



Assessing Sustainability Competences: A Discussion on *What and How*

Francesca Farioli, Michela Mayer, Rick Millican, Aaron Redman, and Paul Vare

Abstract

This chapter comprises a discussion between three international teacher training contexts in Italy, the USA and the UK and explores the tensions that exist in assessing education for sustainable development (ESD) competences. It recognises the need for a form of assessment that is consistent with the aims and values of education for sustainability and that acknowledges the complexity of competences in this field. Debates around the concept of measurement are considered as well as what it is that is being assessed and how judgements are made. Ultimately, the lack of an evidence-based, ‘perfect’ method is acknowledged along with the need for further research to find an approach that might be applied consistently and reliably in order to provide evidence that

those assessed have the competence to make change happen.

Keywords

Education for sustainable development (ESD) · Sustainability · Competence · Competency · Assessment · Measurement

This chapter comprises a discussion between three international teacher training contexts in Italy, the USA and the UK and explores the tensions that exist in assessing education for sustainable development (ESD) competences.

Paul/Rick Colleagues, as you know, all of us have been using educator competences as a way of introducing ESD to educators, trainers and student teachers. Here at the University of Gloucestershire where we have been working with ESD competences over the past four years, we have seen many positive outcomes for our undergraduates. These include increased knowledge, growing self-confidence and a range of positive actions that they have taken. Given that we intend ESD competences to help bring about

F. Farioli · M. Mayer
IASS-Italian Association for Sustainability Science,
Rome, Italy

R. Millican · P. Vare
University of Gloucestershire, Cheltenham, UK

A. Redman (✉)
School of Sustainability, College of Global Futures,
Arizona State University, Tempe, AZ, USA
e-mail: aaron.redman@asu.edu

individual—and ultimately societal—change, we feel this is a good start.

One aspect that we have wrestled with, however, is how to assess these outcomes effectively. As Kerry Shephard illustrates in Chap. 6, the assessment of ESD competences is not as straightforward as might first appear. The whole point of ESD is to stimulate some form of intrinsic motivation to make more sustainable life choices but as soon as we confer credit for demonstrating such outcomes, we immediately introduce a strong extrinsic motivation for students to claim to have made sustainable life choices in a performative manner.

We are aware that we have been tackling this in different ways so we are interested to learn about the approaches you have taken. We, Paul and Rick, know that you worry about the idea that assessment conveys the notion of measurement which you feel is inappropriate. Can you explain why?

Michela/Francesca One thing we should be clear about from the outset is that we need to look for an assessment process which takes account of the *complexity* of ESD and explores the *quality* of the processes and transformations implemented.

The fact that we may define criteria and assign numbers to the observed outcomes and by consequence can order them according to a numerical scale, does not make this a measurement. For instance, Intelligence Quotient (IQ) is not a ‘measurement’, even if it is still in use despite much criticism (e.g. Gould 1981), just as the Michelin stars of restaurants or the ratings expressed on websites are not measurements.

The paradigm that should inspire educational research and practice cannot be the ‘Galilean’ or ‘Newtonian’ one which seeks to simplify and quantify the complexity of the real world in order to establish objective relationships, but rather one like that proposed by Ginzburg (1989) for human sciences: an ‘evidential’ or ‘circumstantial’ paradigm (*‘paradigma indiziario’*). In this paradigm, small differences and small signs enable the historian, the psychologist, the investigator and the

educator, to rationally reconstruct and understand what has probably happened. The aim is not to find simple general rules or to collect defined outcomes but to reconstruct transformation histories that are intrinsically unique.

The consideration of educational evaluation as a measurement has been criticised several times in the past. For example, within environmental education Flogaitis and Liriakou (2000), following a proposal of Robottom and Hart (1993), denounced the predominance in educational evaluation of a positivistic paradigm and proposed a socio-critical paradigm where reality is conceived as a complex matter, knowledge is socially constructed and evaluation is one of the instruments of change. Here, the evaluator is viewed as a social agent of change and uses their judgement, based on stated and shared criteria, to support the transformation process.

