

# *Assessment for learning in Norway. Opportunities for connoisseurship and language games in the globally important work of Gordon Stobart?*

Stephen Dobson\*

How can we grasp the influence of Gordon Stobart, a single and yet pivotal figure in the assessment for learning movement that originated in the UK and became global?

The argument I will make is that he is an *assessment connoisseur* in his work as a teacher, academic and policy adviser the single word I would use would be “self-effacing”. So too in his work in Norway, a society in which a famous proverb called Janteloven underpins the genuine national belief in a culture of equality of all: “du må ikke tro du er noe” (you must never believe you are anything special) (Palamara, 2016, p. 1). Accordingly, it is never about Gordon who is charismatic and a great public lecturer; it is always about the case at hand. In this paper, the case is about the “doing of assessment” by different stakeholders in a country such as Norway.

Some might argue that his influence as a co-founder of the Assessment for Learning movement in UK has been over-shadowed by another front figure; namely Dylan Wiliam (2009). This would be to neglect how Stobart’s ability to be a great communicator and teacher has made him an equally popular exponent of assessment for learning for a varied group of listeners, including but not limited to school teachers, academics, policy makers and the more general public interested in gaining access to new ways of understanding the connection between assessment and learning across many different fields and professions.

We must pay attention to the way in which he communicates, either verbally or in print, where the medium of telling stories is as important as the

*Rec. 7/4/2026, app. 22/5/2026*

\* Norwegian University of Science and Technology, Norway; [stephen.r.dobson@ntnu.no](mailto:stephen.r.dobson@ntnu.no).

*Cadmo* (ISSNe 1972-5019), 2026, 1, DOI: 10.3280/cad2026oa23294

content. He models how assessment can be formative and how it can aspire to meeting the needs of all who happen to be listening. He often uses a touch of humour as he communicates small kernels of advice distilled from his expert noticing and sifting through data, experience, policies and what at times appears to him to be the over-academic theoretical jargon belonging to the world of assessment experts. He echoes the characteristics and goal of the classic storyteller: “In every case the storyteller is a man who has counsel [...] but if today ‘having counsel’ is beginning to have an old-fashioned ring, this is because the communicability of experience is decreasing” (Benjamin, 1973, p. 86).

The author of this comment, Benjamin, lived in Europe and wrote in the 1930s, noting the rise of the information society and its power to influence our thoughts, debates and practice. Stobart is a child of this information society, always up to date on the latest policies, and yet still interested in what he might distil for the benefit of all through his clearly evident storytelling skills. His project is the communicability of assessment knowledge and skills in all its forms and to accommodate the needs of different audiences.

So how does he do it, this teller of assessment stories? Stobart is for want of a better term an “assessment connoisseur”, schooled in the many diverse language games of assessment – for there are many and they change from context to context and in popularity over time. We will define these in more detail in due course, but at this point the term language games is taken to denote the different kinds of assessment language that adopt varying mixes of popular, folk and technical terms e.g. grading, tests, assessment of learning, for learning and as learning. Let me clarify at the outset what I mean by a connoisseur and its connection with the language games of assessment.

In the tradition of Sadler, inspired by Polanyi (2009) not all assessment relevant moments can be easily grasped and put into the language of succinctly defined and linguistically clear and delimited rubrics. They are at times tacit and captured by the common phrase, “I do not know how to define quality, but I know it when I see it” (Sadler, 2012, p. 9). It is from this kind of understanding that the assessor as connoisseur draws – looking for diverse ways to express advice on assessment practices that are at times hard to communicate in all their richness.

However, the idea and concept of the assessor as connoisseur is nothing new. It has arguably always been central in the assessment of creative and performing subjects as varied as music, art and literature and learning that is both expressive and practical. From such a perspective, “Connoisseurship

is the art of appreciation. It can be displayed in any realm in which the character, import, or value of objects, situations, and performances is distributed and variable, including educational practice” (Eisner, 1998, p. 63).

Taking these points on board, Stobart is the living embodiment of the assessment connoisseur. He displays in numerous and varied ways how to communicate the “doing of assessment”. With Norwegian teachers, academics and policy makers he always prepares every presentation in meticulous fashion. In advance, he asks members of the group he is to meet, “what are the hot assessment topics in the Norwegian classroom, among academics in Norwegian initial teacher education or Norwegian civil servants seeking to implement a national policy on assessment for learning?”.

Accordingly, Norwegians know him for selecting different forms of communication, adjusting to the details of the setting, whether a lecture, a conversation or some other professional context. The examples chosen for teachers versed in classroom challenges, might include how to deal with poorly motivated school students showing limited interest in an assessment (for learning) task. He would spice up his thoughts by mentioning a recent Norwegian winter sporting event and how a local Olympic Gold medal cross-country skier had given him a special “honourary” practical skiing lesson the day before the lecture. To policy makers in Norway, he might draw attention to a particularly pertinent scene in a popular TV shown in the country, such as *Desperate Housewives* (one of the many examples he uses in his book *Testing Times. The Uses and Abuses of Assessment*, 2008, p. 19). To researchers in Norway schooled in assessment, validity and reliability might be mentioned. But for sure he would also remember to reflect upon the Norwegian goal of equality (*Janteloven* cited above) rather than hierarchy, making this a key point in reviewing assessment principles suited to the Norwegian context.

