Arguing About Religious Identity and the No True Scotsman Fallacy

Robert Anderson, University of Notre Dame Australia, Australia

Abstract
Anthony Flew critiqued a particular argumentative manoeuvre he dubbed, "The No True Scotsman Move", where a speaker redefines an original claim by inserting the term "true" as an attributive adjective thereby restricting the extension of their first assertion. It is often appealed to in religious-apologetic diatribe. One non-academic book on fallacies names it "The No True Christian Fallacy", suggesting that those who commit this fallacy do so to illicitly defend a particular ideal religious identity. Often the charge of "No True Scotsman fallacy!" is invoked in strong eristic and sectarian contexts. Blamers score points by demonstrating that the opponent who commits this fallacy is evasive, prejudiced, and fails in their epistemic duty – since they refuse to accept falsifying evidence against their beliefs. In this paper I apply a heavy dose of the principle of charity and defend the individual who commits this fallacy and try to show they have something worthwhile to say. I critique the theory of the No True Scotsman Move in debates invoking religious identity. I argue that it is often mistaken to attribute the fallacy to others because of the presumption of a simplistic Aristotelian category theory of class membership. I favor a prototype theory of classification where the alleged committer of the fallacy is thinking about an ideal religious exemplar. If my argument succeeds I have defended this individual by showing that they were only trying to clarify what they originally meant by inserting "true".

Keywords: No True Scotsman move, manoeuvre, religious identity, Antony Flew, argument, fallacy

Introduction

What the philosopher Antony Flew dubbed, 'The No True Scotsman Move' (Flew, 1975) has grown exponentially, cited in many blogs, religious and atheist apologetic web-pages, and other forms of social media. It has become a social epidemic, one which I argue, is destructive to dialogue and reasonable tolerance. Antony Flew never called it a fallacy, nor do scholars working in the fields of philosophy or informal logic. Nevertheless, the 'move' has transformed into 'The No True Scotsman' (NTS) *fallacy* where arguers use it as a tool, if not, a weapon, to reproach their opponent – proving them wrong on the basis that they have committed this 'fallacy' and failing in their epistemic duty of being open-minded to consider objections to their cherished views. Most of the discussion surrounding the No True Scotsman Move, or fallacy is naive as are most explanations of logical fallacies on the web. It is time to put this alleged fallacy to bed perhaps keeping a more articulate and nuanced understanding of Flew's Move in certain contexts.

What I aim to do particularly in this presentation is to defend the Speaker who allegedly has committed the move/fallacy. When a speaker says to another, "you have committed a fallacy", the conversation usually stops, or at least moves in another direction. I defend the Speaker by applying a heavy dose of the principle of charity, a principle whereby the hearer should as far as possible, interpret their interlocutor as making a stronger case than a weaker one. Due to space limitations I am unable to look at particular *alleged* cases of the No True Scotsman Fallacy. There are also more important reasons why I am unable to do this since the proliferation of this particular fallacy on the web is so broad and variable that we are no longer able to grasp what the essential features of this fallacy are. To put it more bluntly and Socratically, I do not want to consider actual *cases* of the fallacy because until we know for sure what the essential features of the fallacy are, we will not be able to identify what is the NTS fallacy and what isn't. Rather, I give a formal explanation of the move/fallacy so that we can at least know when the fallacy obtains on the basis of the essential features, according to its original author.

Antony Flew's No True Scotsman Move

The basic dialogic structure of the NTS move goes like this, explained in terms of Flew's imaginary Scot. A person who self-identifies as Scottish utters the following statement after hearing about a terrible crime by an Englishman,

(U1). *No Scot would do such a thing.*

Another person then responds with a counterexample to this claim, along the lines of,

(Uh). *I know of so-and-so, a Scot who has done the exact same thing.*

The original speaker then utters the modified statement,

(U2). *No <u>true</u> Scot would do such a thing.*

For the sake of easy reference, I refer to the first and second utterance of the speaker, (the one who is charged of committing the fallacy or illicit manoeuvre), as **U1** and **U2**

respectively: utterance 1 and utterance 2. The response by the interlocutor is nominated as **Uh** – the hearer's utterance. This individual raises the counterexample to U1. To distinguish between Speaker and Hearer is a standard convention in the philosophy of language. In this instance, the Hearer also makes a claim, hence Uh.

