These singular terms can be definite descriptions, or class abstracts, or many other sorts of term. A typical example will be the singular term $\iota x(x = a. p)$, that is (dot is conjunction), that object which is a provided p— that is, provided p is true.[3] Then it is plausible that

(1.1) $\iota x(x = a. p) = a \leftrightarrow p$

holds, where '$\leftrightarrow$' is shorthand for *logical* equivalence. That is, it is logically true that if p is true, then the object which is a provided p holds is a, since p holds —it's a if p holds and p does hold; and conversely, if the object which is a provided p holds is a, then indeed p holds— since if p didn't hold, it wouldn't be a, but it is. For the same reason, we have

((1.2) $\iota x(x = a. q) = a \leftrightarrow q$

Since p and q are both true, we can infer from (1.1) and (1.2) that

(1.3) $\iota x(x = a. p) = a$

and

(1.4) $\iota x(x = a. q) = a$

Hence

[3] This is the term used in Gödel [1]; cf. Føllesdal [1: p. 266]. Other terms used are $\hat{x}(x = x. p)$: Quine [5], Davidson [1, 2]; $\hat{x}(Gx. p)$: Anscombe [1]; $\delta p = \iota x(x = 1. p \vee x = 0. \neg p)$ Quine [4: p. 148]; and 'the x such that x is identical with Diogenes and p': Davidson [3].

(1.5) $\iota x(x = a.\, p) = \iota x(x = a.\, q)$

We now apply the substitutivity of identicals (SI) to (1.1), replacing the description term '$\iota x(x = a.\, p)$' by '$\iota x(x = a.\, q)$', for (1.5) says they are identical. We obtain

(1.6) $\iota x(x = a.\, q) = a \leftrightarrow p$

Finally, the substitutivity of equivalents (SE), or simply the transitivity of logical equivalence, applied to (1.2) and (1.6) yields

(1.7) $p \leftrightarrow q$.

Thus, simply on the assumption that p and q are (contingently) true, we seem to have shown that they are logically equivalent. If p and q are both false, we use the singular terms $\iota x(x = a.\, \neg p)$ and $\iota x(x = a.\, \neg q)$ in an analogous way, to show $\neg p \leftrightarrow \neg q$ and so, by contraposition, $p \leftrightarrow q$.[4] Thus any materially equivalent contingent propositions have been shown to be logically equivalent.[5]

This is a paradox. Of course, if '$\leftrightarrow$' merely means material equivalence, there is no problem. But the assumptions made about it, in particular in (1.1) and (1.2) and in the move from (1.2) and (1.6) to (1.7) seem unobjectionable principles for much stronger notions, for example, strict implication, relevant implication, or even entailment. Either we have a remarkable and counterintuitive result, namely, that there is no such strong notion of logical equivalence; or the argument is not sound, that is, either there is a fallacy in the argument, or one of the premises is false. We will see that this is indeed the case.

2. *Church's argument.* As noted, Church's original argument had a different conclusion. It was directed at a very specific point, namely, to show that the designata of sentences are truth-values, not propositions (or thoughts), that is, to defend Frege's original scheme of sense and reference against the changes which Carnap had wrought in it. In particular, he

[4] Of course, (1.7) does not follow from $\neg p \leftrightarrow \neg q$ intuitionistically. However, the converse does hold. So even intuitionistically, we have the paradoxical result that the contradictories of arbitrary materially equivalent propositions are logically equivalent.

[5] The only point in taking p and q to be contingent is that arguably non-contingent propositions cannot be only materially equivalent, and so there would be no puzzle.

objected to Carnap's treatment of the designata of sentences as senses (that is, thoughts or propositions).

Church's counterargument starts by rehearsing (and adapting to Carnapian terminology) Frege's famous argument from "On Sense and Reference" for distinguishing sense from reference. Let T_1 be 'The morning star is the evening star' and T_2, 'The morning star is the morning star'. Then 'T_1' designates 'The morning star is the evening star' and 'T_2' designates 'The morning star is the morning star', for short, Des(T_1, MS=ES) and Des(T_2, MS=MS). Church observes that Carnap accepts an extensionality principle for 'Des', indeed, for the whole metalanguage containing 'Des', so that "synonymous" terms may be intersubstituted in ϕ in the context Des(x, ϕ) —where 'synonymous' for Carnap simply means 'co-designative'. Since MS=ES, we can replace 'ES' by 'MS' in Des(T_1, MS=ES) to obtain Des (T_1, MS=MS). Thus T_1 and T_2 designate the same thing, i.e. Syn(T_1,T_2). (See Fig. I)

I think it is clear that Frege would have accepted this version of his reasoning, given that Syn(x,y) is to be interpreted (as surely was Carnap's intention) as "x and y do not differ in mode of designation."[6] That is, Des(x, ϕ) means that the name which replaces 'x' designates an expression whose designation is the same as that of the expression which replaces 'ϕ'.[7] What Frege, and Church, take it to show is that mode of designation does not respect subject matter, that is, that there must be another element connected with a sign besides its designation, namely, its sense. Syn(T_1, T_2) is true, for all it means is that T_1 and T_2 have the same designation, namely, the True. Sentences designate truth-values, not senses. But how can we show that?

Church's idea is to use the type of singular term we saw in §1 in place of 'ES' in Frege's argument, constructing two sentences S_1 and S_2 for which, mutatis mutandis, he can show Syn(S_1, S_2). He then extends the argument to show that for any true sentence p, Syn(p, S_2), so that for any pair of true sentences p and q, Syn(p, S_2) and Syn(q, S_2), whence Syn(p, q) —since clearly 'Syn' is symmetric and transitive. Thus any two sentences which have in common only that they are both true, designate the same thing. So too for any two false sentences, and so we are forced to accept that "no possibility remains for the designation of sentences except that they be truth-values."[8]

[6] Cf. Frege [1: p. 57].

[7] Cf. Carnap [1: p. 50].

[8] Church [2: p. 300].