Aaron Exactly! There are serious pitfalls when tackling assessment but, in my view, this argues for investing in it rather than ignoring it and the time is ripe for making a serious effort to develop legitimate assessment processes for ESD. Not only is there agreement about the goal of ESD being to support sustainability transformations (Franco et al. 2019), but there is an increasing convergence around specifying what the learning objectives should be for the students who are to be educated in facilitating these transformations (Brundiers et al. 2021; see also Chaps. 3 and 4 of this book). This is important progress and if a university wished to start a degree programme in ESD, they would have a solidly informed base from which to articulate its intended outcomes. However, if they were to ask HOW they should structure their curriculum or their teaching in order to achieve these outcomes, the field of ESD research would have little empirical evidence to offer.

There are certainly plenty of case studies about exciting and interesting programmes, courses and university efforts (Weiss and Barth 2019), but these remain largely descriptive. The intention of the Educating Future Change Agents (EFCA) project (Redman 2020) was to move

beyond this and utilise cases to build empirical evidence about how we can achieve competencies in sustainability. It was immediately apparent that a critical component of that effort was going to be properly and rigorously assessing competencies. Such assessments would indicate whether certain teaching approaches, or curriculum structures, were more or less effective and make empirical comparisons between cases possible. Yet when we looked into the existing research, it offered little guidance and relied heavily on students' self-assessment of their own competence, an approach whose weaknesses are well catalogued (Redman et al. 2021).

As Michela and Francesca point out though, the last thing that we want to fall into is simplifying ESD down to something that can be assessed with a standardised multiple-choice test. Fortunately, one of the advantages of using a competency-based approach for learning objectives is that it preserves that real-world complexity in a way that foils the traditional modes of assessment (Frey and Hartig 2009). Yet, neither can we throw our hands up and say they are impossible to assess. The field of ESD has long argued that novel teaching approaches are vital (Frisk and Larson 2011), yet if we cannot provide evidence (via assessment) that these methods are achieving their stated goals of developing sustainability change agents who can facilitate transformations, then we should (rightfully) expect our calls for these innovative pedagogies to be increasingly ignored.

Paul/Rick It seems that we all feel similarly that assessment of competence in ESD is important so that we can provide evidence of the effectiveness of our teaching. However, the challenge is how to do that in a way that is constructively aligned with our defined outcomes and pedagogical approach.

This notion of transformation complicates matters. If the ultimate aim of ESD is the transformation of society, then presumably assessment should be assessing the non-linear interactions that would need to take place over an extended period of time to see if a given pro-

gramme of educator preparation had led to corresponding classroom actions, which in turn led to their students adopting positive behaviours and attitudes commensurate with a more sustainable society.

Clearly it is not feasible to assess this whole process, in which case we have to determine what we *can* look for and decide whether that can provide sufficient evidence to suggest that our teacher education has made this transformation more likely to occur. In other words, we are seeking specific ingredients that, if evidence can be found for them, would convey the *likelihood* of transformation emerging, possibly over time, at the level of future learners and eventually at system level.

Given that our work is focused on training the educators, then presumably we need to look for this evidence in, or from, them. Demonstration of ESD competences however, is unlikely to be sufficient because they may have *competency* (the ability), but not the *competence* (to put it into practice).

A broad concept of competence (see the discussion in the Introduction of this book) suggests more than just ability, it also encompasses the values that would lead to the motivation to apply their learning and the *agency* to be able to action them. On this last point, Campbell (2009) identifies two types of agency:

Type 1: the ability to operate freely at the individual level, albeit within existing structures. Campbell terms this the *power of agency*; we might also call this competency.

Type 2: the ability and confidence to make changes in the face of structures thereby contributing to societal change—what Campbell terms *agentic power*; we would see this as fully operational competence.