U1 is considered a universal claim in the form of an **E**-proposition, (that is, an **E**-proposition according to the logical square of opposition) - *No S is P*. On Flew's understanding, the hearer who utters Uh directly points to a counterexample by way of an **I**-proposition, *Here is an S who does P*. Since **E** and **I** propositions are contradictory, one must be true and the other false. The default view is that the speaker has uttered a falsehood while the hearer has uttered a truth. Flew's narrative, in the voice of the hearer who points to the counterexample, rebukes the original speaker, since the *move* from U1 to U2 is a *redefinition* which evades falsification.

These concepts, *evasion* and *falsification* are important to Flew's analysis of the NTS move. Indeed, the chapter in Flew (1975) which introduces the move is named exactly that, "Evasion and Falsification." The background to this view was the popular perspective at the time in English speaking philosophy: verificationism in linguistic meaning, and falsificationism in science.

It is thought by Flew that the move from U1 to U2 transforms a fairly average, synthetic and contingent truth, (U1) into an analytic statement that is necessarily true, (U2). Given the analyticity of the new claim uttered by the speaker, it is a tautology, it "says nothing" and most importantly, no evidence can be bought against it in the same way that there can be no counterexample to the claim, *bachelors are unmarried males*. We search in vain, of course, to find a bachelor who is married.

To be sure, and as mentioned, very few, if any, serious academic texts call Flew's move a fallacy. Instead it is an illicit dialogic manoeuvre resulting in a Persuasive Definition, (PD). C.L. Stevenson first wrote on PDs in the 1930's and there has been much written since by philosophers and informal logicians since. Stevenson does not say that persuasive definitions are fallacious but they are illicit in some circumstances, (Stevenson, 1938). That is, persuasive definitions are not always to be avoided. On this point, I'll add in passing that if the academic literature argues that the NTS move is a form of persuasive definition, and if persuasive definitions are not always problematic, then it might follow with some *prima facie* assurance that not all NTSMs are problematic. But what is meant to be wrong with the illicit- kind of persuasive definition? The problem with these is that they disguise an *argument* beneath a definition.

Persuasive definitions are frequently taught in critical thinking courses and text books. For example, Trudy Govier's text explains a persuasive definition as, "… a stipulative definition disguised as a claim or as a reportive definition. In a persuasive definition there is an attempt to change attitudes by keeping the emotional connotations of a word while altering its applications". (Govier, 2009).

The comparison between Flew's NTS move and persuasive definitions is easy to see because of the insertion of the adjective *true* in the speaker's second utterance. Govier explains this where, "terms such as *real, true, authentic,* and *genuine* are often elements of persuasive definitions. If someone claims that modern art is not true

art because true art must depict objects realistically, he is using a premise based on a persuasive definition of "art."…But he offers no reasons to support that conception. Instead of reasons, he offers a disguised definition. Often when persuasive definitions are used, important issues are at stake, (Govier, 2009).

More technically, Stevenson and more recent commentators explain how persuasive definitions roughly work in terms of Frege's *sense and reference* distinction. In the more popular kinds of persuasive definitions, the redefinition keeps the *sense*, (or emotive meaning) of the term, but narrows the class of *reference*. In Flew's example the speaker's manoeuvre from U1 to U2 keeps the normal sense of the term "Scot" but narrows the extension to exclude the counterexample – the one who acts inappropriately, he who is not a Scot after all. This is called high-redefinition when the extension class is narrowed.

I note Govier's last statement where she says that when persuasive definitions are used, important issues are at stake and I view this as a means to ameliorate the disagreement between the speaker and the hearer. For indeed, in the alleged cases of NTS moves and fallacies, the noun that is modified by *true*, etc., is usually a social kind or human category like religion, nationality, political persuasion, or race. Here is another example of a persuasive definition, or at least what some have thought is a persuasive definition: if a person calls out another saying she cannot be a feminist and pro-life at the same time, the speaker is committing a persuasive definition since the term, 'feminist' has kept its emotive meaning, but the extension of the term is narrowed to include only, presumably, pro-choice feminists. That is, a feminist is one who, by definition, cannot be pro-life through a particular theory or ideology. There is no convincing reason why the persuasive definition is not committed in the opposite direction. Persuasive definitions do not take sides between ideologies, only sides of the speakers within an ideology who commit them.