Accordingly, to this assessment connoisseur in each of these applied assessment contexts he selects the most appropriate and familiar communication – what inspired by Wittgenstein (1953/1967, § xl), my colleague and good friend Fuad Fudiyartanto and I would call a specific language game of assessment from the larger group of possible assessment languages games that share family likenesses:

The language that is used in assessment, feedback, questions, commands and so on only means something when the people (the teachers and the students) know the game in which they are taking part. Language games are a vehicle that help us make what we highlight or notice in assessment – quality, standards, excellence and the like – explicit; noticing is thus naming and bringing to light what is assessed.

Thus understood, we can refer to all the (language) activities in assessment as the language game of assessment or simply the assessment game (Dobson and Fudiyar-tanto, 2023, pp. 193-194).

Thus understood, assessment for learning as defined by the Assessment Reform Group constitutes one such language game; discussions on setting a grade in a classroom are another and so on. When we use the term language game inspired by Wittgenstein, we are not inviting the reader to envisage a frivolous, leisure pursuit where the goal is to win and defeat other players. The language game we have in mind is the invitation for participants to adopt and agree to the rules that all have agreed to follow.

The Norwegian Directorate for Education and Training was responsible for implementing a new national formative assessment policy and invited Stobart in 2009 to Norway after a staff member (Vivi Bjelke) had heard him speak of his work in precisely the Assessment Reform Group the previous year at the annual Association for Educational Assessment – Europe conference.

Gordon was my PhD supervisor in the period 2003-2008 and I was only too happy to support his involvement. I also knew Gordon through among other things my work in Norway with Roar Engh and Elli Kari Høihilder. We co-led a national network on pupil assessment (NELVU)<sup>1</sup> for initial teacher educators drawn from Norwegian universities and university colleges. We also published the first book length text in Norwegian on assessment for learning (*Vurdering for Læring*, Engh, Dobson & Høihilder, 2007). For several years we brought the work of the Assessment Reform Group and other prominent assessment experts to the attention of network members (teacher educators in higher education) and the wider community of Norwegian school teachers.

To summarise, the communication of advice through stories, using humour as appropriate, picking the most suitable and appropriate language games of assessment, are a telling sign of Stobart's ability to be an assessment connoisseur. To use a Norwegian phrase, he has a "blikk for øyeblikk" (awareness of the moment) (Skogvoll and Dobson, 2011). It is about a sense of time, place and moment; drawing upon his long experience as a teacher offering formative words to guide students, as the policy adviser listening to the needs of education stakeholders and as an assessment researcher supporting and formulating robust data informed approaches and scholarly works<sup>2</sup>.

<sup>1</sup> NELVU – Nettverk for Elevurdering (Network of Pupil Assessment) financed by the Norwegian Directorate for Education and Training.

<sup>2</sup> The blikk for øyeblikk might remind English speakers of teachers who experience of classroom those heightened "moments of contingency" when they can influence the learning of the student (Black & Wiliam, 2009).

In what follows I will present to the reader how Stobart, based upon his life experience has continually refined his skills as an assessment connoisseur in the three selected applied contexts. Firstly, as the teacher who is an assessment connoisseur, where I will use the example of his work as my teacher when he supervised me for my second PhD, on this occasion at the Institute of Education, now part of University College London. The topic of my research was the Norwegian practice of the viva from final year junior high school exams, to senior high school finals and then in Bachelors, Master and PhDs degrees, with a special emphasis on the university level (Dobson, 2017). Secondly, the researcher as an assessment connoisseur who writes academic books on among other things assessment. His book *Testing Times. The Uses and Abuses of Assessment* (2008) underpins his academic credentials in supporting the assessment work of academics, several of whom have been researchers in assessment in Norway. The last part of this contribution moves to his work as an assessment connoisseur advising Norwegian education civil servants in their implementation of a national assessment for learning policy and framework in schools across the country. They have been supported by among other things his experience and reputation as one of the founding members of the Assessment Reform Group in the UK. He could also draw upon his many years of experience working for examination boards in the UK.

My assertion is simply that the common point connecting his work with teachers, academics and civil servants implementing policy in Norway is his ability to ‘touch’ them all with his experience as an assessment connoisseur.

## **1. The teacher as an assessment connoisseur**

Stobart is a collector of quotes worth sharing that make us think and this epitomises his method of teaching as a connoisseur. Nowhere is this more evident than in his book, *The Expert Learner. Challenging the Myth of the Expert Learner* (2014), where his arguments are supported by well-selected pearls of wisdom from experts in education and by extension those well-schooled in assessment. The title talks of the learner, but I read the book as intertwined with comments on the teacher – you cannot be one without the presence of the other. The experts in this book as learners or teachers are in no way limited to the classroom or researchers on teaching. Stobart includes sport coaches, well-known scientists such as Marie Curie and Albert Einstein and many others from different fields of human activity. But, I am jumping ahead of the narrative.

