

The Inconsistency of the Identity Thesis

Christopher Hom
Texas Tech University
Stanford Humanities Center

Robert May
University of California, Davis

April 18, 2012

1 The Identity Thesis for Pejoratives

In theorizing about racial pejoratives, an initially attractive view is that pejoratives have the same reference as their “neutral counterparts”. Call this the *identity thesis*. According to this thesis, the terms “kike” and “Jew”, for instance, pick out the same set of people. To be a Jew *just is* to be a kike, and so to make claims about Jews *just is* to make claims about kikes. In this way, the two words are synonymous, and so make the same contribution to the truth-conditions of sentences containing them. In other words, the proposition expressed by any sentence of the form ‘ $\Phi(\text{kike})$ ’ is identical to that expressed by the corresponding sentence of the form ‘ $\Phi(\text{Jew})$ ’.

While the fundamental claim for the identity thesis that Jews are kikes *sounds* anti-semitic, it need not be *actually* anti-semitic. The identity thesis is usually bolstered with the further claim that the pejorative aspect of “kike” and other such terms is located elsewhere than in truth-conditional content, so what makes “kike” a bad word is a non-truth-conditional association with anti-semitism that is not shared with the word “Jew”. The exact nature and location of the moral content of pejoratives is a matter of some dispute among identity theorists. Regardless, the thesis holds particular attraction for those who reject a straightforward moral realism for racial pejoratives, including conventionalists, dispositionlists and non-cognitivists of various stripes.¹ But whatever the intuitive appeal of the identity theory for those persuaded by such views, it is nevertheless inconsistent.

2 The Frege Puzzle for the Identity Thesis

Because the identity thesis holds that Jews are identical to kikes, a standard version of Frege’s puzzle arises: why is “Jews = Jews” trivial and cognitively uninformative, while “Jews = kikes” is non-trivial and cognitively informative?

¹For example, see Kaplan (1999), Copp (2001), Hornsby (2001), Potts (2005), Williamson (2009), Lepore and Anderson (forthcoming).

There is an equally standard response open to the identity theorist, by appeal to modes of presentation, expressing ways of thinking. When an agent A thinks of Jews in an anti-semitic way, she employs one mode. Call it m^K . When A thinks of Jews in a non-anti-semitic way, she employs another mode. Call it m^J . The words “kike” and “Jew” are the outward reflections of these ways of thinking, even though they do not differ in the propositional content with which they are associated.

The sentence “ A believes that Jews are Jews” is analyzed as:²

$$(\exists M)\text{BEL} (A, \langle \text{Jew}^*, \text{Jew}^*, =^* \rangle, \langle m^J, m^J, m^= \rangle)$$

(where M ranges over n -tuples of modes), which is true, trivially. There is a mode under which A conceives of the identity proposition that makes it true trivially for her. Under examination of the component parts for the mode of the proposition, we see clearly that the same mode is being employed twice for objects flanking the identity relation, and since the same mode must pick out the same object, the identity follows trivially.

On the other hand, the sentence “ A believes that Jews are kikes” is analyzed as:

$$(\exists M)\text{BEL} (A, \langle \text{Jew}^*, \text{Jew}^*, =^* \rangle, \langle m^J, m^K, m^= \rangle)$$

which may or may not be true, but whichever, it is certainly not true trivially. cursory examination of the component parts for the mode of the proposition reveals that different component modes are employed for the objects flanking the identity relation, allowing that different objects may be represented. It’s conceivable that “kike” stands for a mode of presentation of something other than Jews. Thus, the identity does not follow trivially.

The identity theorist will rest her case, and conclude that the difference in cognitive significance between the identity statements has been successfully explained through the appeal to racist/non-racist ways of thinking encapsulated as modes of presentation.

3 The Initial Inconsistency

Upon cross-examination, it turns out that the identity theorist’s solution to the Frege Puzzle is by no means straightforward.