Let us rehearse the argument formally. Take any two true sentences, p and q. Church writes 'Λ' for the empty set, and '$(\lambda x)\phi$' for the set of x such that ϕ. Consider the set-abstract $(\lambda x)(x = x.\neg p)$.[9] Since p is true, this set-abstract designates the empty set. Let S_1 be the sentence '$(\lambda x)(x = x.\neg p) = \Lambda$' and S_2 the sentence '$\Lambda = \Lambda$'. That is,

$$(2.1)\ \text{Des}\big(S_1, (\lambda x)(x = x.\neg p) = \Lambda\big)$$

and

$$(2.2)\ \text{Des}\big(S_2, \Lambda = \Lambda\big)$$

and as we noted, since p is true,

$$(2.3)\ (\lambda x)(x = x.\neg p) = \Lambda$$

Then by Carnap's principle of intersubstitutivity, from (2.1) and (2.3), we obtain

$$(2.4)\ \text{Des}\big(S_1, \Lambda = \Lambda\big)$$

whence by definition, from (2.2) and (2.4),

$$(2.5)\ \text{Syn}\big(S_1, S_2\big)$$

So far we have simply copied Frege's reasoning into Carnap's notation. But since S_1 is true only because p is (and vice versa)

$$(2.6)\ \text{Syn}\big(p, S_1\big)$$

and so, from (2.5) and (2.6),

$$(2.7)\ \text{Syn}\big(p, S_2\big)$$

But S_2 contains no occurrence of p, and so we can show, for any other true sentence q, in exactly the same way,

$$(2.8)\ \text{Syn}\big(q, S_2\big)$$

[9] If we allow vacuous binding of variables, we could use '$\lambda x.\neg p$'. See, e.g., Quine's use of $\hat{a}.\,p$ in Quine [3: p. 157].

whence

(2.9) $\mathrm{Syn}(p, q)$

as required. All true sentences have the same designation. (See Fig. I)

Frege's argument	*Church's argument*
	Take *p*, *q* both true:
Des (T_1, The MS is the ES)	(2.1) $\mathrm{Des}\left(S_1, \lambda x(x = x. \neg p) = \Lambda\right)$
Des (T_2, The MS is the MS)	(2.2) $\mathrm{Des}\left(S_2, \Lambda = \Lambda\right)$
the MS = the ES	(2.3) $\lambda x(x = x. \neg p) = \Lambda$ since *p* holds
Des (T_1, The MS is the MS)	(2.4) $\mathrm{Des}\left(S_1, \Lambda = \Lambda\right)$ by SI, from (2.1) and (2.3)
Syn (T_1, T_2)	(2.5) Syn $\left(S_1, S_2\right)$
	(2.6) Syn (p, S_1) since p and S_1 are logically equivalent
	(2.7) Syn (p, S_2) by SE, from (2.5) and (2.6)
	(2.8) Syn (q, S_2) since S_2 is '*p*'-free
	(2.9) Syn (p, q) from (2.7) and (2.8)

Fig. I

3. *Quine's version.* So far we have seen the argument presented as a paradox, that any pair of propositions with the same truth-value are logically equivalent; and in support of Frege's thesis, that sentences designate truth-values. But its most common employment has been as an argument in favour of extensionalism, as showing that intensional notions are ill-founded. It purports to show that contexts which might appear to be intensional are not really so. The only alternatives, it is claimed, are full extensionality (truth-functionality) or the complete opacity of quotation.

Consider a sentential context, $\Phi(p)$, about which we make two assumptions. First, we assume that logically equivalent sentences may be intersubstituted in Φ, that is,

$$\frac{\Phi(p) \quad p \leftrightarrow q}{\Phi(q)}$$

(Following Corcoran and Herring [1], I am calling this principle, 'SE'.) Secondly, we assume that $\Phi(p)$ is extensional in singular terms occurring in p, that is,

$$\frac{\Phi(p) \quad s=t}{\Phi(p')}$$

where p and p' result from an open wff q with one free variable x by replacing x by respectively s and t. (This is the substitutivity of identicals, SI.) Now suppose $\Phi(p)$ and $p \equiv q$, that is, p and q are materially equivalent (have the same truth-value). Suppose w.l.g. that p and q are both true. Take one of the now familiar singular terms constructed from p, say, $\iota x(x=a.\, p)$. (We use $\iota x(x=a.\neg p)$ if p and q are both false.) Then as before

(3.1) $\iota x(x=a.\, p)=a \leftrightarrow p$

and

(3.2) $\iota x(x=a.\, q)=a \leftrightarrow q$

and so, as in (1.1) to (1.5),

(3.3) $\iota x(x=a.\, p)=\iota x(x=a.\, q)$

Hence from

(3.4) $\Phi(p)$

we can infer

(3.5) $\Phi(\iota x(x=a.\, p)=a)$

by (SE) using (3.1), and so

(3.6) $\Phi(\iota x(x=a.\, q)=a)$

by (SI) using (3.3), whence

(3.7) $\Phi(q)$

by (SE) again, this time using (3.2). That is, although (SE) suggests we need logical equivalence of p and q to proceed from $\Phi(p)$ to $\Phi(q)$, use of the second assumption, i.e. (SI), on a term like $\iota x(x=a.\, p)$ leads to the intersubstitutivity of material equivalents.

This is paradoxical. Suppose $\Phi(p)$ is 'r is a logical consequence of p', for short, '$r \in Cn(p)$', and suppose p is materially equivalent to q, but does not entail q. Then $p \in Cn(p)$, but $q \notin Cn(p)$. (q is, say, 'p and the earth is flat'.) But logical consequence is closed under entailment, so '$r \in Cn(p)$' satisfies (SI) and (SE) for both 'p' and 'r'. The above argument then takes any form $\Phi(p)$ to $\Phi(q)$ and back again. So p and q are logically equivalent: $q \in Cn(p)$ and $p \in Cn(q)$.

Quine, Davidson and others have all used the slingshot argument to conclude that there are no really intensional contexts. There are opaque ones, that is, contexts which are entirely "opaque" to substitution, quantification and so on, such as quotation. Here (SI) fails. But there can be no contexts, they infer, to which both (SI) and (SE) apply and which are not purely extensional, that is, which don't permit the substitution of mere material equivalents. Contexts are either completely transparent, or completely opaque. This cannot be true.

