The Journal of Applied Linguistics and Applied Literature: Dynamics and Advances, Volume 6, Issue 1, Winter and Spring, 2018, pp. 7-12

JALDA's **Interview with Professor Glenn Fulcher**

Interview by Dr. Bahram Behin

Prepared for publication by: Dr. Davoud Amini



Glenn Fulcher

Glenn Fulcher is the Professor of *Applied Linguistics and Language Assessment* in the English Department at the University of Leicester, UK. He got his PhD in Applied Linguistics and Language Testing from Lancaster University (1993) and his MA in Applied Linguistics from the University of Birmingham (1987). Professor Fulcher has been the editor of Sage's *Language Testing* (2006-2015) and an influential member of the Executive Board of *The International Language Testing Association* (ILTA) for many years. He has extensive experience and expertise in the philosophy of assessment, test design, the development of data-based rating scales as well as teaching language assessment. Professor Fulcher's book *Language Testing and Assessment* co-authored by Davidson (2007) has been the main resource for the ELT masters' courses in language testing in Iran for many years. Among his other publications are *Re-examining Language Testing: A Philosophical and Social Inquiry* (2015, the winner of the 2016 SAGE/LTA Book Award), *The Rutledge Handbook of Language Testing* (2012), *Practical Language Testing* (2010), *Testing Second Language Speaking* (2003) , and *Writing in the English Language Classroom* (1997). In an online interview, Professor Glenn Fulcher has joined **Dr. Bahram Behin** who is a zealous adherent of Fulcher's philosophy of assessment and has presented language testing courses based on his books.

BB: ---- *Professor Fulcher, I appreciate your kindness in accepting my invitation to take part in this dialog.*

GF: ---- It's my pleasure to talk to you.

BB: ---- *I would like to start by touching upon one of your books which I have used as a recourse book for teaching the language testing course. In Fulcher and Davison (2007) you explore new directions in language testing. I think of this as a mutation rather than a continuation of traditional views of language testing. Do you think this is right?*

GF: ---- The first thing to say is that all work in language testing and assessment builds upon the great minds that have gone before us. I have been particularly influenced by Messick, as have many in the field. What always struck me as wonderful about his writing was the concern with the philosophical foundation of his conception of validity. His 1989 essay is probably one of the papers that I go back to most frequently as there is always something new to discover. So, our 2007 book was based on a long discussion between Fred Davidson and me about the nature of classical Pragmatism and what we could learn from that philosophical tradition for the practice of testing and assessment. It came about by chance really. I had visited ETS in Princeton and a friend there had recommended that I read Louis Menand's *The Metaphysical Club*, which had just been published at that time. A few years later when Fred visited me in Scotland, where I lived then, he saw it on my bookshelves, and that got us talking. You will see that our discussion of Pragmatism comes very early in the book on page 10 which is worth reading out here:

> "This takes us to the heart of epistemology and what it means to say that something is 'true' or 'real'. In 1877 C. S. Peirce had put forward a pragmatic notion of meaning: 'Consider what effects, that might conceivably have practical bearings, we conceive the object of our conception to have. Then, our conception of these effects is the whole of our conception of the object' (Peirce, 1877: 146). To translate this into modern English: if we believe something to be true, will the effect be that we are better able to understand the world around us and use the idea to do something practical in a way that results in progress? Or as Messick (1989: 26) puts it (using the term 'instrumentalist' for 'pragmatist'): 'According to the instrumentalist theory of truth, a statement is true if it is useful in directing inquiry or guiding action.'"

Many readers have taken this to mean that Fred and I are relativists, in the same sense that James reinterpreted classical Pragmatism to make it almost a point-of-view approach to the nature of truth. But that isn't what we wished to say at all. With hindsight, we needed to have been a lot clearer. We really wanted to assert something closer to Dewey's warranted assertion, where the acceptance of a standpoint enabled us to better "understand the world", but that this "truth" is contingent upon future research.

However, the answer to your question is really that we were driven by philosophical questions, and I think you can see that much more clearly in our writing since 2007.