Consider a rational agent B who is herself an identity theorist about racial epithets. B is competent with the terms “Jew” and “kike”, and because she is an identity theorist, she is committed to the terms picking out the same extension. By Leibniz’s law (the indiscernibility of identicals), if Jews simply are kikes, then any property instantiated by one is instantiated by the other. Since B is both rational and competent, B must know that there is no property that differentiates Jews from kikes. Since there is no property that distinguishes Jews from kikes, whatever B thinks of Jews, B thinks of kikes, and vice versa.

²We follow here the analysis of beliefs reports in Salmon (1986) and Schiffer (2006)

But if so, there cannot be two distinct modes under which B thinks of Jews and kikes. On the identity theory, to think negatively of kikes is to think negatively of Jews: since Jews *are* kikes, to present Jews *as* kikes is to think of Jews in a negative and hateful way. But this is the mindset of the anti-semite, and of course this way of thinking about Jews is *rejected* by the identity theorist. The proposed solution to Frege's Puzzle would require B to be *schizophrenic* (if rational) with regard to her attitudes towards Jews. One simply cannot be both neutral and hateful, knowingly toward the same thing, at the same time.

By this reasoning, a person who does have distinct modes of presentation, that is, someone who thinks of Jews as Jews and Jews as kikes, will deny the identity thesis. For them, the Jews are not the same as the kikes; rather some are, and some are not. But the Frege puzzle that faces the identity theory is not a just a puzzle for racists of this sort; it is a puzzle for *everyone*. Knowing that Jews are kikes is non-trivially *a posteriori*, but knowing that Jews are Jews is trivially *a priori* does not depend on anyone's particular frame of mind. This precludes the identity theorist's appeal to distinct modes of presentation in order to solve the Frege Puzzle, since what holds of B holds of anyone who is rational and competent.

4 Potential Recourse for Identity

There are two plausible routes for rejoinder on behalf of the identity theory: appeal to modes of presentation as non-cognitive emotions or appeal to modes of presentation as metalinguistic markers. This would be to interpret ways of thinking as either ways of feeling or ways of talking. We articulate each one, but in the end, neither is satisfactory.

On the first proposal, the mode associated with "kike" reflects non-cognitive, anti-semitic sentiments towards Jews. To think of Jews under m^K , is to think of Jews with anti-semitic feeling. To think of Jews under m^J is to think of Jews without anti-semitic feeling. Because feelings are non-cognitive, these modes do not encounter the initial inconsistency with Leibniz's Law.

Generally, however, non-racists do not have racist sentiments toward their targets. So on this construal of modes of presentation, the identity theorist, B , would have to either sympathize with, or imagine being, the anti-semite in order to have distinct modes. Notice that neither sympathetic nor imaginary sentiment is identical to the very sentiment. For example, sympathetic or imaginary mourning for a loved one isn't the same as actually mourning a loved one. What is it to be in such a sympathetic relation? It is to *simulate* the emotional states of the racist. Thus, for B to think of Jews under m^K is to think of Jews as the target of simulated racist sentiments. The move is intended to inoculate the non-racist, as these racist sentiments would be merely simulated, and not experienced first hand.

The proposal runs into problems when we consider more carefully the relation of simulation. In order to simulate the emotional states of the racist, B must be in a cognitive relation to the racist; i.e. B must have certain beliefs

about the emotional responses of the racist to Jews. This is, however, to give up on the non-cognitive aspect to the analysis that was put forward as a way around the initial inconsistency generated by Leibniz’s Law. But once the rationality constraint is reinstated, the non-racist who is in the sympathetic relation to the racist is in trouble. In order to simulate anti-semitism, B must know that Jews are *actually* the targets of anti-semitism – i.e. by the actual racist who is the object of her simulation. Thus, by Leibniz’s Law, when B thinks of Jews as the targets of *simulated* anti-semitism, B is rationally committed to thinking of Jews as the *actual* targets of anti-semitism. But then this is simply to think of Jews as the anti-semitite does (!). The proposal is dangerously close to requiring speakers be racist in order to understand the cognitive significance of racist language, forcing B into a schizophrenic state of mind.