When I first heard Stobart in 2002 as the teacher the occasion was a week-long intensive master unit on assessment held at the Institute of Education. His teaching always commenced with some quotes. Below are some, admittedly not from this intensive unit, but from one of his later lectures for Norwegian teachers a decade later (5.10.2012, Lillehammer, Norway):

If teaching were as simple as telling we'd all be a lot smarter than we are (Mark Twain).

The student knows more than the teacher about what he has learned – even though he knows less about what was taught (Peter Elbow).

What's the best way to improve teaching? Focus on learning. And the best way to increase learning? Move the focus off the teacher and onto the student (David Kolb).

Learners must ultimately be responsible for their learning since no-one else can do it for them (Assessment Reform Group).

Quotes such as these used in his presentations to Norwegian teachers, and also in my own thinking and co-writing on the topic of feedback were a source of important inspiration and guidance (Hartberg, Dobson & Gran, 2012).

His style was to accompany each quote with a story or a joke to make the explanation of the point more easily memorable and understandable. Among the 4 colleagues<sup>3</sup> who collectively taught this unit at the Institute of Education (including Klenowski, 2011 and Wolf, 2011), it was he who piqued my interest. He possessed both broad (e.g. assessment policy implementation), and also more narrowly defined specialist skills in assessment (e.g. margins of error in grading different school curriculum subjects), from inside and outside academia.

Around this time (2002), I was highly motivated to gain entry to this field of academic study and practice. The main reason was as follows: my first PhD on the experience of refugees had been more rooted in sociology and cultural studies, rather than in education. In looking for a permanent appointment in the local Norwegian university's School of Education where I had casually lectured on (post-)modernity and education for over decade there were no openings for a sociologist. They were rightly only interested in somebody who possessed expertise in a recognised sub-discipline of education. I reasoned that taking a second PhD in the field of educational assessment might constitute just such a field. And most importantly, I viewed

<sup>3</sup> All acknowledged academic experts in assessment.

Stobart somewhat selfishly as the perfect expert mentor to support me in this endeavour. The chosen topic would be the viva in Norwegian education, with a specific focus on higher education, from Bachelor to Master and Doctoral studies. The viva in all academic subjects has tended to be more widely used in non-English speaking countries throughout the world (Dobson, 2017). I commenced in 2003 and graduated 5 years later.

So, what pointers does the *Expert Learner. Challenging the Myth of the Expert Learner* (2014) offer to the learners and teachers? He stands as one who knows the way and shows the way of becoming a teacher who is as an assessment connoisseur able to a) self-reflect upon their own teaching and what it must mean for the learner, and b) research and communicate the art of teaching and learning to a broad array of interested parties (not limited to the profession of school or university teacher/lecturer), including my own experience of him as expert doctoral supervisor.

In what follows, I will select a few points from this book in order to emphasise how they are apposite to understanding his role, in particular as a connoisseur of teaching (and learning) during my time as his doctoral student. The journey of the doctoral student is a special experience, it is like a wave moving in slow motion for full time students (3 years) and longer for part-time students, such as myself (a minimum 4 years enrolment is the norm according to most university policies). It arcs through repetitive cycles of writing, submission and feedback.

To be clear at the outset, his book on the expert learner is not specifically about doctoral supervision per se. It is about learning in a more general sense, and yet in my view it still holds key relevance to the doctoral experience for both the supervisor as teacher and the student who is the learner. Accordingly, a typical supervision with Gordon would be initiated by my sending in advance a piece of draft text I had written on say research methodology or the literature review of viva. We would meet to discuss for an hour in his small, cramped office at the Institute of Education; cluttered with piles of papers, books resting at different angles on his shelves and sometimes a bicycle – if he had cycled to work on that particular day.

He rarely offered written comments. He would however, pre-read and think deeply about my submission. In the tutorial the conversation would gradually circle from the bigger picture of where I was seeking to move with the draft text to one or two verbalised questions concerning a key point or two and he wished me to unpack them. It was a deliberate nudge so he could gain an understanding to offer formative feedback.

In the book on the expert learner there is a chapter entitled: “Is everybody clear”, and it opens with the telling of a story about the legendary US



college basketball coach John Wooden. This coach was known for starting a training session by bringing the team together and discussing a particular 'play' to be learnt. This was followed by meticulously breaking it down into its individual parts and working through each of them. This is much like a PhD with its many parts and Gordon was similarly concerned with how a single argument or even a draft contribution might fit into the whole. We both took for granted that a PhD required deliberate and targeted practice with 10000 hours plus of effort and the tutorial was the flag in the ground offering direction.

The communication of respect is a key element mirrored in this book, and it is of particular relevance to doctoral supervision and how he regularly called upon it in his own skills as an assessment connoisseur. In each chapter in the book and equally in every doctoral supervision meeting a deep sense of respect was evident for the specialised relationship between the supervisor (e.g. a teacher or lecturer), the student (a pupil but also a doctoral student) and the more general member of the public reading the book and motivated to learn how to become an expert learner and teacher.