Given the constraints that bind many educators, e.g. operating within prescribed curricula and tightly controlled standardised assessments, the options for Type 2 agency seem limited, yet this is exactly what is needed, a willingness to find—or create—the 'wobble room' required to open up possibilities for implementing their own ESD competences.

To add further complication, if we truly wish to encourage critical thought, autonomy and agency, then we cannot control how this agency will be used and where the critical thinking might lead; indeed, critical thinking has its own dangers (see Chap. 7). As a consequence, it feels like the best we can hope for is to take the ‘evidential’ approach, as suggested by Michela and Francesca, and look for evidence or indicators of (a) competency, that is the *ability* to do things as an individual and (b) competence, which for us includes the *willingness* of student teachers and educators to find opportunities to demonstrate this ability despite the constraints of their professional context.

If we can find evidence of these elements, then perhaps we can infer from this that the ingredients necessary to achieve societal change are present *insofar* as we are able to instil them. The extent to which these changes actually occur *and* contribute towards a more sustainable future will always be determined by the agency of individuals who are subject to a variety of often unforeseen constraints and influences at personal, professional and societal levels.

Trusting this clarifies what we are looking for, we are left with the question of how we find that evidence.

Aaron, you have researched various ways of assessing competences, what did you discover?

Aaron Our systematic review of the literature revealed that the body of research on assessing competences in sustainability has grown rapidly in recent years (Redman et al. 2021). Yet despite this growth, the field is clearly still in its infancy and offers little empirical guidance for either practitioners or researchers interested in effective assessment. There are several ways in which current practice (at least as evidenced in the literature) is hobbled. The first is an underinvestment in development of tools for assessment. This manifests in individual studies where assessment serves merely to produce data about some kind of pedagogical or curricular innovation, as well as in the fact that there are few instances where research groups are building on each other’s (or even their own)

work. Perhaps driven by this underinvestment, the most widely used tool is the weakest, scaled self-assessment, used in well over half of the studies.

But secondly, more fundamentally what we saw hampering effective assessment was touched on by Paul and Rick, which is properly aligning assessment with the desired outcomes of ESD. As they pointed out, the outcome of leading transformations is too ambitious to possibly capture in one (or many) assessments (if possible at all). Currently, this challenge is hand-waved away with limited assessments being used to make broad statements about competence development. However, what is needed is to be explicit about what specific pieces of the overall outcome you are intending to assess and then utilise the approaches which give you the best evidence about those specific pieces. While dispersed, the initial indications of what tools might be best in what circumstances and to measure what components of competence, do exist in the current research.

The typology of tools, which we distilled from the literature, brings together the findings of the field to enable one to evaluate options when selecting assessment tools. We identified eight distinct types of tools: scaled self-assessment, reflective writing, scenario/case test, focus group/interview, performance observation, concept mapping, conventional test and regular course work. These can be clustered into three meta-types, namely self-perceiving-based, observation-based and test-based assessment procedures. This typology provides a framework on which we can layer more findings, explore additional tools and identify the best assessment approach for our specific purposes.

Michela/Francesca This range of assessment approaches is interesting for us and we have actually used many of them for assessing the learning process in our context. However, the identification of appropriate assessment tools, as Aaron states, strongly depends on the main purpose(s) we aim to achieve with our educational project

and on what type of change we expect to stimulate through it.

Therefore, we probably need to start again from the initial question posed by Paul and Rick: WHAT are we aiming to assess? In our context the aim of the assessment was to evaluate the training of educators as change agents by looking for the demonstration of competences in their professional contexts, as Paul and Rick have underlined and as we have explored in detail in Chap. 11.

An important goal in our training programmes was that our learners became more attentive to the complexity of the world and of the educational processes, became competent (in the sense of competency), but were also willing to use their competences to address future uncertainties.

Now, returning to the question posed by Paul and Rick about HOW to assess competences and how to find evidence, the key, for us, is to look at the range of tools and approaches introduced by Aaron not as separate, but interwoven: if self-assessment is fundamental, for example, to look at the increase in consciousness, it is the interweaving with other tools, i.e. focus groups and peer evaluation of reflective writing of experiences, which can provide a more complete and ‘three-dimensional picture’ of what we are trying to assess.

