

# The deceptive use of words

Alan Sokal

*Department of Mathematics*  
*University College London*  
*Gower Street*  
*London WC1E 6BT*  
*UNITED KINGDOM*  
`sokal@math.ucl.ac.uk`

*Department of Physics*  
*New York University*  
*726 Broadway*  
*New York, NY 10003*  
*USA*  
`sokal@nyu.edu`

25 March 2024

Published in *The Critic*, 15 April 2024

Some time back I wrote a short article about a disingenuous linguistic gambit that is sometimes used in social and political argumentation: employing a word in a private technical sense while allowing one’s readers to interpret the word in its ordinary English-language sense along with all its usual descriptive and moral connotations. Here I’d like to look in a bit more detail at the legitimate and illegitimate uses of redefinitions of words, taking two socially salient examples: “woman” and “racism”.

---

From the Middle Ages until a few years ago, English-speakers have universally understood the word “woman” to mean “adult human female”; and even today nearly all English-speakers still understand the word in that way. In order to avoid confusion — precisely the type of confusion that I want to denounce — let’s call the word with that meaning “woman<sub>OLD</sub>”:

woman<sub>OLD</sub>, *n.* Adult human female.

But advocates of gender-identity ideology have given the word “woman” a new meaning, as referring to self-described gender identity (let’s leave aside for a moment the question of what precisely that means) instead of biological sex.<sup>1</sup> Let’s call that meaning “woman<sub>NEW</sub>”:

woman<sub>NEW</sub>, *n.* Anyone who identifies herself/himself/themself as a woman.

But what is the meaning of the final word “woman” in this definition? If it means “woman<sub>NEW</sub>”, then the definition is circular:

woman<sub>NEW</sub>, *n.* Anyone who identifies herself/himself/themself as a woman<sub>NEW</sub>.

That circular definition isn’t *ipso facto* total nonsense: one can perfectly well make a circular definition like

shplunge, *n.* Anyone who identifies herself/himself/themself as a shplunge.

— and this would pick out unambiguously all those people, and only those people, who call themselves a shplunge. The trouble, of course, is that unless we have some *independent* idea of what “shplunge” means, we have no idea what a person *means* when they call themselves a shplunge. In the same way, if the final word “woman” in the definition of woman<sub>NEW</sub> means “woman<sub>NEW</sub>”, then we have no idea what people *mean* by that self-identification.

But of course people *do* have at least a vague idea of what they mean when they identify themselves as a “woman”, and that is because the word, unlike “shplunge”, has a pre-existing meaning: namely woman<sub>OLD</sub>. So what advocates of gender-identity ideology must mean by their definition is:

---

<sup>1</sup>Curiously, they are often coy about this. For instance, the lobby group Stonewall provides an extensive glossary of LGBTQ+ terms, from “abro” to “undetectable”, but the word “woman” is nowhere to be found on it, despite the central role played in Stonewall’s ideology by the assertion that “trans women are women”.

woman<sub>NEW</sub>, *n.* Anyone who identifies herself/himself/themself as a woman<sub>OLD</sub>.

— and this, whether or not the person actually *is* a woman<sub>OLD</sub>. That’s a perfectly legitimate definition, *provided that* one takes care to avoid confusing it with woman<sub>OLD</sub>. And that is also the sense in which the gender-identity mantra, “transwomen are women”, is understandable and true: transwomen (that is, biological males who identify themselves as women) are indeed women<sub>NEW</sub>.

But this raises the question of what it can mean to “identify as” belonging to a group to which one does not in fact belong.

Let’s start with a weaker type of identification, namely “identify with”. It’s easy to come up with legitimate examples of this. A white person who has lived all his adult life in an African-American neighborhood and has tirelessly campaigned alongside African-Americans for racial justice can legitimately *identify with* African-Americans, even while fully recognizing himself as white. Similarly, a man who holds radical-feminist ideas and whose closest friends are women can legitimately *identify with* women, even while fully recognizing himself as a man. Or more mundanely, a die-hard Chelsea supporter can certainly *identify with* the team, even if he is himself perfectly inept at kicking footballs.

“Identify as” is trickier, but there are still some perfectly legitimate meanings. A 60-year-old may say “I feel 40”, and a 14-year-old may say “I feel 20”. Moreover, these subjective assertions may in some cases have an objective factual basis: the 60-year-old’s physical fitness and health may indeed be as good as, or even better than, that of the average 40-year-old; and the 14-year-old may be as emotionally and intellectually mature as the average 20-year-old. But even when that is the case, we still base legal rights — such as the right to vote or to receive retirement benefits — on one’s actual age, not on what age one identifies with or as.