So, to be more specific how did he actualise this understanding of respect in his practice as an assessment connoisseur? In the Expert Learner book there are numerous examples, where the common denominator is the stretching of the student by setting them goals that move them in a scaffolded and supported manner outside of their comfort zone and secondly, offering feedback that is formative on the particular matter at hand, but also additionally offering the learner advice on how to manage the learning trajectory in general.

For me even with one doctorate under the belt, the latter was particularly important in offering support to navigate the liminal space that is a doctoral project. Gordon always sought to find the learner where they were and to emphasise that learning is about embracing and learning from mistakes. A PhD project is about taking what might appear to be so many false steps and seeking advice on different directions so a chapter or an argument might be rewritten in a minor or major way.

## **2. The researcher as assessment connoisseur working with Norwegian academics**

How does Stobart work with Norwegian academics, especially those who are lecturers researching and training teachers in programs offered by Norwegian universities and university colleges? The short answer is that he



draws upon and applies research that you will find in his book, *Testing Times. The Uses and Abuses of Assessment* (2008). The observant reader “notices” in the spirit of Wittgenstein (1953/1967, § xl, § 67) how in working as an assessment connoisseur he selects appropriate examples from his wide understanding of the language games that consider research on assessment capable of supporting and influencing lecturers who are training the next generation of teachers. The specific language game of assessment in which he has been interested is assessment for learning i.e. not the language game of summative assessment separated and after the teaching. Rather, the language game of assessment which is continually interwoven and interspersed by the language game of learning, such that each informs the other in spiral movements in a forward direction.

It might be asked: is he himself a researcher on the one hand, and also on the other hand one who cites and uses the assessment research of others, such as that of his fellow members in the Assessment Reform Group? The answer in Norwegian is “både og”, meaning he is both. Regarding the former, he has been a member of a longitudinal study of variations in teacher effectiveness (the VITAE project, Day et al., 2006), along with participating in evaluations of government strategies in England and Scotland, and the impact of changes in assessment policy (ARIA project, Gardner et al., 2008). Additionally, his research is often theoretical, deepening the reader and practitioner’s understanding of assessment concepts and how they can be applied in different settings. Simply put, as a philosopher of assessment who sifts through a collection or repertoire of concepts deciding how to selectively use and creatively combine them in the manner of a connoisseur.

On the other hand, as the chief editor for a number of years in the 2000s of the disciplinary journal *Assessment in Education: Principles, Policy and Practice*<sup>4</sup> he references the work of contributors to the journal in his book *Testing Times* and in his lectures. The contributions are also referenced in his work with the earlier mentioned Norwegian National Network of Pupil Assessment (NELVU). He attended several of their meetings and a slide he often used to open up the conversation about the focus on research informed knowledge of assessment practices is reproduced below (Figure 1).

The message is clear, and it mirrors the work of the UK based Assessment Reform Group. It is not enough to understand assessment in terms of whether a student has met a particular graded final assessment grade or standard. It is essential to explore the connection between learning and as-

<sup>4</sup> Launched in 1993 by a team of academics from the Graduate School of Education, University of Bristol and the Institute of Education, University of London.

assessment on the journey towards and after any final assessment. To repeat, he is the connoisseur interested in the research that highlights the connection between assessment feedback (i.e. feed-forward) given to support learning in a formative sense (and not merely as a once and for all measure of achieved summative attainment).



Figure 1 – *Learn* (1991, Syracuse Herald Journal)

The research by Black and Wiliams (1998) comes to mind where they identify different research informed understandings of formative assessment. Such use of their research by Stobart lends support to the view that he has used it in an applied sense, of course acknowledging the source.

In what follows, I wish to underpin the argument that he has also been very much the philosopher of assessment researching concepts and theories, and this is also evident in his book *Testing Times. The Uses and Abuses of Assessment* (2008). In what follows, I will accordingly draw from this book to indicate in what way he has been a researcher as assessment connoisseur. One, who in connoisseur style has selected highly debated global (and local/national – read Norwegian) debates on the formulation, implementation and evaluation of assessment policy and practice.

Each chapter in the book engages with a contentious assessment topic of policy and practice, and he proceeds in each case to consider theoretical

concepts that are pertinent to increasing our understanding of assessment practices<sup>5</sup>. The contributions in the book are entitled:

- Assessing assessment.
- Intelligence testing: how to create a monster.
- The resistance movement: multiple and emotional intelligences.
- The lure of learning styles.
- The Diploma Disease: still contagious after all these years?
- The long shadow of accountability.
- Reasons to be cheerful: assessment for learning.
- Reclaiming assessment: becoming responsible for who we are.

In eight chapter he offers a compelling A to Z of what is important in assessment. The only topic missing is the role of AI in supporting or disrupting existing assessment practices since its arrival. But the jury is arguably still out regarding the full impact upon assessment informed by AI, and how this is in turn contributes to a paradigm jump in our research, practice and understanding of learning and assessment. As Schivelbusch (2014) noted in his much-loved book on the history of the arrival of the railway journey, it can take as much as 50 years for the full implications of a technological innovation to play itself out in the consciousness and culture of its users.