On the second proposal, modes of presentation are merely ways of thinking of objects that reflect a competent speaker’s knowledge of linguistic conventions. On this proposal, thinking of Jews under the mode of presentation m^K is just to think of them as being called by the word “kike”, and thinking of Jews under the mode of presentation m^J is just to think of them as being called by the word “Jew”. These are distinct modes without reflecting any moral or emotional content, thus bypassing the instability of conjoining the identity thesis, rationality, competence, and non-racism. Critically, metalinguistic modes are conceptual devices that the non-racist can share with the racist, without taking on the racist’s ideology or sentiments.

The problem is that while initially tempting, the metalinguistic proposal does not account for the Frege puzzle under all types of propositional attitude reports. Consider the following example:

Max disbelieves that Jews are kikes.

which is analyzed as:

$$(\exists M)\text{DISBEL}(\text{Max}, \langle \text{Jew}^*, \text{Jew}^*, =^* \rangle, \langle m^J, m^K, m^= \rangle)$$

What mode of presentation is available for this account? Assuming that Max is rational, he knows *a priori* that Jews are Jews, so there must be distinct modes available to him. Assuming that Max is a competent speaker, he knows that Jews are called “kikes”, so that linguistic fact cannot be what is in disbelief for him. Max is clearly disbelieving something like that Jews ought to be called “kikes”, or that Jews ought to be treated as Jews in the anti-semitic ways that are associated with the word “kike”. But without a normative component to the modes of presentation, this straightforward reading is unavailable.³

To further articulate the point, consider the view that there could be a normatively-loaded, metalinguistic mode for “kike”. Perhaps m^K is a way of thinking of Jews such that they deserve the anti-semitic treatment commonly associated with the word “kike”. Such a proposal might stretch the psychological facts, but is clearly both normative and metalinguistic, and allows Max to

³Worse still, the problem generalizes to other attitude reports such as: doubting, wondering, questioning, and regretting.

coherently disbelieve that Jews are kikes, while also believing that very proposition, i.e. that Jews are Jews, at the same time. Unfortunately, the problem for this move is that it runs squarely into the initial inconsistency set forth in the previous section, namely, that such agents who appear to be denying anti-semitism must actually embrace an anti-semitic way of thinking of Jews (i.e. *as deserving of anti-semitism*), and the resulting state for the agent, if rational, is schizophrenic.

5 The Extended Inconsistency

The situation, it turns out, is even worse for the identity theorist. Without distinct modes of presentation, there seems to be no non-racist way of expressing the identity that Jews are kikes. Accordingly, the identity theorist is committed to linguistic silentism; i.e. they would never make utterances that present Jews in the *kike*-way. Because identity theorists are also committed to rejecting racism, they further adopt a form of *cognitive* silentism. They would not think of Jews in the anti-semitic way; this way of thinking is rejected from their conceptual repertoire. But then the identity theorist is in no position to even formulate the identity theory, for they cannot form the thought that would express the theory. They cannot think the thought that Jews and kikes are one and the same, no less express it.⁴

It might be objected that we have confused use and mention; to formulate the theory we need only mention the anti-semitic way of thinking, and this does not imply a commitment on the part of the theorist to actually thinking it. But this hardly helps. One may choose not to express what one knows, one might even be compelled not to, but this is not the circumstance of the identity theorist. She has expunged the very thought; she has no way of thinking that is represented by the word “kike”. But you cannot mention a thought that you cannot even form in the first-place, no less use.

Still, would it not be sufficient for stating the theory to claim that a non-racist can know what it is to have racist thoughts, even though she does not harbor them herself? Surely someone who is not anti-semitic could know what it would be to present Jews in a hateful way, even though they do not themselves.⁵ Cognitive silentism shouldn’t amount to cognitive nihilism. But then we could formulate the thought that Jews are kikes, with the proviso that only one of the terms reflects an egocentric way of thinking of Jews.

Even if we think we can, we would hardly be any better off. We could only say that there are those who think that Jews are Jews, and those who think that kikes are kikes, but there would still be no way to say that there are those who think that Jews are kikes. This could be expressed only by someone in the

⁴An anti-semitic identity theorist would be no better off. They too are committed to cognitive silentism, but they expunge thinking non-negatively of Jews.