It is not always clear whether there is even any argument in alleged cases of NTS moves, fallacies or persuasive definitions. But an illicit persuasive definition is meant to disguise an argument. It is not clear to this author that in Flew's example, or the feminist example there is any argument or dialogic reasoning taking place. One may be just expressing an opinion. Remembering Govier's suggestion that when persuasive definitions occur, important issues are at stake, I argue that it is far better for the participants in the dialogue to continue the conversation, or argument about just what exactly constitutes a particular case of the social kind under discussion instead of blaming the other for faulty reasoning. I have in mind conversations that carry forth among lay-people like, "oh, why do you think I can't be a feminist and pro-life at the same time?" Or, our original speaker answering back, "You think he is a Scot? Why do you think that? What is *your* definition? Wearing a kilt is not enough".

I make this observation to make the point then to leave it alone: that in abstracted explanations of NTS moves or fallacies, the hearer who points to a counterexample might just as well be begging the question in favor of his or her position as strongly as the one they are bringing the charge against. After all, who is to say, what is a feminist? What is a Scot? What is an Australian? It is not as if there are agreed and uncontroversial necessary and sufficient conditions for these social kinds. The matter may differ when it comes to religion since often soteriological or ecclesiastic

doctrines determine true membership within a religion. Religious identity is a much more complicated social kind, inviting in most circumstances an understanding of kinds determined by the divine.

My Argument

My aim is to defend the speaker against the blame of committing an illicit NTS move, NTS fallacy, or persuasive definition. My reasoning is as follows, where according to the NTS theory of Flew,

- **Evasion** is the fundamental crime committed by the speaker, (since the evasion sidesteps falsifying evidence).
- Evasion depends on **Redefinition**.
- The redefinition is committed by moving from U1 to U2, from a synthetic/contingent utterance to an analytic/necessary truth.
- Therefore, if it can be shown that the speaker makes no such redefinitional manoeuvre, or a less serious kind of move, the charge of evasion should collapse.

Whether the NTS move is illicit or not, fallacious or not, the fallacy is not structural but dynamic. That is, it is an informal fallacy (allegedly) where the speaker strategically manoeuvres to avoid losing "the argument". The fallacy is not to do with form or structure as in, for example, the formal fallacy of affirming the consequent. Instead everything that is meant to be wrong with the NTS is the new posture the speaker takes when redefining a term with *true*.

To resist the charge of redefinition is also to resist the charge of the more serious crime of evasion. The most simple way of doing this is to interpret the speaker more charitably where U1 and U2 amount to saying the same thing, or saying two things that are close enough semantically without the *synthetically-true* to *analytically-true* shift. I believe there are ways we can understand the speaker as not committing a redefinition:

1. Analyzing and evaluating the role of the attributive adjective *real, true, genuine,* etc., inserted in the speaker's second utterance, U2 No *true* X.
2. Analyzing and evaluating the head noun in the utterances. What is the philosophical *kind* of X, natural, artificial, or social?
3. Analyzing and evaluating the illocutionary intent of the speaker on the basis of the copula that completes the predication. Is the illocutionary intent as indicated by the copula verb indicative, subjunctive, or modal? In other words, is the speaker's utterance descriptive or normative?

I have touched on point (2) where it was previously said that most of the alleged cases of the NTSM or fallacy are about social kinds which are by nature intrinsically controversial. Due to space limitations I leave this factor in favor of discussing the other two key points. The strongest being the question of the role or purpose of the interpolated *true* in (1). This is where, I believe, the fundamental disagreement obtains between the speakers in the dialogue. For the hearer, who points demonstratively to the counterexample, understands the interpolated *true,* truth functionally. It is almost as if the hearer makes the speaker mean, *a true Scot is by*

definition, one who does not put sugar on his porridge where *true* strengthens the predicate to make it a necessary condition of being a Scot. In the parlance of philosophy of language theorists, this is a descriptivist account of meaning where *being a Scot* instantiates a collection of conditions.