As an aside on this topic, he gave the inaugural annual Stobart lecture entitled, “Adapting Assessment for the age of Generative AI” (19.11.25). In this he highlighted the need for students to be able to make assessment judgments on the validity and reliability of the AI produced knowledge. It cannot be just, “taken for good fish”, so to speak.

Importantly for the argument at hand, each contribution in the book evidences his skills as an assessment researcher who in expert connoisseur style investigates and unpacks theoretical arguments and concepts. The opening of chapter one entitled *Assessing assessment* is typical in this respect. He sets the scene by citing two very different researchers: Nietzsche the philosopher who writes, “unspeakably more depends on what things are called than on what they are”.

In citing Nietzsche, he is not going down the road of postmodernism or the German philosopher’s well-known belief in the theoretical concept *ubermensch*, where there are presumed to be a limited number of superior humans believing it is their right to exist above the “motley rabble”. Instead, he draws attention to the importance of naming activities that influence how an activity is practiced, understood and valued. When a later contribution

<sup>5</sup> In what follows I draw upon my published review of the book (Dobson, 2009).

is evocatively called, “the diploma disease: still contagious after all these years?”, the reader might recall and ask with the Nietzsche quote in mind: maybe in contemporary society we have too many *qualifications*? e.g. previously to work in business few qualifications were required at entry point and now in an ever-widening move to value credentials a vocational or university qualification might be a requirement. It invites us to think that there might also be alternative qualifications, variously called MOOCs or micro-credentials depending on the naming fashion at particular points in time.

This echoes what Wittgenstein would have called different kinds of languages games associated with different kinds of activities (in our case assessment practices that value activities differently). Accordingly, the book *Testing Times. The Uses and Abuses of Assessment* (2008) can be read as an exegesis on the importance of understanding a limited number of assessment related activities and the different theoretical concepts that underpin and each name a particular language game. According, each chapter title denotes a particular language game of assessment with a selected number of theoretical concepts that inform and lead to greater theoretically inspired research, understanding and practice.

Alongside Nietzsche in the opening of contribution one, he cites Hanson (1994, p. 4) a well-respected contemporary anthropologist who writes: “the individual in contemporary society is not so much described by tests as constructed by them”. This draws attention to the importance of assessment culture and practices that tell us something in addition to how well we perform in the assessment. The identity of the assessment taker (and by implication that of the teacher who has skilfully taught them) can be bolstered positively or less positively by the result e.g. the smart or less smart person can carry this identity with them in future-directed activities in addition to knowledge of the actual grade and the knowledge itself.

The chapter “Intelligence testing: how to create a monster”, is an example that illustrates this. The contribution traces the history of IQ testing to the present day. Every few years the actual IQ test must be recalibrated to be more difficult. Without this adjustment, all would perform better in the test and then the originally IQ smarter individuals would appear to be less smart, as the “motley rabble”<sup>6</sup> catch up with them in their scores. Hanson’s point reverberates through this contribution. Namely, in the popular and professional culture that surrounds IQ, the test taker with a higher scorer is able to confirm and project their identity as “a highly intelligent” person

<sup>6</sup> My choice of phrase. Etymology of the phrase is from the 13-14th century and suggests a speckled, vulgar tumultuous group of people.

who is able and believed to be able to use this intelligence in a wide variety of situations.

Why does Stobart call IQ a Monster? In addition to making the theoretical point that we understand the validity and reliability as technical scores that vary over time (the validity of our inferences based on IQ test scores vary over time due to factors such as people are getting better at taking those tests), we must also understand the culture that praises the identity of the high scores and the obverse, looks down on the lower scorers in a potentially negative manner. Accordingly, the assessment monster is the cultural machinery of IQ that use scores to enhance a person's self-perception and sense of worth.

To summarise, Stobart is the researcher as assessment connoisseur who expertly selects key issues for academics and offers to them reflections on how to combine theoretical concepts. This in turn generates new opportunities for them to research and reflect upon their own work in the field of assessment. In his theoretical armoury, he draws upon a limited number of concepts, each with their own history of usage in assessment. Alongside validity and reliability, he often references fairness, accountability and transparency. There is another he often mentions, but is really more practical than theoretical, namely fitness for purpose. This really bridges the other principals and connects directly with the application of different forms of assessment and the cultural conditions of those designing, monitoring and taking these assessments. Simply put, cultural practices and principles of assessment that speak to the infinite number of types of language games of assessment. Wittgenstein (1953/1967 §18), not specifically thinking of assessment put it as follows: “(And how many houses or streets does it take before a town begins to be a town?) Our language can be seen as an ancient city: a maze of little streets and squares, of old and new houses, and of houses with additions from various periods; and this surrounded by a multitude of new boroughs with straight regular streets and uniform houses”.

### **3. The assessment connoisseur as policy adviser for the Norwegian Directorate of Education and Training**

In a strange quirk of fate, the work of the Assessment Reform Group and its founding members unintentionally (or intentionally) has become over time a global educational commodity exported and appropriated in the assessment policies of other countries. I even heard Stobart lecture in Adelaide (Australia) when based there between 2014-2018. He, like other Assessment Reform Group members embodies what Gulson et al. (2017) have called policy mobility.