Another important aspect to consider in relation to how to find evidence is the necessity to focus attention not only on *expected* results, but also on *emerging, unforeseen* outcomes in order to detect the changes instilled by the educational process which is by nature complex and dynamic.

Consequently, in our training programme in Italy (see Chap. 11) we tried to ‘follow the transformation’ and understand: (1) if we were successful in promoting in our learners a change in their vision and beliefs about ESD and the educators’ role, and if so (2) in which direction and (3) with what level of consciousness this was occurring. This is because we think that the willingness to put competences into practice cannot be generated without an increase in the awareness of being change agents.

To answer these three questions, we had to follow the transformation process from its inception while remaining open to the detection of unforeseen elements and signals. All this had to align with the specific competence framework’s underlying values and concepts (Farioli and Mayer 2020).

In order to gather the required evidence, we used an ‘environmental autobiography’ tool at the start of each training programme. This told us how learners felt about their role as sustainability educators as well as about their emotions, willingness and potentialities to engage as change agents.

It was only by knowing learners’ starting points that we could understand, by the end of the course, the change that had actually occurred and the extent to which this could be attributed to the course itself, rather than being an outcome of the competences and knowledge that the learners already possessed.

Thereafter, the use of interwoven assessment tools, i.e. observation of assigned tasks, analysis and peer assessment of reflective writing and construction of individual portfolios that mapped evidence of experiences and competences achieved at different stages in the process, has been crucial for us in order to ‘follow the transformation’.

Each of the tools that were used captured only some aspects of the learning, but it was the interlacing of the results that provided us with a more accurate and complete picture of the changes that were taking place. For example, a storytelling analysis exercise, carried out in groups, was fruitfully piloted in connection with peer assessment and focus groups and allowed each of our learners to ‘look at themselves from the point of view of others’ and to ‘reflect into others’.

The iterative process of practice (in the sense of carrying out assigned tasks during the course), reflection in action and challenge by others has—and this is for us the most important result we have achieved—favoured a path of consciousness in our learners of the competences that they have developed and acted on and of those that they have yet to develop, improve and put into practice.

The challenge however is in how to tie together the clues and evidence which emerge from the different assessments in place and how to interweave them in order to build a consistent framework for an overall assessment with a rounder sense of purpose. Such a framework should not aim to be an ‘objective’ assessment, since it is never ‘neutral’; even in test-based procedures such as the Programme for International Student Assessment (PISA), subjectivity is always present, for example in the selection, however negotiated, of the questions asked. However, as our experience has demonstrated, it should aim to be useful for learners, allowing them to look at themselves and their professional path with new eyes and to feel more confident in their acquired competences, all of which will probably render them more capable of instigating change.

Quite which clues and indicators are to be used for evidence, and HOW to best combine them for a meaningful and appropriate assessment approach, is the key to a ‘quality assessment’, as well as for assessing the ‘quality of change’. In our opinion, this remains one of the issues in ESD research that requires further investigation and empirical evidence.

Paul/Rick What has become apparent from these contributions is that no single assessment tool can capture the complexity of what we are aiming to assess and that there is no perfect solution. In our own work, we have also relied on a mixed methods approach.

In our context, student teachers need to demonstrate the achievement of specific academic standards in order to be awarded credit *as well as* develop competences. However, we have been fortunate to run a non-credit bearing, competence-based programme for four years and this provided the opportunity to develop our assessment approach before extending the programme to accredited courses.