If we can find another way to understand the interpolated *true* which does not transform a contingent statement into a necessary truth, this should give us points in favor of the speaker. Some recent scholarship understands terms like *true, real, genuine,* which are adjectives in the attributive position, as intersective adjectives. They cannot be converted to the predicative position as subsective adjectives can, (and thus should not be understood truth-functionally). An example of such an intersective adjective/noun combination is "white wine". To refer to "white wine" is not to also say, that "wine is white". To be clear, it might look as if this is possible but the predicate, "…is white" now speaks of the literal colour of wine, (which it is not, it's clear with a yellowish tint). This "whine is white" does not refer to the *kind* of wine – white wine over red wine. Putting this another way, the *white* in "white wine" is syncategorematic. It has no independent meaning apart from its association with the noun it prefixes. As *"white wine"* is a term that cannot be reduced to its component parts predicatively, I argue the same (in the right contexts) goes for *true Scot.* If "true" is intersective then we are unable to meaningfully reduce the sentence, "Angus is a true Scot" to "Angus is a Scot" and "Angus is true." The last statement does not make sense.

What then could *true Scot* refer to if it also is not an element in a truth-functional expression? I suggest the role of *true*, *genuine,* or *real,* amounts to an expression of a good example over a bad example in the speaker's mind. A *true Scot* is a *good Scot.* Understood this way, *true Scot* is an exemplar in the speaker's mind which suggests that there *can be other Scots* who are not good exemplars.

The direction this discussion takes is to posit that a proto-type theory of graded-membership is a better way of interpreting the speaker's utterances in alleged cases of NTS moves and fallacies. This is a controversial claim to which I am unable to devote the time in discussing the relative merits of classical category theory versus prototypes. It is clear however, that Flew and the hearer approach the speaker's utterances assuming a classical category theory where each member of a set must instantiate exactly the same necessary conditions as each other member. Moreover, one either passes or fails the necessary condition test; one is either a Scot or not a Scot. This seems a rather oversimplification, to say the least. The NTS move and fallacy is wholly parasitic on this assumption: that the categories in each of the speaker's mind are classically categorial where the predicates (puts sugar on their porridge, etc.), form a set of descriptions. Again, why should the speaker share this ontological commitment of classical categories and descriptions which form essential properties? Given the speakers in the dialogue tend to break off the conversation when the charge of NTS fallacy is raised, this is some evidence that they are not even on the same page, as far as logical commitments go.

The second point (2) above follows naturally from the previous point made about the referent of *true X.* As the speaker is referring to a good or virtuous example of the social kind in question, the mood of the copula-verb should be modal or descriptive and is better understood as *would, should*, or *ought* instead of the indicative *is.* This

will be explained by an analogous example by combining points (1) and (2) already made. When we think of something about our own national identity and hear an exclamation that, "Peter is un-Australian", what are we to think? The claim is to render Peter unworthy of Australian citizenship formally or informally understood. What might be surprising is that the locution, "Peter is un-Australian" only makes sense if Peter is, in some sense, Australian! In other words, both of these sub-statements can be true at the same time: (i). "Peter is Australian" and/but (ii). "Peter does not act as an Australian *should*" or, "Peter is Australian and un-Australian". Here *negation* does not contradict or provide a refutation of a universal generalization. It makes little sense in everyday language to claim that, for example, "Peter is un-Australian" when Peter is American.

Peter has to be both Australian and un-Australian at the same time, but in different senses. To return to the NTS move or fallacy, a better way to interpret the speaker along these lines might be, "OK though you pick him out as Scottish, he does not act as a Scot should; he is not a *good* Scot, a *true* Scot". Yes, this is a value judgment on the part of the speaker but the disagreement between speakers in alleged NTS moves or fallacies are exactly question of values. The literature on social kinds leans towards understanding social kinds in these evaluative ways. Francesco Guala writes that social or human kinds "seem to be dependent on human classificatory practices…unlike natural kinds, social kinds depend crucially on our attitudes towards them (Guala, 2014). Khalidi summaries three further positions as follows: John Searle's position that social kinds are, "ontologically subjective since they depend on human mental attitudes…"; Ian Hacking: social kinds "are interactive and can change in response to our attitudes towards them"; P.Griffiths: social kinds are, "fundamentally evaluative or normative in nature," (cited in Khalidi, 2013). Schiappa (2003) makes similar claims, all consistent with Trudy Govier's remark about persuasive definitions, that they occur when important issues are at stake.