In the actual translation into different national assessment policies, few have actually critiqued the main policy ideas of the Assessment Reform Group. Other than to say that assessment for learning has on occasion been implemented as mini-assessment of learning steps in classroom practice focusing on a) the “immediate” use of an assessment in say a weekly test, rather than b) its longer time cycle use to inform formative assessment for learning over a longer time cycle of say 6 weeks or a half-full year (Bennett, 2011; Dobson and Fudiyartanto, 2023; Hartberg, Dobson & Gran, 2012).

This apparent lack of critique in implementation might be a misnomer, since each implementation has to respond to and is transformed by the cultural characteristics and needs of the receiving policy framework<sup>7</sup>. As the assessment connoisseur as policy adviser Stobart as already indicated, has been especially attentive to the needs of the end user and lets this colour his comments. In recognition of letting the voices of users come to expression, I have sourced reflections from Norwegian civil servants who worked closely with him for a number of years (Ida Large, Trude Saltvedt, Heidi Paulsen and Reidunn Aarre Matthiessen)<sup>8</sup>:

In the national policy on assessment for learning from 2010-2018, Gordon was a valuable partner for reflection, collaboration and mentor for us at the Directorate of Education, and he participated with presentations at most of the meetings we had with the different groups in these years.

Gordon created tremendously positive engagement and he talked about assessment for learning in a way that made all the pieces fall into place for everyone concerned with student learning. He was an incredibly good lecturer, listening, interested and not the least very full of humour. He presented a complex topic in an understandable way regardless of the target group and their position in the education system, always up to date with the latest research, and he managed to combine theory with practical examples. He came up with concrete tips and advice for teachers and school leaders that they could take with them in their work with pupils in schools<sup>9</sup>.

From this the reader takes the important learning about Stobart as the assessment connoisseur as policy adviser: he listens and provides feedback that is adjusted to the needs of the different educational policy stakeholders, including teachers, school leaders and those responsible as civil servants for leading, managing and delivering the assessment for learning policy agenda.

<sup>7</sup> A point well-made in the empirical study by Hopfenbeck, Flórez Petour & Tolo (2015).

<sup>8</sup> The translations from Norwegian to English are my own.

<sup>9</sup> Personal correspondence received by email June 2025 from Ida Large, Branch Director the Norwegian Directorate of Education and Training.

However, it is important to note that Stobart's participation in the Norwegian policy framework was a timely "hand in glove" moment to address what had been a policy challenge identified from at least 1987. Engh and Gran (2021) looked back and noted the Curriculum Plan from this year had highlighted a lack of a clear relationship between teaching, learning and assessment. Specifically, how some students were perceived as failing even though they were trying their best and were actually progressing in their learning.

Large and her colleagues also highlight a theme found in the book *Testing Times. The Uses and Abuses of Assessment* (2008). In their own words, the importance of how assessment can contribute to the development of a student's self-identity (my underlining)<sup>10</sup>:

Gordon's primary focus was always on student learning. At the same time, he highlighted how assessment is connected to several other factors, such as the importance of formative assessment for a good learning environment, *the development of students' self-esteem*, and the value of working formatively for the inclusion of struggling students. In addition, Gordon's contributions built on the importance of sharing good practice in the school and the collective development of teachers' assessment culture and competence (my emphasis).

The Norwegian Directorate for Education and Training's report (2018) details the intentions of the national policy program highlights a) a cross-curricula approach capable of drawing upon assessment for learning principles including self-assessment, self-regulated learning and pupil and apprentice involvement and b) "to support local school authorities, schools and training establishments in developing more formative assessment practices and cultures. We sought to achieve this by raising skills levels and awareness of assessment as a tool for learning" (2018, p. 6).

In our context, Stobart as the assessment connoisseur as policy adviser provided advice and support to participants in this national education project. The principles of self-assessment, self-regulated learning and pupil and apprentice involvement in their own self-assessment belong to the particular language game of this Assessment Reform in Norway; but also overlapping with the language game of the Assessment Reform Group in the UK. In the seven cohorts selected by the Directorate between 2010 and 2018, each local authority took part in the Directorate's program of activities and received between 16 and 24 months of funding. To unlock this a local authority had to guarantee the participation of three primary/lower secondary schools, while

<sup>10</sup> As Hanson noted and cited by Stobart: "The individual in contemporary society is not so much described by tests as constructed by them".



the corresponding county level ensured the participation of five upper secondary schools and three vocational training institutions.

Was the Norwegian policy a success? Many participants in their feedback to the Directorate indicated a) experience of a more learning driven form of assessment and b) students who identified greater involvement in self-assessment (through conversations with teachers that were more dialogical and feedback oriented and not one way emphasising summative assessment).