4. *Russellian contextual definition.* To see the error in these arguments, we need to think carefully about the role of complex singular terms in our logic. The terms which play a crucial role in the arguments have come to be called "variable bound terms" (see Hatcher [1: pp. 65-6]; also in the revised edition, Hatcher [2: pp. 59-61]). There are two sorts of complex singular terms: function terms and variable bound terms (vbts, for short). The former are familiar; the latter, somewhat surprisingly, less so. For in everyday, mathematical and non-mathematical, language they are very common.[10] In informal mathematics, we commonly speak of the set of objects satisfying some condition; in ordinary speech, we talk about "the man who robbed the bank at Monte Carlo" or "a typical day at the races". These terms are formed from an open formula by binding its free variable with an operator —a "variable-binding term operator" (vbto) in Hatcher's terminology. '^', 'ι' and 'ϵ' all form complex terms from open formulae, forming respectively a set abstract, $\hat{x}Fx$ (the set of x such that Fx, otherwise written $\lambda x. Fx$ or $\{x|Fx\}$), a definite description, ιxFx (the object x such that Fx) and an indefinite description, εxFx (Hilbert's epsilon symbol, an x such that Fx).[11]

Most commonly in formal logic and mathematics, such terms are ignored and make no appearance. Most often, when they do, they are contextually defined, that is, they are treated as abbreviatory definitions, to be elimi-

[10] See, e.g., Rosser [1: ch. VIII, esp. p. 195].

[11] Quine's 'δ' mentioned in fn. 3 above is also a vbto.

nated. Thereby, there is no need to provide either proof rules or semantics for them. Any occurrence of a vbt, νxFx, is eliminated by definition, replacing $G(\nu xFx)$ by some formula in which 'ν' does not occur. The usual definition is a contextual one, of the sort introduced by Russell. Take the definite description operator. Following Neale [1: p. 45] (cf. Sharvy [1: fn. 5]), it will be useful to indicate these terms as quantifiers. Thus

$$(4.1)\ [\iota xFx]Gx$$

is read: 'of the (unique) object which is *F*, it is *G*', that is, there is some object which is uniquely *F*, and it is *G*. Note that all three occurrences of '*x*' which are exhibited in (4.1) are bound by 'ι'. We see immediately that (1.1) is ambiguous, between

$$(4.2)\ [\iota x(x = a.\, p)](x = a \leftrightarrow p)$$

(that is, there is one and only one thing which is a and p holds, and it is a iff p holds) and

$$(4.3)\ ([\iota x(x = a.\, p)]x = a) \leftrightarrow p$$

(that is, *p* holds iff there is one and only one thing which is *a* and *p* holds and it is *a*).

In (4.2) the description has wide scope; in (4.3) it has narrow scope. (4.2) treats it *de re*; (4.3) *de dicto*. Most importantly, (4.2) is false, while (4.3) is true.[12] For (4.2) asserts the existence of an object which is *a* iff *p* holds, and although if *p* holds there is indeed such an object (namely, *a*), its being *a* is not logically equivalent to *p*, for *a*'s being *a* is necessarily true, whereas *p* was taken as a contingent truth. Recall that it was important for the paradox in §1 that '$\leftrightarrow$' in (1.1) and (1.7) be *logical* equivalence. Of course, if *p* is true, then *p* and '$a = a$' are *materially* equivalent, and if *q* is also true, so are *p* and *q*. (4.2) is true if '$\leftrightarrow$' is material equivalence, but false if it is a stronger notion.

On the other hand, (4.3) only asserts the object's existence and identity with *a* on condition that *p* holds (and vice versa), and that can indeed be established logically. That is, (4.3) is (logically) equivalent to

$$(4.4)\ (\exists x)((\forall y)(y = a.\, p \equiv y = x).\, x = a) \leftrightarrow p$$

[12] The ambiguity was clearly spelled out, even as Quine took over the argument from Church, in Smullyan [1].

which is a valid logical equivalence.

But we want (1.1) to join with (1.5) to entail (1.6). (1.5) cashes out as

$$(4.5)\ [\iota x(x = a.\, p)][\iota y(y = a.\, q)]x = y;$$

and (1.6) can be read as either

$$(4.6)\ [\iota x(x = a.\, q)](x = a \leftrightarrow p)$$

or

$$(4.7)\ ([\iota x(x = a.\, q)]x = a) \leftrightarrow p$$

But although (4.2) and (4.5) entail (4.6), (4.3) and (4.5) do not yield either (4.6) or (4.7). That is, (SI), which allows the replacement of genuine singular terms, can only be used to replace contextually defined descriptions if they have wide scope, and cannot affect the narrow scope descriptions in (4.3) and (4.7). The inference from (1.1) and (1.5) to (1.6) is either invalid—taking (1.1) as (4.3)—or has a false premise—taking (1.1) as (4.2), even on the best case scenario of reading (1.6) as (4.6).

The same diagnosis applies to the arguments in §§2 and 3. If (3.1) is true, then the description has narrow scope there and hence so too in (3.5), now read as

$$(4.8)\ \Phi([\iota x(x = a.\, p)]x = a),$$

and so (SI) is inapplicable, and (3.6) does not follow. Again, in Church's argument to (2.5), there was an equivocation over the scope of the class term, if it is also to be contextually defined, reading (2.1) as

$$(4.9)\ [\iota y(\forall x)(x \in y \equiv (x = x.\neg p))]\,\mathrm{Des}(S_1,\, y = \Lambda)$$

or as

$$(4.10)\ \mathrm{Des}(S_1,\, [\iota y(\forall x)(x \in y \equiv (x = x.\neg p))]y = \Lambda).$$

(4.9) entails (2.4), since (2.3) is true, but is false, while (4.10) is true, but cannot interact appropriately with (2.3) to yield (2.4).

