



# Language Teaching Research Quarterly

2022, Vol. 29, 6–19



## An Interview with Glenn Fulcher: Looking Back and Moving Forward

Lee McCallum<sup>1\*</sup>, Christine Coombe<sup>2</sup>

<sup>1</sup>Coventry University, UK

<sup>2</sup>Dubai Men's College, Higher Colleges of Technology, Dubai, UAE

*Received 12 March 2022*

*Accepted 17 April 2022*

The interview took place on 16<sup>th</sup> February 2022 at 11:00am (GMT) via Zoom. The interview focused on three central themes of Professor Fulcher's work in language assessment: (i) Reflecting on his career as an educator and scholar, (ii) Experiences and lessons learned when working with industry testing companies, and (iii) Future opportunities and challenges for the language assessment community. The interview was transcribed by Dr. Lee McCallum.

CC: Christine Coombe

GF: Glenn Fulcher

LM: Lee McCallum

### Reflecting on your career as an educator and scholar

**LM:** Can you tell us how you first became involved in language assessment?

**GF:** The first thing to say about this is when asked what do you want to do? at school or university, no one says 'oh I want to be a language tester'. This is unheard of, and in fact I think it's the same with Applied Linguistics, which probably accounts for why there are just so few Applied Linguistics courses at Undergraduate level. Just the Applied Linguistics first, I started out studying theology and philosophy and my major was 1<sup>st</sup> century Greek manuscripts. My career trajectory

was actually to join the Anglican priesthood. When I was an undergraduate, I got an exchange scholarship to Dartmouth College in New Hampshire, in the US. That was, dare I say way back in 1979. And I went to study philosophy. But while I was there, I had to make up extra credits, and I took courses in linguistics. When I got back to King's College London, I was working on the problem of Q. This has nothing to do with Q in Star Trek. This is Q as in Quelle, which is the unknown source that the authors Matthew and Luke were drawing on together, as well as Mark, in the evolution of the Greek manuscripts that we now know as the New Testament. And I actually applied what I'd learned in the US to that problem, the problem of Q. But eventually, I didn't follow a career in the church because when I went for the job interviews before training, I was told, "you have more questions than answers! And priests should be able to give people answers." And I think that has stuck with me throughout my career. So, it was very perceptive of them.

So, I packed my bags, and I went off to Cyprus and I got a job eventually in the English Institute there, and one of the things we had to do was predict the grades of our students. I could never get it right. And the principal of the school said to me, 'why can't you predict the students' grades properly when we give them to parents?' And I thought to myself, well if I can't do this, perhaps I'm using different criteria from the examination boards. And I started to investigate why I couldn't get it right. But then I thought, this needs research to answer the questions about which I'm curious. And that early stuff into teacher assessment compared with exam boards, was actually published in a book edited by Tim Caudery who was my predecessor as Director of Studies at the English Institute in Cyprus, and it was called '*The Role of Assessment by Teachers in Schools*', and that was back in 1991. So that was my very early involvement.

Whilst I was still in Cyprus, I also got involved in a campaign to get away from the old London GCE system, which was purely traditional essay writing: you know the kind of abstract "discuss this topic" and then a very obstruse reading passage with questions, and I was trying to persuade the board of the school to allow me to change the examinations and the curriculum simultaneously. So that was an early kind of language assessment literacy project with the board of management; and eventually they said: 'Okay you can run a two-year pilot', because they were frightened of change for financial and recruitment reasons. But I persuaded them. The project went really well. And we started using the international IGCSE. Other schools followed, and Cyprus still uses the IGCSE today. And back in the day in the late 80s/early 90s this was in the newspapers actually (although it's probably forgotten in Cyprus now, as these things always are). But that's how it happened, and that was how my interest in testing and assessment developed, back in those days as a young teacher in my 20s.

**LM: In your early teaching and research career, who had an influence on your work? Where did your ideas for furthering/building on that work come from?**

**GF:** Well, when I was doing all this in Cyprus, of course, all my background was philosophy, theology, and this sort of stuff. And I realized that as an English teacher, I wasn't really doing a great job. I was okay in the classroom as I'd done a PGCE at Cambridge, but I didn't really understand what I was doing regarding language. So, I did what most people do, and that was to enrol on an MA in Applied Linguistics. And that was with what was then called 'English Language

Research' at the University of Birmingham. Now this was the mid 80s when I was there – '85 to '87. Now if you think about who was at Birmingham during that period it was kind of like the golden era. John Sinclair was the Head of Department. Mike McCarthy was there. Malcolm Coulthard, David Brazil, Tim Johns, he was developing 'Micro Concord' - early research into corpus-based analysis.