The national initiative had a significant reach: 310 local municipalities (approx. 73% of the total number all in the country) and all of Norway's 19 counties i.e. 1,500 schools and 55 vocational training establishments, along with 96 adult education centres across 60 municipalities. The Directorate in its report notes how the OECD commended them for not adopting the expert role, "instead making it clear that participants at all levels were involved in a shared process of learning and development... several school leaders found that they were engaged in a genuine dialogue and partnership with the Directorate rather than a top-down process" (2018, p. 10).

Even though the policy did not espouse civil servant as 'experts,' it did have the vital support of Stobart and at times other international assessment for learning experts, such as Dylan Wiliam. Also, let us return to how in the implementation of the Assessment for Learning framework it evidenced cultural aspects at play in a country such as Norway. For example, the Norwegian education system, unlike in many other parts of the world is funded solely by the taxpayer. In Australia, to take an opposite example, there are private, Catholic and state-run schools; each with its own special characteristics, mission statements and levels of funding contributions from parents, state or federal funding bodies.

Surely, this would have meant the language games of the Assessment Reform Group and those voiced by Stobart were more easily shared and implemented evenly across all participants – all are employees of a basically tax funded school system? The report might be interpreted in such a way, suggesting as it does that there is evidence of a greater shared teacher understanding of the connection between learning and assessment. And yet, researcher evaluation of the policy recommends, "continuing to improve the quality of the language teachers use when talking about assessment with their pupils – a language the pupils understand, and which can serve as a tool for both teacher and pupil in the pupils' learning process" (2018, p. 20). Simply put the level of understanding of the language games of assessment for learning can differ significantly between participants; even in Norway after this national assessment policy seeking to raise assessment skills and understanding. As Wittgenstein, suggested there can be so many different streets

and houses in a town or city and they can accumulate over time organically, rather than uniformly in a planned grid and cross fashion.

#### 4. Summary

Stobart personifies his own argument inspired by Hanson, “that the individual in contemporary society is not so much described by tests as constructed by them”. Let us remember that he began his career as a classroom teacher and this has given him lifelong insights into the needs of the teacher and the student in the classroom or other settings, such as a doctorate student (my needs at one point). Accordingly, it has equipped him well in his role as a connoisseur imparting advice to teachers and by extension to learners. Similarly, working in higher education and in writing books that contribute to assessment research in different forms (e.g. the expert learner and the uses and abuses of assessment) he has been a researcher as assessment connoisseur able to offer timely advice to Norwegian academics.

Stobart as an assessment connoisseur as policy adviser was never interested in the adoption of assessment for learning in exactly the same way by all participants. He was too experienced a policy adviser to believe something so simplistic. In his lectures and policy advice he returned again and again to the following point: “Communicating the principles of assessment for learning is like putting them on a piece of paper in a bottle. Throwing it in the sea. And waiting for whoever picks it up when it is washed up on the beach”.

Like the peanuts cartoon above, just because you have been told to learn, it does not mean you will learn it.

It might be postulated that in recognition of this kind of argument, the Norwegian Directorate for Education and Training learnt from its first cohort of participants who felt they had not received enough support. In 2015, they approached the Centre for Lifelong Learning<sup>11</sup> in what is now Inland Norway University and asked them to develop a number of assessment for learning MOOCs in Norwegian open to all who were interested in the topic.

As I write today, we hear little of MOOCs and this kind of old wine in new bottles is re-packaged as “micro-credentials” (or micro-programs) of differing lengths – from an hour to longer, even over 50 hours. But arguably, it is still the same dropping of a bottle in the sea with a message – let us call

<sup>11</sup> The team led by Egil Hartberg, the co-author of *Feedback i skolen* (Feedback in the School, 2012) continue to this day to work in a number of assessment related policy implementation initiatives in Norway.

it a language game of assessment – and waiting for others to take it up and make up their own minds on what it means. Stobart if he was making this very point would for sure have referenced a famous pop song released by The Police in 1979: *Message in a Bottle*<sup>12</sup>. We might be alone, a castaway on an island, but assessment for learning is the message hoping to join us to others in sharing stories and language games of assessment.