To see why (4.9) is false: consider S_3, say, which is to mean 'Someone is a spy'. Then $\mathrm{Des}(S_3, (\exists x)x \text{ is a spy})$. But there is no particular person, x, such that $\mathrm{Des}(S_3, x \text{ is a spy})$. The same is true of S_1. Although, on the Russellian expansion, S_1 means 'there is a set which is empty just when p is true', there is no particular set which S_1 says is empty just when p is true.

What set could that be? Not the empty set, for that is always empty, regardless of whether p is true. Nor the universal set, since that is never empty. (4.9) is false.

5. *The Fregean analysis.* Vbts may, however, be treated as primitive terms, that is, as fully legitimate terms in proof theory and semantics. Again, let us work with the example of definite descriptions. For that case, the approach is essentially Frege's. First, let us look at the proof theory for extensional languages, as formulated by da Costa [1: p. 138]:

I: $\iota xFx = \iota yFy$

II: $(\forall x)(Fx \equiv Gx) \supset \iota xFx = \iota yGy$

(these hold of any vbt)

III: $(\exists! x)Fx \supset (\forall y)(y = \iota xFx \supset Fy)$

where $(\exists! x)Fx$ abbreviates $(\exists x)(\forall y)(Fy \equiv x = y)$, and

IV: $\neg(\exists! x)Fx \supset (\iota xFx = \iota x. x \neq x)$

This ensures that ιxFx denotes the F when there is one and only one F, and denotes some arbitrary object (one and the same for all F) when there isn't. So ιxFx always denotes, and can appear as an ordinary singular term, s or t, in the law of substitutivity of identicals (SI)—see §3; in the Law of Identity,

$$t = t$$

for all terms t, including descriptions; and in the quantifier rules: Existential Generalization,

$$\frac{F(t)}{(\exists x)F(x)},$$

and Universal Specification,

$$\frac{(\forall x)F(x)}{F(t)}.$$

There are two ways to extend this theory to intensional languages. One is that chosen by da Costa and Mortensen [1], which amounts to treating descriptions as flexible designators varying their designation across possible worlds. We will look at it later (§7). First, let us consider an alternative approach which keeps much closer to the extensional model by treating descriptions as rigid designators whose designation is fixed by how things actually are. It transpires that in that way we can keep the above four axioms (for definite descriptions), SI, Identity and the quantifier rules. Nonetheless, something has to give, on pain of paradox.

Here is the problem: suppose $H(\iota xFx)$ is true, for certain H and F, and further, suppose there is exactly one F. Then by (III), Identity, Universal Specification and detachment, we obtain $F(\iota xFx)$. Since by hypothesis there is at most one F, it follows that

$$(\forall y)(Fy \equiv y = \iota xFx).\, H(\iota xFx)$$

whence by Existential Generalization and Addition,

$$(5.1)\ (\exists x)\big((\forall y)(Fy \equiv y = x).\, Hx\big) \vee \neg(\exists x)(\forall y)(Fy \equiv y = x).\, H(\iota x.\, x \neq x)$$

But if there isn't exactly one F, then by (IV), $\iota xFx = \iota x.\, x \neq x$, so $H(\iota x.\, x \neq x)$ by (SI), whence (5.1) again follows by Addition. Indeed, the steps are reversible, so we have the equivalence

$$(5.2)\ H(\iota xFx) \equiv \big[(\exists x)\big((\forall y)(Fy \equiv y = x).\, Hx\big) \vee \neg(\exists x)(\forall y)(Fy \equiv y = x).\, H(\iota x.\, x \neq x)\big]$$

For the paradox, consider an intensional language with '$\Box$' for necessity, and take the wff $\Box G(\iota xFx)$. By (5.2) we have

$$(5.3)\ \Box G(\iota xFx) \equiv \big[(\exists x)\big((\forall y)(Fy \equiv y = x).\, \Box Gx\big) \vee \neg(\exists x)(\forall y)(Fy \equiv y = x).\, \Box G(\iota x.\, x \neq x)\big]$$

but also

$$(5.4)\ G(\iota xFx) \equiv \big[(\exists x)\big((\forall y)(Fy \equiv y = x).\, Gx\big) \vee \neg(\exists x)(\forall y)(Fy \equiv y = x).\, G(\iota x.\, x \neq x)\big]$$

The paradox emerges by asking if we can strengthen the equivalence in (5.4) (and 5.2) to a co-entailment, that is, do we have

$$(5.5)\ \Box\big[G(\iota xFx) \equiv \big((\exists x)\big((\forall y)(Fy \equiv y = x). Gx\big) \vee \neg(\exists x)(\forall y)(Fy \equiv y = x). G(\iota x. x \neq x)\big)\big]$$

If so, then by the Kripke axiom

$$\Box(p \supset q) \supset (\Box p \supset \Box q)$$

we would obtain

$$(5.6)\ \Box G(\iota xFx) \equiv \Box\big[(\exists x)\big((\forall y)(Fy \equiv y = x). Gx\big) \vee \neg(\exists x)(\forall y)(Fy \equiv y = x). G(\iota x. x \neq x)\big]$$

and so by (5.3) and (5.6)

$$(5.7)\ \big[(\exists x)\big((\forall y)(Fy \equiv y = x). \Box Gx\big) \vee \neg(\exists x)(\forall y)(Fy \equiv y = x). \Box G(\iota x. x \neq x)\big] \equiv \Box\big[(\exists x)\big((\forall y)(Fy \equiv y = x). Gx\big) \vee \neg(\exists x)(\forall y)(Fy \equiv y = x). G(\iota x. x \neq x)\big]$$

But (5.7) is false. Let 'Fx' read 'x is a natural satellite of the Earth' and 'Gx' read 'x is a physical object', and let '$\iota x. x \neq x$' denote the number zero—some arbitrary object which is necessarily non-physical. Then the left-hand side of (5.7) is true: there is in fact a unique natural satellite of the Earth, the Moon, and it's necessarily physical; but the right-hand side is false: it is quite possible that the Earth should have had no natural satellite (or more than one), while its second disjunct is always false, given the choice of designation for '$\iota x. x \neq x$'.[13]

There are undoubtedly several solutions to this paradox—one of them is da Costa and Mortensen's, to be considered in §7. What the paradox shows is that it is not altogether straightforward to extend the treatment of vbts as primitive in extensional logic to the intensional case. Something has to give. If we are to continue to treat the vbts *de re*, the solution has to be to

[13] See Marti [1: p. 581]. $H(\iota xFx)$ is unambiguous in extensional logic: see, e.g., Kalish and Montague [1: pp. 258, 264].

restrict application of the rule of Necessitation, in particular, to deny that (5.5) follows from (5.4).[14] Necessitation converts *de re* truths into *de dicto* ones, as we see by tracking (5.3) through (5.5) to (5.6). But this move is fallacious. Reading the description *de re*, 'The number of planets is necessarily odd' is true—that object, 9, which numbers the planets is of necessity, odd; but read *de dicto*, it is false: we cannot infer that there is necessarily an odd number of planets, as (5.7) would invite us to do.