I remember I was asked to trial an early copy of 'Micro Concord', and we used it in Cyprus to analyse the Bush-Dukakis debates, and then used the analysis to teach some of our students how to hold the floor in a debate. I mean it was really quite an exciting time when we were using this stuff in our teaching. Tony Dudley Evans was there, as was Chris Kennedy. The COBUILD project was in its early stages before the first dictionary was launched. A lesser-known member of staff was Charles Owen, and he was a great guy and a brilliant mind, but probably isn't remembered as much as the others because he didn't publish very widely. But he taught the language testing course, and really got me interested in assessment. I already was, but he introduced me to lots of the ideas of the time. And he also taught me my first statistics course. So, it was Charles Owen primarily, who was an early mentor.

Being there with all these other amazing people in Birmingham was just wonderful. And it was in that really rich environment that I started thinking, 'hey discourse analysis and language testing, these two fields, haven't been brought together' and they hadn't in the mid 80s at that stage. And I thought, this is somewhere where I can make a difference. My 1987 paper '*Tests of Oral Performance: The Need for Data-based Criteria*', was actually my language testing assignment which Charles Owen guided me through on my MA program. So, it kind of just goes to show that MA students can do this sort of stuff. I come across lots of MA students now, who have really publishable stuff. They just don't have the confidence to do it. That became the basis of my 1993 PhD thesis in data-based rating scale design. But the other person who really encouraged me was Caroline Clapham. Of course, who isn't with us anymore, sadly, but I met Caroline at a conference in Edinburgh in 1989. And I thought, well, if Caroline is anything to go by, this is a really welcoming community of scholars and researchers, and so she encouraged me a lot to take it further.

**LM: Your work has always had clear practical application (e.g., in the form of test design, instrument and rubric design), what challenges have you faced in contributing something 'new' in these areas as a scholar?**

**GF:** The practicality first, because I think that you're pointing out that what we do is practical and that is really important. And language testing is a practical activity. But it's also an application of theory. When Fred Davidson and I did the 2007 book, the advanced resource book, one of the things that came in the introduction was, and I hope you don't mind me quoting this: "sociolinguists do not create sociolinguistic things, discourse analysts do not create discourses, phonologists do not create spoken utterances, language testing in contrast is about doing, it's about creating tests". So, whatever we do always has practical application. But that doesn't mean to say that theory is not important. And this is why some of the stuff that I've done in language assessment literacy concentrates on theory. And that's where my Greek influences come in. The original Greek word

for craftsman was ‘Demioergoi’, from the word public (demios) and work (ergon), and the God of craftsmen was Hephaestus. And the word symbolised the joining together of theory and practice in the process of creation. It was only post Aristotle that they changed the word to ‘Cheirotechnon’ which is literally translated as ‘hand worker’. And I think this still lives with us, people who do practical things are not expected to be particularly theoretical and surprising enough, I’m not entirely sure that I’m right about this, but think about how Applied Linguistics at universities is often not considered to be a real academic subject, and we’re still living with that. And I think it’s because of this word “apply”, and people think, this is not really seriously academic, which, when I worked at the University of Surrey, was the reason for closing down the English Language Institute and the MA in Applied Linguistics because people working in Literature and Languages didn’t think what we were doing was real work - real academic work. I think we’re still living with this problem.

But we need to reverse that. And I’ve recently argued this case, in a paper I wrote that kind of sums up my view of how we teach language testing, which was in Dina Tsagari’s (2021) recent volume: *‘Language Assessment Literacy: From Theory to Practice’*. I wrote a paper on operationalising assessment literacy. But that’s the practicality, now to the new.

I suppose here that I must say that, there are mainly two areas in which I’ve done new stuff. And just reflecting, when you’re asking this question, I think I’ve had something that I’m now going to term the “combination principle”. So, going back to my time at Birmingham, discourse analysis was going on, corpus linguistics was going on. And there was some language testing that was totally separate. And what I did at Birmingham was to say: ‘all this stuff can be combined’. I can look at the discourse in a corpus of spoken interactions because I can use tools from discourse analysis to analyse transcripts. And I can use the tools of language testing to improve on all these defective rating scales and descriptors that were around at the time by looking at what actually happens in speech. And all of that stuff came together.

Now this, I don’t think I’ve ever told anyone this before, but the paper I published in *System* in 1998 which was Wilson’s model of communicative competence and the testing of reading, was a shortened version of my MA dissertation. That was an attempt to apply discourse analysis to reading texts to see whether I could validate Widdowson’s model of reading and communicative competence in a reading context. Now, the external examiner just happened to be Charles Alderson. And Charles said, this is exactly what I’m thinking about at the moment ‘How can we use discourse analysis in testing and assessment?’. ‘Why don’t you come to Lancaster and do a PhD with me?’ And I said, ‘Oh, okay’. And the rest is history, I suppose.