BB: ---- *Ethics and the question of ethical concerns seem to have more importance in language testing than before. If that's the case, would you please comment on this?*

GF: ------ This is certainly true. In older models of validity there is really no concern at all with ethics or even test use. Again, Messick was a ground-breaker here, including social consequences in his model. But we do have to remember that adverse social consequences were only to be considered if they arose directly from validity issues with the test. Now we are concerned with the social consequences of test use even if all the evidence suggests that the test can reasonably be used for its intended purpose. Our view of the role of the language tester in ethical test use begins with the test design and development, and it is fundamentally pragmatic – in the philosophical rather than popular use of this term. It is best expressed on page 144 of our 2007 book where you read:

> "The task for the ethical language tester is to look into the future, to picture the effect the test is intended to have, and to structure the test development to achieve that effect. This is what we refer to as *effect-driven testing*."

If test developers state in advance what the effect of the test is intended to be – both on individuals and society – then we can evaluate whether or not it meets these goals once it is introduced. Drawing on excellent washback methodologies developed by researchers like Wall and Horak, we can collect baseline data from societies before tests are introduced and compare with what happens after the introduction of the test. We are all familiar with the notion of "unintended consequences", and this notion of effect-driven testing takes these side-effects much more seriously. Let me give you one very practical example of an unintended consequence – the IELTS test is intended for use in making entry decisions to higher education; but we know that in recent years it has been widely used for immigration purposes. In India there are many young men who would like to live and work in Australia, but they don't have the English language level to obtain a work permit. An industry has grown up around legal firms that match men from wealthy families with girls who have achieved the necessary IELTS score to enter Australia, but don't have the funds to study. A marriage is arranged so that the man's family pays for the girl's education, and she takes her husband with her as he can obtain a temporary work permit and has the opportunity to make it permanent while his "wife" is studying. This new industry exists only because of the test and its use, and it therefore seems reasonable to suggest that the testing agencies and the policy makers directly responsible for social engineering using language tests take some ethical responsibility for these unintended consequences.

BB:---- *There are studies nowadays that challenge the validity of large-scale tests such as TOEFL and IELTS for different reasons. These studies seem to me to be in line with your 2007 book. What would be your view of such an assumption?*

GF:---- Large-scale standardized language tests will always be questioned by independent researchers. This is a good thing. But we have to remember that a fundamental notion that we work with is *test purpose*. Both of these tests were designed to provide an indication of whether an individual could function in English medium higher education. They do not do this "perfectly" – in the sense that no test can be absolutely reliable. Errors will be made. If the test providers make available the error rates, then decision makers can decide whether they wish to maximise false positives or negatives. We also have to remember that admissions officers should use other sources of evidence in decision making – school history, references, personal statements, and so on. They should be just one piece of evidence in the decision-making mix.

Where the real problem lies is if the test is used for a purpose for which it was not intended, and no new validation evidence is provided for this new use. We could argue that immigration is such a purpose. It seems to me on a-priori content grounds that what is contained in the two tests you mention is not relevant to the survival needs of new immigrants in an English medium environment. So why are they used for this purpose? Fred Davidson and I have written about this particular problem and compared the situation to that of using buildings for new purposes. This often happens, but architects need to re-design the building to consider the health and safety issues of its new use, adding new features and removing others, and carrying out checks to ensure that it will be fit for purpose. We call this "test retrofit", which requires evolution and validation for new contexts. Your readers can download our article on this here: http://languagetesting.info/articles/store/Test%_20Architecture.pdf. Again, you can see that this fits in with our Pragmatic notion of consequences, and the need to provide warranted assertion for all test purposes.