6. *The Fregean diagnosis.* We saw that, when eliminated by contextual definition in the Russellian manner, complex singular terms either rendered invalid the superficially correct instances of (SI) in the Slingshot, or falsified the minor premise, depending on the scope accorded to them. But that diagnosis seems set to fail now, for (SI) applies universally to vbts as to any other singular terms and those vbts are being treated *de re*. So the minor premise is false.

Nonetheless, the argument is still unsound. The reason is that the major premise of one application of (SE) now fails. For it is crucial to the argument that we be able to establish the truth of the premises (1.1) and (1.2). Only thereby may we use (SE) to replace 'p' by the identity claim

$$\iota x(x = a.\, p) = a$$

and the identity claim

$$\iota x(x = a.\, q) = a$$

by 'q'. The latter move is no longer sound, however, since

$$(6.1)\ \iota x(x = a.\, q) = a \to q$$

fails to hold if the description is treated as primitive and *de re* and '$\to$' is logical implication.[15] It depends on the truth of q itself, and so holds only materially.

[14] One can ensure that Necessitation only applies to logical truths, not those deriving from uses of (I) - (IV), by, e.g., dropping it as a rule, and replacing the logical axioms by their necessitations. The Kripke axiom (see above) then serves to transmit necessitations through applications of detachment.

[15] This was recognised by Taylor [1], p. 36.

What may seem puzzling is that of course if $\iota x(x = a.q) = a$, it is natural to infer that q must be true. How else could $\iota x(x = a.q)$ be a? But this is to use

$$(6.2)\ (\forall y)(y = \iota x(x = a.q) \supset y = a.q)$$

and its instance with 'y' replaced by '$\iota x(x = a.q)$'. But to do so we need to detach (6.2) from an instance of (III), and to do that we need the truth of $(\exists! x)(x = a.q)$, which in turn requires the truth of q. So the argument is circular. We must realise that, on taking descriptions as primitive, it follows that they always denote, even when there is nothing of which they are true. So q can be false even when '$\iota x(x = a.q)$' denotes a.

We can put the point formally as follows. (I) - (IV) are true in a singleton domain containing a alone. Then every description 'ιxFx' denotes a. We can take 'q' to be false, without contradiction. Nonetheless, '$\iota x(x = a.q)$' denotes a, so the antecedent of (6.1) is true while its consequent is false.

Let's think how a proof of (6.1) will fail. Suppose $\iota x(x = a.q) = a$. We need to deduce q. So suppose $\neg q$. Then $\neg(\exists! x)(x = a.q)$. So by (IV),

$$\iota x(x = a.q) = \iota x(x \neq x),$$

whence

$$a = \iota x(x \neq x).$$

What we would like to conclude here is that $a \neq a$ (for *reductio*, to conclude $\neg\neg q$ and so q) by Universal Specification from

$$(6.3)\ (\forall y)(y = \iota x(x \neq x) \supset y \neq y)$$

given that we know that $a = \iota x(x \neq x)$. But (6.3) is not available. It is the consequent of an instance of (III), and we know that the antecedent of that instance is false in the case where Fx is $x \neq x$. Indeed, the existential antecedent of (III) is crucial to the avoidance of contradiction. For suppose (6.3) were universally valid. Then we could infer

$$\iota x(x \neq x) = \iota x(x \neq x) \supset \iota x(x \neq x) \neq \iota x(x \neq x)$$

and so the Law of Identity would invalidate itself.[16]

[16] So we should avoid reading $\iota x(x \neq x)$ as 'that object which is not self-identical'. Every object is self-identical, including the object which $\iota x(x \neq x)$ denotes. It denotes an

We cannot avoid this impasse by choosing v (the vbto) to be set abstraction. For we may no more state generally that

$$(6.4)\ (\forall y)(y \in \hat{x}Fx \equiv Fy)$$

than we may say for the description operator that $(\forall y)(y = \iota xFx \supset Fy)$. (6.4) holds only for those set abstracts $\hat{x}Fx$ which genuinely denote sets. But if Fx were $x \notin x$ (the notorious case of Russell's paradox) we could infer

$$\hat{x}(x \notin x) \in \hat{x}(x \notin x) \equiv \hat{x}(x \notin x) \notin \hat{x}(x \notin x),$$

and land in contradiction.

It follows that argument (1) fails. (1.3) and (1.4) still hold if p and q are true, and so (1.5) and

$$p \rightarrow \iota x(x = a.q) = a$$

follow. But that is no problem. The point is that the step from there to '$p \rightarrow q$' fails in the absence of (6.1). Similarly in argument (3), from $\Phi(p)$ we proceed as usual by using (3.1) from right to left and (3.3) to obtain

$$\Phi(\iota x(x = a.q) = a)$$

But without (6.1) we cannot proceed to (3.7) and the triviality the anti-intensionalists desire.