Of course, the data-based approach to rating scale design that I developed back then for my PhD has evolved, including the EBB method and we’ve recently done work on performance decision trees and stuff like that. But no one is going back to apriori rating scales. Okay, so that’s one area. Now, the second one, the combination principle is bringing areas together, that haven’t interacted before. And the next one of course is no surprise. It’s the philosophy. You know, it’s where I started out. And when I said earlier that I spent some time in Dartmouth in New Hampshire. I was actually there on a Descartes scholarship from King’s College London. And not a bit of history – one of those things that happen in life and change your direction of travel. I was

professionally involved in language testing working in UK universities. Fred Davidson, and I worked together on numerous committees, primarily the TOEFL committee of examiners. And in 2005 he got in touch with me and said, ‘we’ve done all this stuff together but we’ve never done a language testing project together, why don’t we do that?’. And I was living in Scotland at the time. I said come across and let’s talk about it and he came across to Scotland for a holiday, because he’d always wanted to go and see the Davidson castle up near Inverness. So, I was kind of like a tour guide and hotel combined. I’ll never forget this first day, he was in my flat in Scotland, and he was looking at one of my bookcases, and he said, ‘Hey, you’ve read Louis Menand’s (2001) *‘The Metaphysical Club’*. And this turned out to be one of his favourite books, and it was also one of mine. This was the start of our long and (I think) productive collaboration.

Menand’s book (2001) is the story of the start and evolution of pragmatism, that’s Pragmatism with a capital P, the philosophy in America, through William James, John Dewey and of course, Charles Sanders Peirce (1878, 1958). And this goes back to the practicality issue. In the introduction to the book Menand (2001) explains that they all believe that ideas are not out there waiting to be discovered but are tools like forks and knives and microchips - devices people use to cope with the world in which they find themselves. Hence, Effect-Driven testing, which is an application of Peirce’s (1878, 1958). Pragmatic Maxim to testing, and all the philosophical frameworks for assessment that we started to write about, back in 2007. I think that Effect-Driven testing, and the stuff that we’ve done on the philosophy behind what we do, has kind of become an idea for people now to look at the outcomes of research projects and testing consequences. And I like to think that’s an important contribution.

Sometime ago Benjamin Kremmel very perceptively said to me, ‘all this stuff you’re doing, the philosophical stuff, is it associated with any methodology?’ It took me aback this question, you know, bright new scholars you see coming up. But I think now, on reflection, the answer is no, it really can’t be. It’s a thought tool for evaluation, and for guiding practice. And that comes out in the epilogue of the second edition of the handbook: *‘The Routledge Handbook of Language Testing’* (Fulcher & Harding, 2022) which I’ve just done with Luke Harding at Lancaster.

But just to finish off the answer to this question, which is the second combination principle. Of all the stuff that I’ve written, the one thing that I think I’ll probably go back and read again in retirement, to see if I can find anything new in it, is my 2015 book, *‘Re-examining Language Testing’*; and one question I’ve got for myself when I reread it is, is that book about language testing?, or is it a book about philosophy?. And the answer is: I really don’t know. But I’ll try and answer the question for myself in retirement. So, finding things that are new, I think it’s being able to see connections, to combine, and to build on the work of other people. And then to offer things that colleagues in the field will find useful in their own work.

**LM: When you’re combining different theories different methods, do you think there’s an extent we can go to, or do you think we could take it too far in trying to make connections.**

**GF:** Yes, I think sometimes there are. I think I’ve come across an example of that recently actually. At the moment, many people in universities now have a decolonization agenda. I’m sure it’s the case in the US but this is increasingly becoming the case in the UK. This is beginning to affect

assessment practice and I attempted to respond to this recently. In order to respond to this, we have to take a philosophical starting point because it's not a technical issue over assessment. And I thought, probably the best place to start in dealing with this was the work of Hilary Putnam, who of course is a modern Pragmatist. And I attempted to apply some of Putnam's arguments about how people can be marginalized through decolonization practices. And I wrote this up as a draft and showed it to someone and they came back with a number of questions on detail. I suddenly realized that this could all be seriously misinterpreted. And I dropped it very quickly. But, yeah, but that's what colleagues and reviewers are for. People will soon tell you if they think you're making a mistake, which is why reviewers and other readers are really important. We never stop learning.