BB: ---- *According to my personal experience both as a student and as a teacher, most language teachers and language students find mainstream language testing materials and terminology confusing and frustrating. One of the reasons is that they are confronted by methods that are derived from natural sciences. Treating language students and language learning within scientific statistical frameworks perhaps sounds like testing robots to see whether they behave according to their programming. But human beings are not robots and language learning is not a computer programme. I wonder what your reaction would be towards such reasoning. Could one rely on your concept of "context" in your 2007 book to defy language learning and language testing based on the modern sense of science?*

GF: ---- There is absolutely nothing wrong with the use of "scientific methods" in language testing or applied linguistics. I don't have much time for point-of-view epistemologies. If we assert that we all have different experiences and points of view, that we construct (or co-construct) reality in each new context or interaction,

then we throw generalisability out of the window. Why would I wish to conduct a research study if the results were only relevant to the context in which I conducted it? Knowledge would not be transferable or useful for anyone else, and it would be radically contingent to the local environment. If we are to do useful research we need to try and establish trends and patterns from which other people can learn, and from which predications to other contexts can be made. In fact, this is at the heart of the language testing endeavour. We want to make reasonable predictions from a test, which is an abstraction of real life, to what someone is likely to be able to do in real life. So, we must have methodologies that are consistent and compare like with like. And we must also have replication studies, which are sadly lacking in much of applied linguistics and language testing.

But this doesn't treat humans like robots. We share a common evolution, and our languages are structured in similar ways so that they are learnable, as Chomsky has shown us. So, we have much in common that we can study and describe. This does not deny cultural differences, or individual variability. But when we employ someone to work as an air traffic controller who must be able to communicate in English with pilots and others who use English as a second language, what we are interested in is whether they have the language abilities and skills to do their job in a way that keeps us all safe when we are travelling. If individual or cultural differences get in the way, this is noise.

So, the answer to your question is that the scientific method – to the extent that it can be used in social science research – is essential to our tasks. But this does not mean we ignore the humanity of the test takers or any of the other stakeholders. After all, that is why we are so interested in ethics and fairness!



BB: ---- *As my last question, I would like to see how you characterise the traditional dichotomy of "objectivity" and "subjectivity" in scientific discussions?*

GF: ---- An interesting final question! And it very neatly brings us back to Pragmatism. On page 11 of our 2007 book we have this quotation from C. S. Peirce (1877) which reads:

> This great law is embodied in the conception of truth and reality. The opinion that is fated to be ultimately agreed to by all who investigate, is what we mean by the truth, and the object represented in this opinion is the real. That is the way I would explain reality.

One of my colleagues here at the University says that I'm very fond of quoting "dusty Victorian writers"! Indeed so! Combine this quotation with Dewey's notion of warranted assertion, and what we get is the view that at any point in time we have evidence that supports theory, which attempts to explain the world in ways that allows us to make predictions. We can test these predictions. So, we have an empirical approach to finding out "the truth". But is this objective? Well, no, because the data could always support more than one theory. 2000 years ago, Lucretius was quite prepared to entertain two theories on the basis of the same evidence: (a) The sun disappears at a certain time of day, and reappears at the start of the next day; (b) The sun is extinguished at a certain time of the day, and re-ignites at the start of the next day. For us (a) – however simplistic – is clearly "truer" than (b). But given the evidence available to the Romans, both were plausible. But it didn't really matter – because both theories led to predictions THAT WORKED for the purposes of their daily lives. Of course, as time passed one theory became established as the better explanation and this led to better predictions of all kinds of things. We therefore use warranted assertion, *given our current state of knowledge*. What is truth? Peirce states that ultimate truth is what we would all ultimately agree upon, but we will never reach that point because all human knowledge is contingent.

So, to answer your question, we do not abandon the notion of objective truth, but we are not so arrogant as to believe that we objectively know the truth. As humans this is fundamentally satisfying, as we are on a question to better understand our world, how we fit into it, and what the human condition is. And that includes our unique ability to use language to create and manage relationships and societies.

If your readers would like to know more about my thinking in this area, I can only recommend my latest book, which is a philosophical and social investigation of language testing and how it all fits in with my understanding of the world in which we live. They can find more information on my website, at this URL: http://languagetesting.info/RLT/home.html.

BB: ---- *I am sure that JALDA's readers will find the information helpful. Thank you for sharing your insights with our readers.*

GF: ---- My pleasure.