Returning to my argument wherein I aim to justify the position that there is no real difference between the speakers first and second utterance. The syntactical move, if it can be considered as a move at all is merely to hedge or to precisify what the speaker had in mind originally. The second *true X* utterance need not be understood as a new claim where the speaker's original utterance U1 can be paraphrased as, a good Scot does not do such a thing.

Is there a way to understand the speaker and hearer as making statements that are not contradictory so as to defend the speaker against the charge of evading falsification through high-redefinition? I understand speaker's second claim, *No true S would P* not as a Universal Negative proposition commonly understood as an **E**-proposition, but as a Particular Negative **O**-proposition. The speaker seems to be saying, *"that* is not a true Scot – the one that you have pointed to, to contradict me." On the basis of the logical square of opposition with its **A, E, I,** and **O** propositions, if the interlocuter and speaker are making **I** and **O** propositions respectively then *both can be true* since they are sub-contraries of each other. Therefore, it is possible that *some Scots put sugar on their porridge* **and** *some Scots do not*, but those who do not are the truer, better Scot. Or in the parlance of the Prototype theorist Eleanor Rosch, *he is a real Scotty Scot.*

Arguing About Religious Identity: Essentialism versus Constructivism

Most of the above can be applied, *mutatis mutandis* to cases of disputes in disagreement about religious identity. I make the provisional observation that intra-religious disagreement about religious identity is often, at least in the Judeo-Christian traditions, a matter of doctrinal orthodoxy where one Anglican deems another as not being properly Christian, or Anglican because of a failure of belief. On the other hand, inter-religious disagreement about religious identity tends to be about orthopraxy, for what is observable, and what counts as important whether one is truly Christian, Moslem, Buddhist, or Atheist, is *what they do*. Hence, I speculate that inter-religious/secular disputes that invoke the NTS move or fallacy blame the other for bringing up an example of a believer, or atheist, who fails in their behavior of being good. For example, to consider that atheists can be good without God or revealed religion, and to point to Stalin as a counterexample is to show demonstratively that some atheists are not good.

The NTS move or fallacy in religious argument *depends* on essentialism about religion. This is because Flew's narration and the position of the hearer presume a Descriptive/truth-functional account of essential properties, (a true Scot is one who does not put sugar on his porridge). I do not argue against essentialism about religious identity *per se* but wish to avoid the further complication involving individuals with different perspectives of essentialism merely arguing from their own essentialist standpoint. Essentialism is a problem, but not *the* problem; disagreement about what is essential is the problem that motivates the argument. An over-simplified dichotomy is to pit essentialism against social constructivism but this results in a dilemma. As mentioned, the NTS move/fallacy obtains because of essentialist attitudes but the other horn of the dilemma may be even more undesirable for some, where constructivism leads to a slippery slope to the view that religious identity is just "in our heads" and can be created and taken away by just thinking about it. Outlined this way, essentialism is too rigid while contructivism too liberal. Is there a way of avoiding this dilemma?

Religious Exemplarism

Linda Zagzebski has constructed an exemplarist virtue theory (Zagzebski, 2010) using the general Kripke/Putnam account of direct reference of natural kind terms, (avoiding descriptivism). Ian James Kidd has also created approaches to understanding *religious exemplarism* in particular, (Kidd, 2016). In what is also known as the causal theory of reference, natural kind terms like "water" and "gold" refer to the same thing in each case when indexed by a demonstrative such as "that" just as long as there is a proper chain of communication between speakers. This chain of communication should reach back far enough to the original *baptism* of the object or kind, with the name. What is important for our discussion is that speakers do not need to know the proper descriptive conditions that refer to these objects. A speaker can use the term "water" to correctly refer to H_2O without knowing that water is H_2O. The nature of water is discovered empirically but once discovered is deemed a necessary truth that *water is H_2O*. Hence, we can have necessary/*a posteriori* truths.