**LM: In the spirit of contributing something 'new', what research skills have you developed over the course of your career and how important has that been to your development?**

Well, I think someone coming into language testing from my background, which was kind of really very theoretical, apart from the practical issues of analysing texts and manuscripts, and text being redacted, you know, I had all those kinds of skills, but the one thing I didn't have anything in was statistics. Originally with statistics, I knew it was a gap and Charles Owen really got me interested in it.

And coming from a background, particularly the theological side, I found when I first started studying statistics, I was gung-ho and I thought, this is it. This is the answer. We get the numbers and they tell us the truth. So, I really got into all the statistics stuff. But it's been a constant challenge. But I think it's also been one of the most rewarding things to have studied. Because in school my maths was absolutely atrocious. And it's now a lot better because I can see practical applications. I said it's been a career long challenge but that's partly because I've learned to link it back into the philosophy.

And if you look at philosophers who discussed early social research using numbers like Ian Hacking's (1990, 2014). *The Taming of Chance*, it raises questions like, (which I don't think language testers really think about), 'Why do we use the population standard deviation estimates, to calculate the standard error of a score of an individual?' Now that's an amazing philosophical question. And it really helps you to understand what we're doing in calculating the standard error. So, it's one of the things I do actually say, 'Well, how would you do it without using the population standard deviation?'. And the answer to that question is, you have to conceive of a world (and this is going into science fiction), in which individuals could take the same test in parallel universes (this is just mind-blowing stuff!), which is why I like statistics. I will never be a creative original statistician, I mean I use statistics to do things and I think it's great, but it's been a lifelong challenge.

But the other side of the coin is one that I learned from Fred again. And he talks a lot about statistical determinism: that you cannot let the numbers alone determine what you do. It has to be an interplay between meaning, measurement, and the values that you bring to the testing. Philosophy and numbers are just tools like ideas in doing what we do. And arriving at what Messick would have called an 'integrated evaluative judgment.' Now, you know we trot off this

quotation from Messick, again and again and again but I think we stopped really trying to understand what Messick was doing, because Messick got this, that everything is evidence, and it all contributes to us as human beings making decisions. And that's why last year when I was invited to give the Messick lecture, I argued there that Messick really got this right - he really understood it. And some of the other players now, working on validity theory from the wider world of education measurement - I don't think they get this. I think part of the reason for that is that they come from a background in psychometrics. I think that's a real problem. So, given a choice. I would prefer to come from philosophy and learn statistics, rather than come from a psychometrics background, and then try to have to acquire the rest, but it has been a lifelong challenge.

**LM: You've supervised many doctoral candidates in language assessment projects, what are some of the important lessons you feel they must learn about assessment as they travel through their respective journeys? What have you learned from them?**

I find this a very difficult question to answer, actually for a variety of reasons, but partly because of my own approach to how I supervise. I have colleagues who are very directive. It's kind of on a model of the natural sciences perhaps: Someone has a really big grant, and they have individual questions they want to address. So, they recruit doctoral students and then direct them to answer that question. I know people who do that in language testing, but I've never done that.

I think that it's much better if individual projects stem from curiosity. So, how do you teach people to be curious? I'm not sure that I know, but I don't want to tell people what to do. I want them to be curious, and then passionate about it. And I think that's what I try to do in my MA teaching. So that some of them will go on and be curious enough to want to learn more. So, if I could extend my answer to the students. That's why, in that paper in Dina's book, (I'm coming back to that now) - 'Operationalising assessment literacy'. That's why I came up with this model of acquiring assessment literacy as an apprenticeship. So it's moving from apprenticeship to craftsman, but in the sense of theory and practice combined.

So, I guess what I'm saying is that this has a number of implications. First of all, that the students are at the centre of the process and what I have to do as a supervisor is help them to develop the skills, the confidence, and the curiosity. Secondly, that they learn by doing. And by doing, they uncover problems and issues that they become curious about.

The third one is of course, that you kind of place yourself as the master craftsman who's guiding the apprentices. But knowing that you're not perfect, of course. And I've got a few quotations to bring in now because I was reading about this when I was doing that paper for Dina's book. I was rereading some of Charles Sanders Peirce's (1878, 1958), more obscure writings. And he actually wrote a paper about teaching and learning physics, which I had never come across before. And he said about his students (when he was a university lecturer for a short period of time in his life in the late 1800s): "they should be made to feel that they are doing real and important work, which was to appear in the digests of science, and for the accuracy of which they are responsible". I think that's absolutely wonderful. If you can get students to feel like that, I think you've achieved an awful lot.

