Deflationism and Gödel's theorem – a comment on Gauker

Panu Raatikainen

In his recent article Christopher Gauker (2001) has presented a thought-provoking argument against deflationist theories of truth. More exactly, he attacks what he calls 'T-schema deflationism', that is, the claim that a theory of truth can simply take the form of *certain* instances of the T-schema:

(T) 'p' is true $\leftrightarrow p$.

Gauker's main claim is that 'if there is any satisfactory way of spelling out the instances of (T) comprising the theory of truth of our language, then the theory contradicts Gödel's first incompleteness theorem' (130).

Gauker's argument is based on a trick due to Vann McGee (1992), which provides, for every sentence S, an instance of (T) that is materially equivalent to S. (This is achieved by diagonalizing the open formula 'x is true $\leftrightarrow$ S'; one then obtains a sentence G such that $G \leftrightarrow$ ('G' is true $\leftrightarrow S$); and by propositional logic, $S \leftrightarrow$ ('G' is true $\leftrightarrow G$)). Gauker's own argument begins with the requirement that 'it is surely necessary to explain how the instances of (T) that express the theory might be *effectively enumerated*' (130). He then points out that the set of true sentences of arithmetic is not recursively enumerable. But, so the argument continues, it would be so if the instances of (T) expressing the deflationist's theory of truth were recursively enumerable. 'So the deflationist cannot after all have what he or she needs', Gauker concludes (134).

Although I feel no need to defend deflationist theories – I think they face various difficult problems – I think that Gauker's argument is in fact less efficient than it appears: its exact relevance is much less clear than Gauker's presentation suggests.

My basic reason for thinking so is that Gauker's consideration ignores totally one very natural version of T-schema deflationism, namely the one which takes seriously the Tarskian distinction between object language and metalanguage, and requires that the sentence p in (T) must be a sentence of the object or base language. That is, let us assume that we have a base language L (the language of arithmetic or its extension). One then forms the metalanguage by adding a new predicate $Tr(x)$ and a name 'S' for every sentence S of the base language, and adds the schema

(T) $Tr(\text{`}S\text{'}) \leftrightarrow S$,

Analysis 62.1, January 2002, pp. 85–87.

where *S* is required to be a sentence of the original base language; let us call the version of T-schema deflationism thus restricted *minimalistic deflationism*.[1] Now clearly the set of such instances of (T) is recursively enumerable (in fact even primitive recursive). Hence Gauker's argument has no force against minimalistic deflationism. His argument only works for the cases where one is allowed to iterate the truth predicate; or, in other words, the sentence *S* in (T) is itself allowed to contain the truth predicate (*G* in the argument requires this). And Gauker gives no reasons why a deflationist should use such an approach instead of the minimalistic one.

Moreover, Gauker begins his argument by assuming that every sentence is (absolutely) either true or false or *neither*. Indeed, if one is allowed to apply the truth predicate to a sentence containing the truth predicate, as Gauker assumes throughout his paper, it seems to be necessary to allow something like truth-value gaps in order to avoid the Liar paradox etc.

But Gauker seems not to realize that this already refutes T-schema deflationism (see e.g., Horwich 1998: 76–77) and makes any further, more complicated argument redundant. That is, let *B* be any sentence (not necessarily paradoxical) which is neither true nor false. Then '"*B*" is true' is false. But it then follows that the equivalence '"*B*" is true ↔ *B*' cannot hold, for the left-hand side is false but the right-hand side is not (cf. Schmitt 1995: 136–41, Stoljar 1997; the argument goes back to Dummett 1959). And the restriction of the theory of *truth* to the instances of (T) where *S* is either *false* or *true* would certainly beg the question.

If, on the other hand, one sticks to bivalence, then the 'Tarskian' approach amounting to minimalistic deflationism appears to be very natural and appealing; but against it Gauker's argument fails, as we have already seen.

In conclusion, Gauker should at least point out an interesting version of deflationism, one which would not allow problematic truth-value gaps, but which would allow sentences like *G* above, and against which his argument would apply. It remains to be seen if there exist any.[2]

P.O. Box 9, FIN-00014
University of Helsinki, Finland
panu.raatikainen@helsinki.fi

[1] Minimalistic deflationism has an attractive (from the deflationist point of view) property that it is consistent relative to and even conservative over the base theory; that is, the addition of thus restricted instances of T-schema does not entail any sentence of the base language that is not entailed already by the base theory (see e.g., Ketland 1999).

[2] I am indebted to Jeffrey Ketland for our useful exchange of views on these matters.

References

Dummett, M. 1959. Truth. *Proceedings of the Aristotelian Society* 59: 141–62.
Gauker, C. 2001. T-schema deflationism versus Gödel's first incompleteness theorem. Analysis 61: 129–36.
Horwich, P. 1998. *Truth*. 2nd ed. New York: Oxford University Press.
Ketland, J. 1999. Deflationism and Tarski's paradise. *Mind* 108: 69–94.
McGee, V. 1992. Maximally consistent sets of instances of Tarski's schema (T). *Journal of Philosophical Logic* 21: 235–41.
Schmitt, F. 1995. *Truth: A Primer*. Boulder: Westview Press.
Stoljar, D. 1997. Truth, Deflationary theories of. In *Stanford Encyclopedia of Philosophy*, ed. Edward N. Zalta. URL = http://plato.stanford.edu/entries/truth-deflationary/.

Anti-individualism and analyticity

ANTHONY BRUECKNER

Jessica Brown (1995), following Michael McKinsey (1991), has defended an argument aimed at establishing the incompatibility of anti-individualism and privileged access. The basic idea of the Brown-McKinsey *reductio* argument is that if anti-individualism is true, then one can come to know some substantive proposition about one's external environment purely by knowing one's own thoughts and reflecting on the consequences of anti-individualism. In Falvey 2000 and McLaughlin & Tye 1998, there are persuasive objections to Brown's defence of such an argument. Her subsequent response (Brown 2001) depends upon the following claim:

> (C) A subject can know a priori (that is, without depending in a justificatory way on empirical investigation) that (1) he is unsure about whether a certain concept applies to a type of thing, and (2) there is a determinate fact about whether it does so apply.

The way in which (C) fits into her considered *reductio* strategy can be seen by considering the following situation. Suppose that S is unsure of whether his non-natural-kind term 'sofa' is true of a certain range of cases, say, large armchairs. Suppose further that the concept sofa determinately fails to apply to the cases in the range. Then, according to Brown, we can, given anti-individualism, correctly attribute the concept sofa to S *only if* S is a member of a community of speakers who possess sofa. Given that there is no relevant natural kind in S's environment, given S's lack of dispositions regarding the use of 'sofa' in the range of cases even though sofa does deter-

ANALYSIS 62.1, January 2002, pp. 87–91.