We asked participants to keep a reflective journal throughout the programme outlining how they had applied the competences in their profes-

sional, social and/or private life and, where applicable, how they had helped develop the competences in others. A thematic analysis of these journals provided evidence of three key outcomes:

- Understanding of the competences and the issues they raise
- Action taken on the basis of the competences
- Reflection on the competences themselves and on their own engagement with them

Each of these outcomes were broken down further into three sub-categories or ‘learning aspects’, which we have listed elsewhere (Vare and Millican 2020). In our case, we were working with the twelve Rounder Sense of Purpose (RSP) competences and recognised that seeking evidence of nine learning aspects for each competence would be unrealistic and would sacrifice depth of engagement for breadth of coverage. We decided that a meaningful indicator of the *extent* of a student’s learning across all twelve competences would be if they provided evidence of at least four of the nine learning aspects under each competence, with at least one in each category (Understanding, Action & Reflection). We also sought evidence of each of the nine aspects in at least four competences.

Unsurprisingly, analysis of students’ reflective journals revealed qualitative differences in the *depth* of engagement or levels of ability. Using exemplar statements from the journals we drafted descriptors for different levels of achievement in relation to the nine aspects of learning. This enabled us to create a marking grid similar to those used by colleagues on our other accredited courses. By shading the ‘best fit’ descriptors for each aspect of learning, an assessor builds an impression of a learner’s competence; this allows for a composite grade to be reached thus fulfilling university requirements.

Used together, these two tools can be used to assess a range of evidence including reflective journals, videos and formal essays. While any form of assessment will give an incomplete picture and be based on the professional judgement of the assessor, we hope that this balance between

extent (quantity) and *depth* (quality) can go some way towards assessing competence and competency as well as an indication of transformation with its promise of sustained change.

Aaron As already mentioned, for the EFCA project we found a need to develop our own assessment approach and ultimately deployed ten different tools which spanned the whole range of types described earlier. Similar to Paul/Rick and Francesca/Michela, we were able to make the most robust assessments of students' competencies by triangulating results of different assessments in cluster 1 (student self-perception). One particularly strong approach was to ask students to rate their level of competency and write a short justification (Birdman et al. 2021) which was then used as a starting point for interviews. This process was repeated four times throughout a two-year degree programme and gave a robust overview of the students' individual development, but did not enable a comparative 'measurement' or empirical comparison between students.

Two studies also attempted to construct domain-specific, yet holistic, test-based assessments and significant time was invested in developing and piloting these tools. One of the tests used real-world curriculum and expert judgement to assess the students' Pedagogical Content Knowledge (Brandt et al. 2019), while for the other a full in vivo simulation was run inspired by advanced approaches taken in medicine (Howley 2004). In this live simulation, students were confronted with a mock city council to which they had to offer their advice on the sustainability of an economic development plan (Foucrier 2020). These assessments gave insight into the competency development of the group, but little in terms of tracking individuals. They also suffered from the fact that they were not 'graded', which certainly influenced student effort. As other studies have found, variance in scores may be largely driven by effort invested (Zamarro et al. 2019).

In conclusion, our three cases give an example of the rich variation of assessment

approaches being taken with ESD, but highlight the critical need for a more comprehensive and coordinated approach to be taken. The experience at EFCA was that even with significant resources and explicit support from both institution and instructors, it was not possible to administer a consistent and robust set of assessments across its studies. This chapter therefore serves as a starting point for the necessary conversation between academics and practitioners to both learn from, and build upon, each other's work in order to develop holistic and effective approaches to assessing students' development of competency, so that they can be effectively supported to become the change agents that the world needs.