So, I guess the way I'm answering your question is kind of like an interaction between how I feel as a supervisor and what I think students need to acquire. They don't always match up. But we do our best. But I have to say, I enjoyed the teaching probably more than anything else I've done in my career. I love teaching. I find all that very rewarding and seeing our students do new stuff that I haven't thought about keeps me in my place.

### **Experiences and lessons learned when working with industry testing companies**

**CC: What are some of the key things you've learned about large-scale assessment from working with industry testing companies?**

**GF:** Most of the work I've done has been with ETS. I have done work for other testing companies and examination boards, a handful, but the only sustained consultancy work has been with ETS. So, I'm really going to direct my comments at ETS. I have to say that all of my engagement with ETS has been instructive and positive. And the first thing that I learned about ETS as a company, was that there is *a lot* of creative thinking that goes on in these companies. As academics we have these external roles, so we have a kind of symbiotic relationship with them, things to learn both ways.

We were kind of there to critique and Fred used to say this when he was on the TOEFL Committee of Examiners. He used to say: 'We are the custodians of TOEFL, the external members of the committee. And we are there to ensure quality, and to give ETS advice on how to go forward'. Now, although the relationship between externals and the company has changed over the years, I don't believe that the companies who use externals and consultants have essentially changed that relationship. And I know that ETS is exceptionally open to criticism, and to change as a result of what externals do.

Now, going back. My first involvement with ETS was actually on the TOEIC advisory panel. And I was drafted on to that by Gary Buck, when he worked at ETS, before he went off and created his own testing company. Gary, Steve Ross and Miyuki Sasaki were on that panel and they are great people to work with. I still stay in touch with both of them, and I've been working with Steve Ross recently. I remember in the great storm of 1996, when the whole of New Jersey was covered in snow, and we couldn't get out. We were trapped in a hotel in Princeton. The three of us with Gary developed an internet-based business writing test, during the two weeks we were trapped in the hotel. We ran the test from the University of Surrey. Steve collected the data in Japan. And then we tried to sell the idea to ETS for the evolution of TOEIC because there were no writing components in those days. And ETS entertains this kind of creative thinking.

The second thing is what Fred would call 'stasis'. In all companies there is a reluctance to change. Because, and you can understand this, they also have to run themselves as businesses. They are commercial enterprises, even if like Cambridge and ETS they are actually officially charities. So, they don't generate profits, they generate surpluses, but they do have a purpose to generate surplus, so there is this stasis. One example is the reluctance to change scales on tests. So, even if the construct changes, they're reluctant to change the scale, or to change what's in the test if they think the users won't accept it. But ultimately, what happened at ETS when they changed



from paper and pencil to the CBT, they changed the scale. And then when they introduced the iBT, they changed the scale again. If we look back at Charles Alderson's paper in 1998, I think it was, in Brian North's book on bands and scores. Charles in that paper tells us that the IELTS scale hasn't changed because they're unwilling to change the meaning of 6.5 or 7, irrespective of construct change. So here we have examples of stasis, built into the companies because of business. And also an example of how they take on board research informed decision making. And for me the greatest experience as a consultant, after Fred nominated me for the TOEFL Committee of Examiners, was the huge privilege of being Chair of the research subcommittee during the development of the TOEFL iBT. And they trusted me with the commissioning of research that would affect what they did. And we commissioned a lot of studies in those days, one of the greatest was the collection of remarkable washback studies, from Dianne Wall, in the TOEFL Research Report series.

So, what you can do as an external does actually genuinely change the approach to testing and teaching. I learned a lot about the process of test development. Sometimes I think I learned more from them than they did from me. I learned a lot about the process of test development and particularly item writing, and the care with which it is done. And lastly, I suppose working within a commercial environment, you learn about the resources and the limits on what is practical and what isn't. You see in the books, you know, one of the features of testing is practicality but I don't think you really understand it until you see how these people work.

**CC: Reflecting back on your experiences, what have been some of the biggest challenges?**

**GF:** Well, you're asking this in the context of working as a consultant or working with large testing companies, and I think I have to say I don't think I faced any major challenges at all. That was because it was all very collaborative and when I think about some of the people I've been able to work with, Bob Mislevy, Don Powers, Brent Bridgeman, Susan Nissan, Mike Kane, Xiaoming Xi, Spiros Papageorgiou, Yeonsuk Cho, to name but a few, all these people are wonderful to work with, they're great colleagues. Not forgetting all the other consultants who have become friends as well. And over the years we've met in the most amazing places as well. I've travelled to most of North America for ETS meetings.