7. *A Fregean contextual definition.* There is yet a third way of treating descriptions (and other vbts). However, although similar to the first way in treating them contextually, it turns out to be equivalent to the second way, that is, to treating them as primitive terms subject to (SI). Postulates (III) and (IV) in §5 tell us that if $(\exists! x)Fx$ then $F(\iota xFx)$, while if $\neg(\exists! x)Fx$, $F(\iota xFx)$ only if $F(\iota x.x \neq x)$. On Russell's contextual definition, that is, the first way, all apparent claims about the F are rendered false when there is no unique F. But on the alternative contextual definition, the different approach is taken of interpreting such "empty terms" as denoting an arbitrary object. This approach, the Fregean contextual definition, is, therefore, equivalent to the second, the treatment of descriptions as true singular

arbitrary object, the denotation of every absurd description (among others). But it is not itself absurd, since no object is absurd!

terms subject to postulates (III) and (IV). (5.2) is now taken as the defining axiom.

It follows that arguments (1) and (3) fail for the same reason as was brought out in §6, namely, the failure of the major premise of the step of (SE)—(1.2) and (3.2) respectively. It remains to expose the error in Church's argument, as articulated on the second (and third) accounts of vbts.[17] As one will by now expect, there is no quarrel with the application of (SI), and so no quarrel with that part of the argument which constitutes Frege's original argument for distinguishing sense from reference, that is, for the claim that (2.5) Syn(S_1, S_2). Given the Russellian contextual definition, this argument fails, as Russell observed,[18] to demonstrate the need for a notion of sense. But with descriptions (and class terms) treated as Fregean (primitive and *de re*, or given the Fregean contextual definition) the original argument goes through: S_1 and S_2 are indeed co-designative or "synonymous", and so if we see a difference in meaning between them, it must lie in some other notion of "sense" beyond matters of designation.

The fault must now be found in the argument for (2.6) Syn(p, S_1). The reason given for accepting it was that S_1 is true if and only if p is, and by that was meant that they "are L-equivalent and therefore synonymous".[19] But they are not logically equivalent—if the vbt is treated in the second or third ways. The whole argument exhibits the fallacy of four terms—it equivocates on the middle term. Either the vbt is given a Russellian definition, in which case Frege's argument, depending on (SI) fails; or it is taken as primitive and *de re*, or given a Fregean definition, in which case the second part of the argument, due to Church, fails. For on this account, p and S_1 (i.e. $\lambda x(x = x.\neg p) = \Lambda$) are not logically equivalent, and so we are not entitled to assert (2.6), and so (2.7) does not follow. The error lies in exactly the same place as was identified with arguments (1) and (3) in §6.

For the sake of completeness, a few words should be given to a fourth and final method of treating vbts in intensional contexts, namely, that of da

[17] I leave as an exercise for the reader the diagnosis of Church's later version of his argument in Church [3: pp. 24-5]. It uses two vbtos, ι and the 'number of' operator, f, defined by, say, Hume's principle. (On Hume's principle, see Heck [1: p. 579 ff], and on 'f' see da Costa [1: p. 151].) The argument moves from $s = \iota x.Axw$ to $s = \iota x.fy(Axy.Wy) = 29$ by (SI), $fy(s = \iota xAxy.Wy) = 29$ by (SE) and finally by (SI) again to $fyUy = 29$, where s means Scott, w means *Waverley*, Axy means 'x is the author of y', Wx 'x is a Waverley novel', and Ux 'x is a Utah county'.

[18] See Russell [1: p. 483 ff].

[19] Church [2: p. 300].

Costa and Mortensen [1].[20] Descriptions are again treated as primitive, but *de dicto.* This time, the paradox of §5 is blocked by restricting (SI), and indeed the necessity of identity fails. Instead, substitution in intensional contexts requires a suitably strong identity: with S5-modality, (SI) becomes

$$\frac{\Phi(p) \quad \Box s = t}{\Phi(p')}$$

with p and p' as in §3 (though in weaker modal logics, the identity premise requires stronger prefixing).[21] Under this treatment, descriptions are treated flexibly—'ιxFx' designates in each world whatever is uniquely F in that world (if there is such a thing), otherwise the arbitrary and constant referent of absurd descriptions. Unsurprisingly, therefore, the derivations in §§1 - 3 fail in the same way as was shown in §4 for narrow-scope Russellian descriptions.

A case analysis of the ways of treating vbts, and how the slingshot argument fails (in different ways) in each of them, is given in Fig. II.[22]

[20] See also Carnap [2: §§ 7-8]. Liu [1] shows how to treat descriptions as primitive singular terms equivalent to the Russellian contextual definition.

[21] See da Costa and Mortensen [1: Theorem 3].

[22] Although Frege treated all names as denoting in his preferred scientific language, he conceded that in ordinary discourse many names lack reference. The various ways of treating atomic sentences containing such empty names—e.g., lack of truth-value, or always false—and their compounds are collectively called "free logic". The analysis of the slingshot argument's cogency in a free logic context must be the subject of a separate and future study.

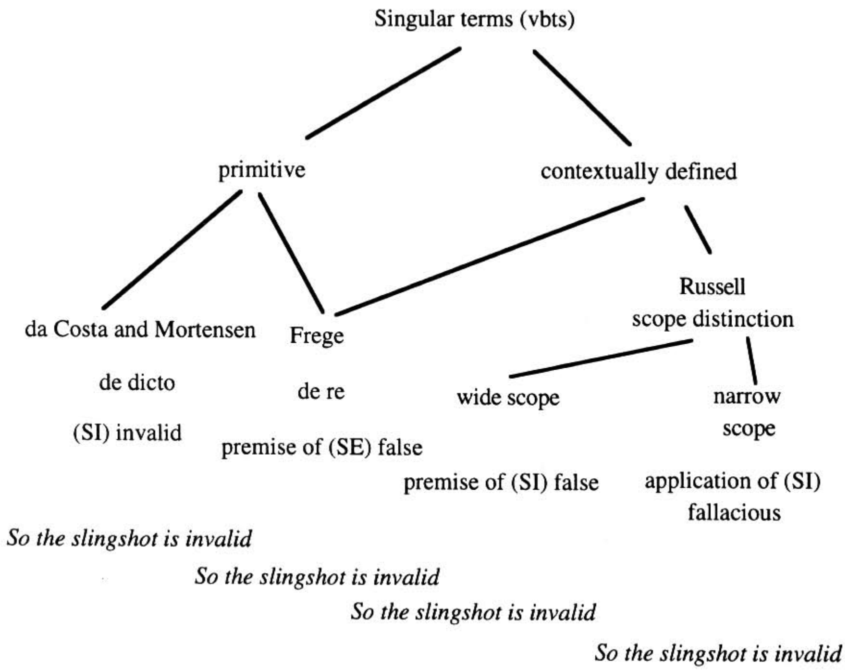

Fig. II

8. *Substitutivity.* Those who recall Quine's general attack on intensional logic will be puzzled by the treatment of vbts in §§5, 6 and 7. How can (SI) hold universally? Has Quine not shown that it leads to absurd results, in particular, to being forced to accept the truth of