What is interesting is that on the Kripke/Putnam account, as Zagzebski remarks, "we do not need to know the nature of the referent, and yet we know how to construct a definition that links up with its nature," (Zagzebski, 2010). This should seem familiar at least in terms of the speaker who allegedly commits the NTS move/fallacy. He knows that so-and-so is a Scot but does not necessarily know what it is that makes him one. It makes little difference to him, therefore, when a pesky observer raises a counter-example by way of another Scot who supposedly instantiates some property thought to be impossibly Scottish. The speaker is not defining a Scot by descriptive content but evaluating what is a good Scot.

Zagzebski constructs her moral theory along these Kripke/Putnam ideas where moral concepts "are anchored in exemplars of moral goodness, direct reference to which are foundational in the theory. Good persons are persons *like that,* just as gold is stuff *like that*. Picking out exemplars can fix the reference of the term, "good person" without the use of descriptive content." (Zagzebski, 2010). Zagzebski's theory goes a bit further than I require since it is a theory of moral concepts grounded in exemplars who are the "most imitable". They are most imitable "because they are most admirable" where admiration is the emotion learnt through the emotions of other individuals. For the purposes of going between the horns of the essentialist/constructivist dilemma, I believe that what I have already alluded to from Linda Zagzebski's moral exemplarist theory can be transformed as a way for disputants in religious arguments about identity to know *that* a particular person in an exemplar, good role model, without knowing *what* it is about them that makes them such. By identification with the speaker over religious identity, the proto-typical example is a religious exemplar in his or her mind. The outsider along the periphery is still religious but a bad example. This is to view the radius of a prototype-circle as a gradient of *good* X to *bad* X rather than a categorial exclusion of *either X* or *non-X*. For example, a good Catholic might be Mother Teresa, a bad Catholic might be a bad Pope from history. Yet they are still Catholic in the same way that robins are good examples of birds, and ostriches are not so good examples of birds yet there is no contention despite this difference that both robins and ostriches are considered birds by competent speakers.

An objection to my use of the Zagzebski/Kidd approach was made by Stephen E. Gregg, (in conversation, 5th July, 2017), where my configuration of the religious prototype/exemplar is still a case of essentialism. With this point, presumably the bias we have in society and religions towards the virtuous, the good, and the orthodox renders us intolerant to outliers and radicals who exemplify another set of conditions. This is a very good objection to which I can only respond now by stating that the prototypical religious exemplars we have are "voted in" by the societies and cultures in which religions are formed. In other words, my appeal is just *ad populum* – to the masses. I have no other answer yet beyond this since my appeal to a causal-historical account of direct reference presupposes a social *ad populum* appeal.

Conclusion

I have not provided any evidence for my claim that the uses and abuses of the No True Scotsman move, or fallacy, are divisive. Space does not permit me to demonstrate this but I urge the reader to casually search the world wide web for this "fallacy", especially when indexed with religious terms. It is used too often in the

unfortunate role resembling Lakoff's 'Argument is War' metaphor - thrown at others to win arguments about religion. The NTS always concerns, in some way, an individual's true identity. My argument is to resist the manoeuvre where the speaker commits a redefinition that evades falsification. If it can be understood in some way that the speaker does not commit a fallacy or illicit move, the argument can continue. The speakers may never agree due to their differing values but I consider it more worthwhile and conducive of mutual respect through difference, than a full-stop and a Socratic *puzzlement* where each speaker wonders just what happened when one is blamed of committing the No True Scotsman fallacy.

References

Flew, A. (1975). *Thinking about thinking.* New York: Harper Collins.

Govier, T. (2009). *A Practical study of argument.* Boston: Wadsworth.

Guala, F. (2014). On the nature of social kinds. In M. Gallotti, & J. Michaels (Eds.), *Perspectives on social ontology and social cognition* (pp. 57 − 68). Dordrecht: Springer.

Kidd, I. (2016). What is a religious exemplar? 4th Glasgow Philosophy of Religion Seminar, May 2016, University of Glasgow.

Schiappa, E. (2003). *Defining reality: definitions and the politics of meaning.* Carbondale: Southern Illinois Press.

Stevenson, C.L. (1938). Persuasive definitions. *Mind 47*, 331 − 350.

Zagzebski, L. (2010). Exemplarist Virtue Theory, *Metaphilosophy*, *41,* 41-57.

Contact email: robert.anderson@nd.edu.au