The real challenges I'm going to talk about are in other areas. It's not with working with colleagues in language testing and assessment, or even Applied Linguistics. All the challenges that I faced in my career is because of the reluctance of UK universities, and I'm sure it's the case in universities in other countries, the reluctance of universities to recognise the importance of Applied Linguistics and language testing research. We get no credit for the work we do with companies. We've had no resources to build upon what we do, we're actually for the most part invisible. And in my own university, just as an example, I have been in my current institution for 14 years, and in those 14 years, not once have I been asked for my opinion on assessment or language assessment, in particular, not once; and yet there is constant, top down, administrative change. When I've tried to advise, it is ignored. It's as if I didn't exist.

The only person I can see who has really overcome this is Cyril Weir, I have to say, and to Cyril's credit. He's a great language tester, was a great colleague and friend. But he's the only

person who has persuaded multiple universities to establish serious test centres/departments that focus on language testing and to invest in them, to build up the research centres to get critical mass. And I think Cyril has to be given all credit for doing something that I could never do. He had this ability to deal with policymakers and engage with and overcome this administrative wall that we run up against.

The second challenge is a personal failing. And that is despite my early start with the analysis of Greek manuscripts, I've always had a problem with attention to detail. And so, when I started co-editing *Language Testing*, I found that really, really hard. That was a huge challenge. But, Cathy Elder, and then April Ginther, helped me hugely and I got better at it.

It also helped with my own academic writing. And often students find this difficult to believe that people like us, you know who they think of as these great teachers, we struggle with academic writing, and we do all throughout our careers. So, if I can just get this in, John Read wrote a review of my 2010 book '*Practical Language Testing*', and he could have rubbished it, I wouldn't have minded. But he said one thing, which I think was the most moving compliment I've ever received. He said this book is "literate".

I thought, 'wow! all this effort, trying to focus on expression and get the detail right, it's kind of paid off and perhaps I am getting somewhere with my own academic writing'. So, you asked about challenges, nothing to do with testing companies. It's the environment in which I've worked and overcoming these problems I had with attention to detail and moving forward with my own writing.

**CC: If you could give industry testing companies one piece of advice for the future, what would it be?**

**GF:** Here it is testing companies: Do not allow the constant drive to increase test volume, and hence income, to trump validity. Scores only have meaning when they're used for defined purposes that informed the design of the specifications, or if the test has been retrofitted for new purposes: subverting validity does no service to test takers or score users.

**CC: How can we bridge the gap between researchers (that is the research output) and teachers/practitioners (real-life educational practice)?**

**GF:** Well, I find this really hard. I think it has to do with whether you're, perhaps what I call a one or two-paradigm person. I'm a two-paradigm person, but I know that others in the field aren't. Atta Gebril (2021) has just done a new book '*Learning-oriented Language Assessment: Putting Theory into Practice*' which is a great book. And he invited me to do a paper for that. And I argued in there that there are two paradigms. The large-scale high stakes test paradigm and the Learning - oriented formative or assessment for learning paradigm. I genuinely believe that these two are completely different and should remain separate. Now, if that's the case and that argument holds up, then, the one thing that we can do is spend more time looking at how classroom assessment is used for the purposes of learning. I think that will go a long way to bridging that gap. Teachers need to be able to select tests, I think, and understand what large scale tests do, but I don't think they need the kind of detail that we might work with, or people in the testing companies. So, I'm not sure that gap

needs bridging, as long as they understand what people in testing companies do. If there is a validity criterion for what teachers do in classrooms, it should be *change*, and large-scale testing isn't interested in whether we change anybody. I suppose the other thing is making research information in easily digestible forms freely available for teachers, when they're doing professional development. And I think that's always been the idea behind my website. I don't know how many of those materials are used. I wish I did, to be honest, because I'd be able to improve them. But I think there are lots of people out there working in this area, and how we bridge the gaps is probably one of the questions that research faces in the future.

### **Future opportunities and challenges for the language assessment community**

**LM: Like all fields, language assessment will continue to grow and face challenges, what opportunities would you like teachers/scholars to take advantage of in the future?**

**GF:** They need to grasp opportunities for personal individual development.

And it's no different between teachers and others for that matter, as language testers, and their learners. What's important is our ability to change - our personal growth. And, you know, given what Fred and I have done is grounded in Pragmatist theory - that's exactly what John Dewey argued: the primary goal is personal growth and change. I'm not going to go on too long about this because I think lots of people are doing this but again, this is personal reflection. I started out when you asked me the first question, 'how did I get into language testing'? by saying I became curious about why I couldn't predict scores. And I said, 'and then I did some research on this, and I published it in this book'. And I think that teachers, learning to develop what they do and how they do it in a process of locally developed research, which they then share with others locally, even within their own school or publish if they so wish, is important. Getting teachers into the habit of thinking that what they do personally should be research-led would be a really important step. Because it would also allow teachers to take control of what they do and give them some agency as well.