⁵Note that we would not want to say that this reflects recollection of long-distant memories of thoughts no longer held. This would have the disturbing consequence that to form the identity theory one would had to have been an anti-semite at some point. This is no progress.

state of mind of the anti-semite who *rejects* the identity theory. Since only *they* have the distinct ways of egocentrically thinking of Jews, only they could have distinct modes of presentation. Without distinct modes, there is no non-racist way of expressing the identity; and hence, no non-racist way of expressing the identity theory.

6 What is to be Done?

The inconsistencies we have derived in the identity theory arise from its interaction with the rationality and competence of speakers, and the rejection of anti-semitism. The natural reaction to this, in way of resolving the problems for identity theory, is to reject some one or more of these assumptions on which the derivation depends, and so circumvent the conclusion. Of these, the last, the rejection of anti-semitism, is of course non-negotiable, and since we take our colleagues to be as logically acute as they are morally upstanding, we would hardly question that any of them would fail to recognize the elementary consequences of Leibniz's Law. Left to deny is that speakers know the meanings of pejorative words, but this hardly seems an attractive strategy, as it is a core entailment of the identity theory that it is part of our linguistic competence that pejoratives are co-extensive with their neutral counterparts.

This leaves only one possibility. Reject the identity theory. This is our recommendation. Doing so has a tonic effect. If we reject the identity theory, there is no Frege Puzzle. If Jews are not kikes, (because no one is), then "Jews = kikes" is false, although "Jews = Jews" is true, indeed analytically so. One can certainly believe that Jews are kikes - this would be a belief typically of an anti-semite - but one would believe something false, and hence it would be to believe a different proposition than that Jews are Jews. The latter is a belief held indifferently by anti-semites and non-anti-semites alike.⁶

References

- [1] Almog, J. and Leonardi, P., eds., 2009, *The Philosophy of David Kaplan*, Oxford: Oxford University Press.
- [2] Copp, D., 2001, 'Realist-Expressivism: A Neglected Opinion for Moral Realism', *Social Philosophy and Policy*, 18: 1-43.

⁶Note that rejecting the identity thesis does not imply rejecting that sentences containing pejoratives have truth-conditions. If we do reject that they have truth-conditions - as recommended by expressivists and other non-cognitivists - then of course there will be no Frege puzzles, although there will be analogues; cf. Schroeder (2008). An alternative, suggested by Anderson and Lepore (forthcoming) is that our competence of the word 'kike' consists in knowing that it is a bad way of speaking of Jews, and that it would violate social conventions to utter a word so associated with such vile attitudes. To attribute the belief that kikes are Semites would then be make an utterance that breaches these social taboos, and having said this, according to Anderson and Lepore, there is nothing more to say about the use of slurs.

- [3] French, P., Wettstein, H., eds., 2001, *Midwest Studies in Philosophy, Volume 25, Figurative Language*, Oxford: Blackwell Publishers.
- [4] Hornsby, J., 2001, 'Meaning and Uselessness: How to Think About Derogatory Words', in French, et al. [2001], 128-141.
- [5] Kaplan, D., 1999, draft, 'What is Meaning? - Explorations in the theory of Meaning as Use', Brief Version - Draft no. 1.
- [6] Lepore, E. and Anderson, L., forthcoming, 'Slurring Words', *Noûs*.
- [7] Potts, C., 2005, *The Logic of Conventional Implicatures*, Oxford: Oxford University Press.
- [8] Salmon, N., 1986, *Frege's Puzzle*, Ridgeview Publishing Company.
- [9] Schiffer, S., 2006, 'A Problem for a Direct-Reference Theory of Belief Reports', *Noûs*, 40 (2):361-368.
- [10] Schroeder, M., 2008, *Being For: Evaluating the Semantic Program of Expressivism*, Oxford: Oxford University Press.
- [11] Williamson, T., 2009, 'Reference, Inference, and the Semantics of Pejoratives', in Almog, et al. [2009], 137-158.