References

- Birdman, J., Redman, A., & Lang, D. J. (2021). Pushing the boundaries: experience-based learning in early phases of graduate sustainability curricula. *International Journal of Sustainability in Higher Education*, 22(1), 237–253. <https://doi.org/10.1108/IJSHE-08-2019-0242>
- Brandt, J. O., Bürgener, L., Barth, M., & Redman, A. (2019). Becoming a competent teacher in education for sustainable development: Learning outcomes and processes in teacher education. *International Journal of Sustainability in Higher Education*, 20(4), 630–653. <https://doi.org/10.1108/IJSHE-10-2018-0183>
- Brundiars, K., Barth, M., Cebrián, G., Cohen, M., Diaz, L., Doucette-Remington, S., Dripps, W., Habron, G., Harré, N., Jarchow, M., Losch, K., Michel, J., Mochizuki, Y., Rieckmann, M., Parnell, R., Walker, P., & Zint, M. (2021). Key competencies in sustainability in higher education—toward an agreed-upon reference framework. *Sustainability Science*, 16(1), 13–29. <https://doi.org/10.1007/s11625-020-00838-2>
- Campbell, C. (2009). Distinguishing the Power of Agency from Agentic Power: A Note on Weber and the “Black Box” of Personal Agency. *Sociological Theory*, 2009;27(4):407-418. <https://doi.org/10.1111/j.1467-9558.2009.01355.x>
- Farioli, F. & Mayer, M. (2020), *Transformative Evaluation for Sustainability Educators: the experience of a Competence-based Educational Research project*, 2nd GUNI International Conference on SDGs: Higher Education & Science Take Action, Summary Report, p.101 - 104
- Flogaitis, E. & Liriakou, G. (2000). Quelle évaluation pour quelle Education relative à l'Environnement? [What Evaluation for What Environmental

- Education?] Education relative à l'Environnement, *RegardRechercheRéflexions*, 2, pp. 13–30
- Foucrier, T. (2020). *Training future entrepreneurs – developing and assessing sustainability competencies in entrepreneurship education* (Order No. 27955218). Available from Dissertations & Theses @ Arizona State University; ProQuest Dissertations & Theses Global. (2406963829)
- Franco, I., Saito, O., Vaughter, P., Whereat, J., Kanie, N., & Takemoto, K. (2019). Higher education for sustainable development: actioning the global goals in policy, curriculum and practice. *Sustainability Science*, 14(6), 1621–1642. <https://doi.org/10.1007/s11625-018-0628-4>
- Frey, A., & Hartig, J. (2009). Assessment of competencies. *Studies in Educational Evaluation*, 35(2–3), 55–56. <https://doi.org/10.1016/j.stueduc.2009.10.001>
- Frisk, E., & Larson, K. L. (2011). Educating for Sustainability: Competencies & Practices for Transformative Action. *Journal of Sustainability Education*, 2 (March).
- Ginzburg C. (1989). *Clues: Roots of an Evidential paradigm in Clues, Myths, and the Historical Method*, The John Hopkins University Press
- Gould, S. J., (1981) *The Mismeasure of Man*. New York: W. W. Norton & Company.
- Howley, L. D. (2004). Performance assessment in medical education: Where we've been and where we're going. In *Evaluation and the Health Professions*. <https://doi.org/10.1177/0163278704267044>
- Redman, A. (2020). *Assessing the Development of Key Competencies in Sustainability* [Arizona State University]. In ProQuest Dissertations and Theses. <http://login.ezproxy1.lib.asu.edu/login?url=https://www.proquest.com/docview/2445352259?accountid=4485>
- Redman, A., Wiek, A., & Barth, M. (2021). Current practice of assessing students' sustainability competencies: a review of tools. *Sustainability Science*, 16(1), 117–135. <https://doi.org/10.1007/s11625-020-00855-1>
- Robottom, J. & Hart, P. (1993). *Research in environmental education. Engaging the debate*. Victoria, Deakin University
- Vare, P and Millican, R (2020) Educating for a 'Rounder Sense of Purpose' - Lessons learned from practice: assessing ESD competences. Presentation to TEESNet Conference, 2020. Accessed January 2021 at: <https://eprints.glos.ac.uk/id/eprint/10434>
- Weiss, M., & Barth, M. (2019). Global research landscape of sustainability curricula implementation in higher education. *International Journal of Sustainability in Higher Education*, 20(4), 570–589. <https://doi.org/10.1108/IJSHE-10-2018-0190>
- Zamarro, G., Hitt, C., & Mendez, I. (2019). When students don't care: Reexamining international differences in achievement and student effort. *Journal of Human Capital*. <https://doi.org/10.1086/705799>