## References

- Benjamin, W. (1973), The storyteller: Reflections on the Works of Nikolai Leskov. In H. Arendt (ed.), *W. Benjamin, Illuminations*. Waukega: Fontana Press, pp. 83-109.
- Bennett, R. (2011), “Formative Assessment: A critical Review”, *Assessment in Education: Principles, Policy and Practice*, 18 (1), pp. 5-25, <https://doi.org/10.1080/0969594X.2010.513678>.
- Black, P., Wiliam, D. (1998), “Assessment and Classroom Learning”, *Assessment in Education: Principles, Policy & Practice*, 5 (1), pp. 7-74, <https://doi.org/10.1080/0969595980050102>.
- Black, P., Wiliam, D. (2009), “Developing the Theory of formative assessment”, *Educational Assessment, Evaluation and Accountability*, 21 (1), pp. 5-31, <https://doi.org/10.1007/s11092-008-9068-5>.
- Day, C., Stobart, G., Sammons, P., Kington, A., Gu, Q., Smees, R., Mujtaba, T. (2006), *Variations in Teachers’ Work, Lives and Effectiveness*, DfES Research Report No. 743.
- Dobson, S. (2009), “Review of Testing Times. The Uses and Abuses of Assessment, by Gordon Stobart”, *Nordic Studies in Education*, 30 (1), pp. 63-71, <https://www.scup.com/doi/10.18261/ISSN1891-5949-2010-01-05>.
- Dobson, S. (2017), *Assessing the Viva in Higher Education. Chasing Moments of Truth*. Cham: Springer.
- Dobson, S., Fudiyartanto, F. (2023), *Transforming Assessment in Education – The Hidden World of Language Games*. Cham: Springer.
- Eisner, E. (1998), “What do the Arts teach?”, *Improving Schools*, 1 (3), pp. 32-36.
- Engh, R. Dobson, S., Høihilder, K. (2007), *Vurdering for Læring* (Assessment for Learning). Oslo: Cappelen Damm Akademisk.
- Engh, R., Gran, L. (2021), *Vurdering for læring i skolen. På vei mot en bærekraftig og demokratisk læringskultur* (Assessment for Learning in Schools. Towards a Sustainable and Democratic Learning Culture). Oslo: Cappelen Damm Akademisk.
- Gardner, J., Harlen, W., Hayward, L., Stobart, G. (2008), *Changing Assessment Practice: Process, Principles and Standards*, <https://www.nuffieldfoundation.org/project/analysis-and-review-of-innovations-in-assessment-aria>.

<sup>12</sup> The lyrics of the song: <https://www.sting.com/discography/lyrics/188>.

- Gulson, K., Lewis, S., Lingard, B., Lubienski, C., Takayama, K., Webb, T. (2017), "Policy mobilities and methodology: a proposition for inventive methods in education policy studies", *Critical Studies in Education*, 58 (2), pp. 224-241.
- Hanson, F. A. (1994), *Testing Testing: Social Consequences of the examined life*. Oakland: University of California Press.
- Hartberg, E., Dobson, S., Gran, L. (2012), *Feedback i skolen* (Feedback in the school). Oslo: Gyldendal.
- Hopfenbeck, T. N., Flórez Petour, M. T., Tolo, A. (2015), "Balancing tensions in educational policy reforms: Large-scale implementation of assessment for learning in Norway", *Assessment in Education: Principles, Policy & Practice*, 22 (1), pp. 44-60.
- Klenowski, V. (2011), "Assessment for learning in the accountability era: Queensland, Australia", *Studies in Educational Evaluation*, 37 (1), pp. 78-83.
- Norwegian Directorate for Education and Training (2018), *Observations on the National Assessment for Learning Programme (2010-2018). Skills development in networks*, [https://www.udir.no/contentassets/0b8163f632bd48ea9552ae3693513654/the-norwegian-assessment-for-learning-programme\\_final-report-2018.pdf](https://www.udir.no/contentassets/0b8163f632bd48ea9552ae3693513654/the-norwegian-assessment-for-learning-programme_final-report-2018.pdf).
- Palamara, D. R. (2016), "Social status in Norway and the law of 'jante': An analysis of ISSP social inequality data", *European Journal of Interdisciplinary Studies*, 8 (1), pp. 1-21, <https://www.proquest.com/docview/1879179899?pq-origsite=primo&searchKeywords=Social+status+in+Norway+and+the+law+of+%27jante%27%3A+A+n+analysis+of+ISSP+social+inequality+data&sourcetype=Scholarly%20Journals>.
- Polanyi, M. (2009), *The tacit dimension*. Chicago: University of Chicago Press.
- Sadler, D. R. (2012), "Assuring academic achievement standards: from moderation to calibration", *Assessment in Education: Principles, Policy and Practice*, 20 (1), pp. 5-19, <https://doi.org/10.1080/0969594X.2012.714742>.
- Schivelbusch, W. (2014), *The railway journey: the industrialization of time and space in the nineteenth century*. Oakland: University of California Press.
- Skogvoll, V., Dobson, S. (2011), Connoisseuren – med blikket for øyeblikket. Et kroppsfenomenologisk essay om det profesjonelle (Connoisseur – awareness of the moment. A corporeal phenomenological essay about professionalism). In Ø. Haaland, S. Dobson, G. Haugsbakk (Eds), *Pedagogikk for en ny tid* (Education a new time). Son: Oplandske Publishers, pp. 161-173.
- Stobart, G. (2008), *Testing Times. The Uses and Abuses of Assessment*. London: Routledge.
- Stobart, G. (2014), *The Expert Learner. Challenging the Myth of the Expert Learner*. Milton Keynes: Open University Press.
- Wiliam, D. (2009), *Assessment for Learning: Why, What and How?*. London: Institute of Education, University of London.
- Wittgenstein, L. (1953/1967), *Philosophical investigations*. New York: Macmillan.
- Wolf, A. (2011), *Review of vocational education – the Wolf Report*, [www.gov.uk/government/uploads/system/uploads/attachment\\_data/file/180504/DFE-00031-2011.pdf](http://www.gov.uk/government/uploads/system/uploads/attachment_data/file/180504/DFE-00031-2011.pdf).