**LM: What challenges do you think the field might face in the future?**

**GF:** This is so hard to predict. All I can tell you is what worries me at the moment. And that is around the world, we're seeing lots of what we do, in education, and then testing and assessment, starting to be driven by ideologies. And, of course, the one at the moment is the decolonization agenda, which I mentioned earlier. In my own university, a position paper has recently been published that actually suggests that we should change assessments to allow people who are identified as marginalized to pass. I wrote a piece for the Times Higher Education supplement (Fulcher, 2021), which was published in October last year, saying why my university had got it wrong. I think that when educational policy and assessment and validity theory is hijacked by ideologies, we're moving into a very dark place. One of the little stories I told in the newspaper article was that when I engaged in a dissemination activity within my university, and I explained what bias was and how we might tackle that in the case of our university, the question I immediately got was not about the paper or about bias or about how we make sure no one's being

unfairly treated, the question was: Was the person who conducted the research white? I find that exceptionally scary because it's identity politics at its worst. And it is a standpoint epistemology. It means that we can only know things from our membership of certain communities. In other words, it trivializes knowledge. And I'm just going to have one more quote from Louis Menand (2001) where he said pragmatists hold, "the belief that ideas should never become ideologies, either justifying the status quo, or dictating some transcendent imperative for renouncing it was the essence of what they taught". So, basically, I hold that ideas are the tools for "skepticism for how we cope with life in a heterogeneous industrialized mass market society". And as academics and language testers, our job is to challenge this ideological nonsense. Because I think this is going to be the greatest challenge of the next decade. And it will affect what we do if we don't stand up against it.

**LM: In which areas do we need to do more research?**

**GF:** The first one is the interaction between policy making and assessment research. I've never done anything like this myself. And I've always railed against people who make policies and never listened to me! But then I know it's easy isn't it? But, you know, I did that podcast with Bart Deygers from the University of Ghent. His work on looking at how policymakers think, and why they think that what we do is irrelevant to their work, was kind of an eye opener. I thought this really does have to be taken a lot further, because without that research, we're not going to be able to make an impact. So, I think that that whole area really needs opening up.

Second area. And this is because of growing commercialization and the need to make more money on the part of the testing companies, is the field of automated assessment.

Now, I'm not just talking here about doing it. I'm talking about looking at what it can do and what it can't do. And this goes back to some of my own research. Lots of the stuff that's being done at the moment goes under the heading of "Cognitive Linguistics" or "cognitive approaches" to language testing. Translated, what that means is: this research uses low inference category analysis to make processing claims about human "computers". Now what I mean by that is, let me take a simple example, in the area of fluency research, automated assessment counts the number of pauses and how long they are (among other things). And then an algorithm is created that makes a judgment about how fluent the person is. That is a low inference category. The fact of the matter is that we *use* pauses. And we have different lengths of pauses for communicative purposes. Sometimes it's for amusement, and great comedians know that, knowing when not to speak, is just as funny as what you say. People use pauses to indicate that it's time to change speaker. There are all kinds of reasons, and these are high inference categories. The fact of the matter is that human communication is a high inference activity. And we need to say: 'okay, we can do this low inference stuff'. And for some learners that might be appropriate. You know, if I have beginning learners, the low inference categories may indicate problems with processing grammar and selecting lexis. It's still a surrogate, but it may help, and I think we need to not just come up with better algorithms and so on that tend to correlate with human judgments. We need to look at what it is we're actually doing with language because correlation, as we all know is not causation; nor is it validity. These people who are working in these areas need to stop thinking this is cognitive,

they're treating humans as mere processors - asking how fast their chips are. I'm beginning to get a little bit irritated by it. So, this is where values come into validity. And we need to explore the meaning of language, and recently I've kind of talked about this in terms of language being part of our evolutionary heritage. And, you know, Data in Star Trek is a wonderful speaker. But the reason I love Data in Star Trek is because of all the cases where he doesn't understand what people say, because he can't process illusionary force. And he doesn't understand the emotive meanings behind words in context. And there are things computers are never going to be able to do, and I think we just need to recognize that. And last but not least, this exploring the meaning of language. And then, what kind of systems we develop in what context to assess it. I wish people would read Lado (1961). He's not on anyone's reading list anymore but should be – and I'd like to quote from page 276: all his research was done on “a basic assumption of belief in the unity of all mankind”. How great is that. I wish language testers would make their assumptions as clear as this.