(8.1) The number of planets is necessarily greater than 7,

which would follow by (SI) from the uncontentious truth

(8.2) 9 is necessarily greater than 7

given that

(8.3) The number of planets is 9.[23]

On the Russellian analysis, of course, (8.1) is ambiguous, true when the description is taken with wide scope in its eliminative definition, false when read with narrow scope. Crucially, the false narrow scope reading does not follow from (8.2) by a correct application of (SI). The description is not a true singular term, and (SI) can only correctly be applied to expressions in primitive terms.[24]

But when vbts are treated as primitive and *de re*, (SI) applies to all uses of such singular terms, so it seems that (8.1) will follow validly from (8.2) by (SI). Yet surely it is false?

It is not false, for what we must realise is that when 'the number of planets' is treated as primitive, its function is simply to refer to a certain object. That is why, when vbts are treated as in §5, there is some justice in the claim that they are universally treated as having "wide scope". (8.1) says of a certain object that it is necessarily greater than 7, and that is true, for 9 is so, as (8.2) says.

What then lies behind Quine's rhetoric, which construes (8.1) as false, and resists the substitution of 'the number of planets' for '9' in (8.2) to obtain (8.1)? "What is this number which, according to ['$(\exists x)(x$ is necessarily greater than 7'] is necessarily greater than 7? According to [8.2], from which [it] was inferred, it was 9, that is, the number of planets; but to suppose this would conflict with the fact that [8.1] is false."[25] When the vbt 'the number of planets' is treated *de re*, (8.1) is true; for the singular term is then construed simply as referring to 9, and there is no reading of (8.1) on which it is false. What is false is 'It is necessary that whatever number of planets there is, it is necessarily greater than 7', when it is read (with narrow scope) as

$$(8.4)\quad \Box(\forall y)(Ny \supset y > 7),$$

where 'Ny' abbreviates 'y numbers the planets' or 'there are y planets'. For (8.4) is false in virtue of the possibility that there be, say, 6 planets. But it is essential to the falsity of (8.4) that 'N' lie within the scope of '$\Box$'; and when the description 'the number of planets' in (8.1) is taken as a *de re* singular term governed by the da Costa postulates, that is impossible. It lies outside the scope of the phrase 'necessarily', is open to substitution, and serves simply to pick out the actual number of planets, *viz* 9.

23 See, e.g., Quine [1, 2; and 3: §41].

24 See Church [1].

25 Quine [3: p. 24].

9. *Conclusion.* Intensional logic depends on the existence of contexts $\Phi(p)$ which are not so opaque that 'p' may not be replaced, and so in a certain sense not seen, at all; nor so transparent that 'p' may be replaced merely by any proposition with the same truth-value.[26] The Slingshot Argument seems to show that there can be no such contexts, that if in $\Phi(p)$, 'p' may be replaced by logical equivalents, and its singular terms replaced by coreferential ones, then Φ is truth-functional. But the argument is unsound; and where its fallacy lies depends on exactly how complex singular terms are treated. If they are eliminated by a Russellian contextual definition, then the apparent use of (SI) is ill-formed. This is not to say that (SI) itself fails in any way, but that, since the singular terms to which it was applied have been eliminated, its apparent form is illusory. Correctly applied at the level of real singular terms, the conclusion is not forthcoming. The argument fails.

In contrast, when descriptions are treated as *de dicto* vbts, (SI) really does fail. It is an unacceptable rule, as are other standard inference patterns. So again the argument is fallacious in relying on an illicit use of (SI).

Alternatively, the singular terms may be treated either as *de re* variable bound terms, or eliminated by a Fregean contextual definition. Then no problem arises in the application of (SI). In each case, however, we find the flaw now lies in the major premise of the second step of (SE). The logical equivalence of p and $\iota x(x = a.\, p) = a$ fails. If p is true, $\iota x(x = a.\, p)$ is certainly a. But from the fact that $\iota x(x = a.\, p)$ is a, it does not follow that p is true. p could be false, and a the object which "empty names" denote.

It is important to recognise that the failure of (SE) only affects that part of Church's argument which is not found in Frege. What is really "Frege's argument", the part using (SI), is perfectly sound (on Fregean principles, of course, not Russellian). The error emerges in the second half of the argument, when (SE) is applied.

Although the Slingshot Argument is fallacious, some may discern in this diagnosis an ultimate and ironic justification of Quine's attack on substitutivity. For (SI) was what Quine wished to fault in attacking the coherence of quantified modal logic. Opacity was by definition the failure of (SI). But opacity would seal off not only the terms from (SI) but the variables from "quantifying in". However, either one contextually defines the terms by the Russellian or Fregean definitions, in which case there is no term to which to apply or from which to withhold (SI); or one treats the vbts as genuine terms, in which case one may not unrestrictedly assume they denote objects

[26] See Wolenski [1] and Neale [2: p. 111].

with the property expressed by the open formula from which they are constructed. That was the moral of the reduction of (6.3), or more generally of

$$(\forall y)(y = \iota x Fx \supset Fy),$$

i.e. $F(\iota xFx)$, to absurdity. Perhaps the reason why so many logicians continued to employ the Slingshot Argument after it had been diagnosed as fallacious when given the Russellian analysis, was that they rejected that analysis and wanted to take the terms used in it as primitive and *de re*.[27] Certainly, their practice suggests this. But we now see that the argument then turns on a purported logical equivalence which the theory of vbts and vbtos cannot endorse. Intensional logic emerges unscathed from attack by the Slingshot Argument.[28]