Philosopher Alex Byrne, in his recent book *Trouble with Gender*, gives some further examples of what it could mean to “identify as” a member of a group to which one does not in fact belong. He starts with the weaker “identify with”:

Someone — Gordon, to give him a name — might have a strong sense of kinship with Canadians. ‘I *feel* Canadian, these are *my* people’, Gordon says. He loves Canadian friendliness and humility, not to mention the Canadian Rockies and poutine. Among Canadians is where Gordon belongs, they are his in-group ... Regimenting language slightly, we can put this by saying that Gordon identifies *with* Canadians. (174)

But, Byrne goes on,

In social contexts, Gordon might encourage others to classify him as Canadian. Perhaps he pronounces ‘about’ *aboot*, celebrates Canada Day, drinks Moosehead beer, and so on. Perhaps he affects a Canadian style of dress — a parka and woollen hat in the winter. An observer would take Gordon to be Canadian, or at least to be making the effort. ... Again regimenting language slightly, we can put this by saying that Gordon identifies *as* Canadian. (174)

These two self-identifications, Byrne points out, are distinct from actually *being* Canadian (e.g. by legal nationality) or *believing oneself to be* Canadian; and the four distinct

concepts can in fact be mixed and matched in all sixteen possible ways. (Exercise for the reader: Give examples of each. Byrne does some of them: for instance, a Canadian who doesn't believe he is because his parents deceived him into believing that the family is American. He may or may not identify with, or as, Canadian. That gives four of the sixteen.)

Perhaps — as philosopher Kathleen Stock also suggests — the nebulous concept of “gender identity” can best be understood as a similar kind of “identifying as”. There are, after all, clear statistical differences, in many aspects of behavior, between men and women. The most obvious and radical of these is sexual attraction: approximately 96% of men are primarily sexually attracted to women, while approximately 97% of women are primarily sexually attracted to men. Another clear difference is physical aggression: worldwide, men commit the vast majority of homicides. So some people whose interests, predilections and behaviors are more typical of the opposite sex — or alternatively, match stereotypes of the opposite sex, whether or not they are statistically valid — in a strong enough way may feel impelled not only to “identify with” the opposite sex, but even to “identify as” a member of it.<sup>2</sup> On this reasoning, it can make sense for a biological male to “identify as” an adult human female, and hence to be — since the definition is purely subjective — a woman<sub>NEW</sub>.

But here is the catch: despite the superficial linguistic similarity, woman<sub>NEW</sub> is a quite distinct concept from woman<sub>OLD</sub>, and it is important not to conflate them. Members of the class woman<sub>NEW</sub> do not automatically acquire all the legal and social rights (or obligations) that attach to members of the class woman<sub>OLD</sub> — for instance, the right to compete in women's sports or to use single-sex changing rooms reserved for women. Rather, those legal and social issues need to be debated case by case. The gender-identity ideologues' redefinition of “woman” is an attempt to circumvent that debate by linguistic fiat: inducing people to confuse woman<sub>NEW</sub> with woman<sub>OLD</sub> by using the same word for both.

---

Before discussing the word “racism”, we need first to clarify the meaning of “race”. Some people take it to be an objective biological classification based on geographical ancestry; others take it to be a socially constructed and historically and culturally variable classification; and still others take it to be (most plausibly in my view) an amalgam of the two, namely, a socially constructed and historically and culturally variable classification that is founded partly, but not entirely, on objective biological facts about geographical ancestry.

But however you understand the word “race”, the meaning of “racism” has traditionally been clear: for instance, the Oxford English Dictionary defines “racism” as

---

<sup>2</sup>Note, however, one grave danger in this type of reasoning: precisely because sexual attraction is strongly correlated with biological sex, teenagers who feel same-sex attraction may decide to “identify as” members of the opposite sex — and even undergo hormone treatments and surgery — before realizing that they can be happy as gay men or lesbians. Indeed, the same is true for teenagers who are gender-nonconforming in other ways. Gender-identity ideology — and its medical correlate, “gender-affirming care” — can constitute a worrisome type of “conversion therapy” for gay and gender-nonconforming youth.

Prejudice, antagonism, or discrimination by an individual, institution, or society, against a person or people on the basis of their nationality or (now usually) their membership of a particular racial or ethnic group, typically one that is a minority or marginalized.

Also: beliefs that members of a particular racial or ethnic group possess innate characteristics or qualities, or that some racial or ethnic groups are superior to others; an ideology based on such beliefs.

Let's call those meanings "racism<sub>OLD</sub>". And of course, a racist<sub>OLD</sub> is a person who engages in racism<sub>OLD</sub>.

Recently, however, there have been proposals to redefine "racism" and "racist", expanding their ambit in some ways and contracting it in others. Both the expansion and the contraction can be seen in this definition, which formed part of a mandatory training program for students at the University of Delaware:

**A RACIST:** A racist is one who is both privileged and socialized on the basis of race by a white supremacist (racist) system. The term applies to all white people (i.e., people of European descent) living in the United States, regardless of class, gender, religion, culture or sexuality. By this definition, people of color cannot be racists, because as peoples within the U.S. system, they do not have the power to back up their prejudices, hostilities or acts of discrimination.

This definition — let's call it "racist<sub>NEW</sub>" — contains quite a bit to unpack.