**LM: Reflecting on your career, if you could start all over again, is there anything you would do differently?**

**GF:** I think that depends on whether I'd be starting again from today or in some kind of anti-Milan Kundera universe where I could rewind the tape and start again from the beginning. So, I'm going to assume it's the latter, you know if I could rewind and do it again. So, the answer is “no”, because I think I've been exceptionally lucky, in many ways, and I've had two careers. And that newness question that you asked me, you know, it has come out of being able to do different things in different fields. I also think being a voracious reader in other areas kind of helps as well. You know it's during lockdown and knowing the time is coming up to early retirement, I've been reading Marcus Aurelius' *Meditations* again (see Aurelis et al., 2020). I also found a wonderful book: first century AD, called 'The Shortness of Life'. It's actually very good. I think people should read it before retirement, really good tips on how to prepare. So, I think I've been very lucky and so I don't think I'd do anything else and I'm just happy and grateful that I've been able to do bits which colleagues feel have been useful.

**CC: What's next in retirement?**

**GF:** Music Theory. I have never, ever had the opportunity to study music properly and yet I've always played. I want to be a rock guitarist by the age of 70. (We can but dream). I was in a band when I was at school, and the drummer, actually became a professional musician, and we're still in touch. We're getting together and starting to jam again, and I have some nice guitars. I love reading music and I love music theory.

I live right on the edge of a beautiful 900-acre deer park, and I'm going to spend more time in there. And spending time on my house, but it's almost finished now. When I got the Chair at Leicester, I decided I would finally settle down and buy myself somewhere to live. I took a huge risk, I bought a disused flour mill that was built about 1790 and bringing it up to the standards of a modern home for a listed building that you can't do much to because of its historic importance, has been absolutely wonderful over the last 10 years. Yeah, so this, and being able to stay at home a bit more, go for walks and do that sort of stuff. I'm also an avid board game player, and

particularly strategy games - I also make my own games. I design and paint them. So, these should keep me busy as well as reading. But I'll be staying in touch with the field too. Leicester University is giving me an Emeritus title, and I've just been elected a Fellow of the Academy of Social Sciences. I think a visiting professorship is also on the cards.

**LM and CC:** Glenn, thank you so much for taking the time to do this interview with us.

### **Works referred to in the interview**

- Aurelius, M., Seneca, L. A., & Epictetus (2020). *Three stoic classics: Meditations by Marcus Aurelius; The Shortness of Life by Seneca; Selected Discourses of Epictetus*. Benediction Classics.
- Gebril, A. (2021). Learning-oriented language assessment: Putting theory into practice. Routledge.
- Hacking, I. (1990). *The taming of chance*. Cambridge University Press.
- Hacking, I. (2014). *The taming of chance*. Cambridge University Press.
- Lado, R. (1961). *Language testing*. Longman.
- Menand, L. (2001). *The Metaphysical Club: A Story of Ideas in America*. Farrer, Straus, Giroux.
- Peirce, C. S. (1878/1958). Research and teaching in physics. In P. Wiener (Ed.), *Values in a university of chance: Selected writings of Charles S. Peirce*. Doubleday & Co.
- Tsagari, D. (2021). *Language assessment literacy: From theory to practice*. Cambridge Scholars Publishing.

### **Professor Fulcher's work referred to in the interview**

- Fulcher, G. (1987). *Tests of oral performance: The need for data-based criteria*
- Fulcher, G. (1991). The role of assessment by teachers in schools. In T. Caudery (Ed.), *New Thinking in TEFL. The Dolphin Series, 21*, (pp. 138-158). Aarhus University Press.
- Fulcher, G., & Davidson, F. (2007). *Language testing and assessment: An advanced resource book*. Routledge.
- Fulcher, G. (2010). *Practical language testing*. Hodder Education/Routledge.
- Fulcher, G. (2015). *Re-examining language testing: A philosophical and social inquiry*. Routledge.
- Fulcher, G. (2021). Calls to decolonise assessment do students a disservice. *Times Higher Education*. Retrieved from: <https://www.timeshighereducation.com/opinion/calls-decolonise-assessment-do-students-adiservice>
- Fulcher, G., & Harding L. (2022). *The Routledge handbook of language testing* (2nd Edition). Routledge.

### **Acknowledgments**

Not applicable.

### **Funding**

Not applicable.

### **Ethics Declarations**

### **Competing Interests**

No, there are no conflicting interests.

### **Rights and Permissions**

### **Open Access**

This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. You may view a copy of Creative Commons Attribution 4.0 International License here: <http://creativecommons.org/licenses/by/4.0/>.