University of St Andrews, Scotland

BIBLIOGRAPHY

G.E.M. Anscombe [1] "Causality and extensionality", *Journal of Philosophy* 66, 1969, 152-9.
J. Barwise and J. Perry [1] *Situations and Attitudes*, M.I.T. Press 1983.
J. Barwise and J. Perry [2] "Semantic innocence and uncompromising situations", *Midwest Studies in Philosophy* 6, 1981, 387-403.
R. Carnap [1] *Introduction to Semantics*, Harvard U.P. 1942.
R. Carnap [2] *Meaning and Necessity*, Chicago U.P. 1947.
A. Church [1] "Review of W. Quine, 'Whitehead and the rise of modern logic'", *Journal of Symbolic Logic*, 1942, 100-1.
A. Church [2] "Review of R. Carnap, *Introduction to Semantics*", *Philosophical Review* 52, 1943, 298-304.
A. Church [3] *Introduction to Mathematical Logic* 7, Princeton U.P., 1956.
J. Corcoran and J. Herring [1] "Notes on a semantic analysis of variable binding term operators", *Logique et Analyse* 14, 1971, 644-57.
R. Cummins and D. Gottlieb [1] "On an argument for truth-functionality", *American Philosophical Quarterly* 9, 1972, 265-9.

[27] See, e.g., Scott's remarks in Scott [1: p. 191].

[28] I must acknowledge a particular debt to Graham Priest for invaluable discussions in the course of composing this paper.

N. da Costa [1] "A model-theoretical approach to variable binding term operators", in *Mathematical Logic in Latin America*, ed. A. Arruda et al., North-Holland, 1980, 133-62.

N. da Costa and C. Mortensen [1] "Variable binding term operators in modal logic", *Pré-publicações do Centro de Lógica, Epistemologia e História da Ciência—CLE*, no. 4, Campinas-SP, Brazil, 1980.

D. Davidson [1] "Causal relations", *Journal of Philosophy* 64, 1967, 691-703.

D. Davidson [2] "Truth and meaning", *Synthese* 17, 1967, 304-23.

D. Davidson [3] "True to the facts", *Journal of Philosophy* 66, 1969, 748-64.

M. Davies [1] *Meaning, Quantification, Necessity*, Routledge and Kegan Paul, 1981.

D. Føllesdal [1] "Quantification into causal contexts", *Boston Studies in the Philosophy of Science* II, 1965, 263-74.

G. Frege [1] "On sense and reference", tr. in *Philosophical Writings*, ed. P. Geach and M. Black, Blackwell, 1966.

K. Gödel [1] "Russell's mathematical logic", in *The Philosophy of Bertrand Russell*, ed. P.A. Schilpp, Northwestern U.P., 1944, 123-53.

K. Gödel [2] *Collected Papers*, vol. II, ed. S. Feferman et al., Oxford U.P. 1990.

W. Hatcher [1] *Foundations of Mathematics*, W.B. Saunders Co. 1968.

W. Hatcher [2] *The Logical Foundations of Mathematics*, Pergamon Press 1982.

R. Heck [1] "The development of arithmetic in Frege's *Grundgesetze der Arithmetik*", *Journal of Symbolic Logic* 58, 1993, 579-601.

H. Hochberg [1] "Facts and truth", in his *Logic, Ontology and Language*, Philosophia Verlag, 1984, 279-95.

D. Kalish and R. Montague [1] *Logic: techniques of formal reasoning*, Harcourt, Brace and World Inc., 1964.

S.-C. Liu [1] "A theory of first order logic with description", *Bulletin of the Institute of Mathematics, Academia Sinica* 2, 1974, 43-65.

W.G. Lycan [1] "The extensionality of cause, space and time", *Mind* 83, 1974, 498-511.

J.L. Mackie [1] *The Cement of the Universe*, Oxford U.P., 1974.

G. Marti [1] "Do modal distinctions collapse in Carnap's system?", *Journal of Philosophical Logic* 23, 1994, 575-93.

A. Morton [1] "Extensional and non-truth-functional contexts", *Journal of Philosophy* 66, 1969, 159-64.

S. Neale [1] *Descriptions*, M.I.T. Press, 1990.

S. Neale [2] "Term limits", *Philosophical Perspectives* 7, 1993, 89-123.

S. Neale [3] "Are there facts worth representing?", typescript 2 May 1994.

W.V. Quine [1] "Notes on existence and necessity", *Journal of Philosophy* 40, 1943, 113-27.

W.V. Quine [2] "The problem of interpreting modal logic", *Journal of Symbolic Logic* 12, 1947, 43-8.

W.V. Quine [3] "Reference and modality", in his *From and Logical Point of View*, Harper Torchbooks, 1953, 139-59.

W.V. Quine [4] *Word and Object*, M.I.T. Press, 1960.

W.V. Quine [5] "Three grades of modal involvement", in his *Ways of Paradox*, Random House, 1966, 156-74.

J.B. Rosser [1] *Logic for Mathematicians*, McGraw-Hill, 1953.

B. Russell [1] "On denoting", *Mind* 14, 1905, 479-93.

R.M. Sainsbury [1] *Logical Forms*, Blackwell, 1991.

D. Scott [1] "Existence and description in formal logic", in *Bertrand Russell: Philosopher of the Century*, ed. R. Schoenman, George Allen and Unwin, 1967, 181-200.

R. Sharvy [1] "Truth-functionality and referential opacity", *Philosophical Studies* 21, 1970, 5-9.

A. Smullyan [1] "Modality and description", *Journal of Symbolic Logic* 13, 1948, 31-7.

B. Taylor [1] *Modes of Occurrence*, Blackwell, 1985.

J. Wallace [1] "Propositional attitudes and identity", *Journal of Philosophy* 66, 1969, 145-52.

J. Wolenski [1] "Remarks on extensionality and intensionality", in *Philosophical Logic in Poland*, ed. J. Wolenski, Kluwer Academic Publishers, 1994, 321-5.