The idea underlying the expansion — that all white people living in the United States are racists — seems to be that, whether or not one personally holds racist beliefs or engages in acts of racist discrimination, anyone who profits from a system of white supremacy is complicit in that system. That's a tenable view, but it elides several important questions:

1) To what extent is passively benefiting from an unjust system morally culpable? One can certainly argue that it is; but it is undoubtedly *less* morally culpable than actively engaging in racist discrimination. It is therefore wrong to conflate racism<sub>NEW</sub> with racism<sub>OLD</sub>; they raise very different questions and give rise to very different moral judgments. Furthermore, there are many forms of social injustice from which one could passively benefit, not only those connected with race: for instance, the accident of being born in a rich country rather than a poor one. One needs an analysis that handles all these situations in a consistent way.

2) To what extent *is there* a system of white supremacy in the US (or the UK) today? That is an empirical question, and one that needs to be answered by careful analysis of data, not by mere assertion. Concerning one key aspect — the large black-white income gap in the US — see here for a careful recent study, and here for a brief summary.

The contraction of the concept of racism is expressed concisely in this same document's redefinition of that word:

**RACISM.** Racism is race prejudice plus power.

Since people of color do not have power, they cannot be racists.

But even if one accepts this redefinition of the concept of “racism” — let’s call it “racism<sub>NEW</sub>” — is the premise of the claim true? In the United States 75 years ago it certainly was: few if any Black people, even the wealthiest ones, held any significant social power over white people. But today that is far from the case: many Black Americans (and Britons) hold significant social, economic or political power over people of all races, even if they are proportionately less numerous than their white counterparts. When people of color use their power “to back up their prejudices, hostilities or acts of discrimination”, that is indeed racism even according to the definition racism<sub>NEW</sub>.

So the quoted “definition” of racist<sub>NEW</sub> is not, in reality, purely a definition; it also contains explicit or implicit, and highly debatable, assertions of alleged fact. It is illegitimate to smuggle in assertions of fact as part of a “definition”; rather, like all assertions of fact, they need to be substantiated by evidence, after which they can be debated.

So let’s strip off those assertions of fact to obtain the pure definition:

racist<sub>NEW</sub>, *n.* One who is both privileged and socialized on the basis of race.

That’s a perfectly legitimate concept, but one must take care to avoid confusing it with racist<sub>OLD</sub>. The terms “racism” and “racist” in the sense of racism<sub>OLD</sub> have descriptive and moral connotations that do not necessarily carry over to racism<sub>NEW</sub> and racist<sub>NEW</sub>; rather, those issues need to be analyzed and debated case-by-case. But the likely effect of using the same word for both concepts is to induce people to conflate the two concepts and thereby to circumvent that analysis and debate — and in particular to induce people to attach the negative moral and emotional connotations of racism<sub>OLD</sub> to the new concept racist<sub>NEW</sub>, without conscious reflection or appropriate evaluation. Indeed, that is presumably the intention of those who proposed the redefinition.

---

**Postscript.** As Paul Connuck has kindly pointed out in comments on my earlier article,

the rhetorical sleight of hand that Professor Sokal identifies has been investigated at some length by rhetoricians, argument theorists, and philosophers of language as the deceptive use of *persuasive definitions*. See Charles Leslie Stevenson. (1938) “Persuasive Definitions”. *Mind*, Vol. 47, No. 187, pp. 331–350.

Dr. Connuck went on to explain:

Persuasive definitions are engineered to vary in meaning from the contemporary dictionary (lexical) definition and common understanding of the terms they define, while retaining those terms’ positive (or negative) “emotive meanings” (i.e., emotional evocations and associations). This altering of the conventional definition may consist of broadening it to include referents not ordinarily within its compass, narrowing it by applying new conditions (e.g., defining “racism” as requiring the existence of a hierarchical or other unequal distribution of power), or even by standing the definition on its head (e.g., “War is peace” in Orwellian Newspeak).

The deceptive effects of persuasive definitions come into play when the rhetor uses a term without disclosing that the meaning he or she ascribes to it *differs materially* from that which the term customarily conveys. The listener or reader may then unwittingly transfer to the redefined term the positive or negative valuations, attitudes, or associations (i.e., the emotive force) the term ordinarily evokes when understood *according to its conventional meaning*. Thus, by the “careful selection of terms with emotive meanings,” one may seek to bypass any need to provide persuasive support for a position — and avoid any deliberative examination of or challenge to that position — having already won over one’s audience by force of the definitional tweaking and use of axiologically and emotionally charged words designed to gain that audience’s automatic, unreflective acceptance. See Douglas Walton. “Deceptive arguments containing persuasive language and persuasive definitions.” (2005). *Argumentation* **19**:159–186 at p. 161.

The furtive use of such altered definitions thus has all the manipulative and deceptive elements of the classic “bait and switch,” and those suspected of engaging in this practice should be challenged to state specifically what they mean by the words they use.

*The author is Professor of Mathematics at University College London and Professor Emeritus of Physics at New York University.*