



**OFFICIAL REPORT**  
AITHISG OIFIGEIL

# Education and Skills Committee

**Wednesday 9 January 2019**

**Session 5**



The Scottish Parliament  
Pàrlamaid na h-Alba

© Parliamentary copyright. Scottish Parliamentary Corporate Body

Information on the Scottish Parliament's copyright policy can be found on the website - [www.parliament.scot](http://www.parliament.scot) or by contacting Public Information on 0131 348 5000

---

**Wednesday 9 January 2019**

**CONTENTS**

	<b>Col.</b>
<b>INTERESTS</b> .....	1
<b>DECISION ON TAKING BUSINESS IN PRIVATE</b> .....	2
<b>SCOTTISH NATIONAL STANDARDISED ASSESSMENTS INQUIRY</b> .....	3

---

**EDUCATION AND SKILLS COMMITTEE**

**1<sup>st</sup> Meeting 2019, Session 5**

**CONVENER**

\*Clare Adamson (Motherwell and Wishaw) (SNP)

**DEPUTY CONVENER**

\*Johann Lamont (Glasgow) (Lab)

**COMMITTEE MEMBERS**

\*Dr Alasdair Allan (Na h-Eileanan an Iar) (SNP)

\*Jenny Gilruth (Mid Fife and Glenrothes) (SNP)

\*Iain Gray (East Lothian) (Lab)

\*Ross Greer (West Scotland) (Green)

Gordon MacDonald (Edinburgh Pentlands) (SNP)

\*Rona Mackay (Strathkelvin and Bearsden) (SNP)

\*Oliver Mundell (Dumfriesshire) (Con)

\*Tavish Scott (Shetland Islands) (LD)

\*Liz Smith (Mid Scotland and Fife) (Con)

\*attended

**THE FOLLOWING ALSO PARTICIPATED:**

Professor Sue Ellis (University of Strathclyde)

Juliette Mendelovits (Australian Council for Educational Research)

Professor Christine Merrell (Durham University)

Gil Paterson (Clydebank and Milngavie) (SNP) (Committee Substitute)

Mhairi Shaw (Association of Directors of Education in Scotland)

**CLERK TO THE COMMITTEE**

Roz Thomson

**LOCATION**

The Robert Burns Room (CR1)



## Scottish Parliament

### Education and Skills Committee

Wednesday 9 January 2019

*[The Convener opened the meeting at 10:00]*

### Interests

**The Convener (Clare Adamson):** Welcome to the first meeting in 2019 of the Education and Skills Committee. I wish everyone a happy new year.

I remind everyone to turn mobile phones and other devices to silent in case they interfere with the broadcasting system.

Agenda item 1 is a declaration of interests. We have received apologies from Gordon MacDonald, whose substitute is Gil Paterson. Mr Paterson is attending the committee for the first time, so I welcome him and invite him to declare any interests.

**Gil Paterson (Clydebank and Milngavie) (SNP):** I refer to my interests that are on the public record. I have no additions to that with regard to my attendance here.

## Decision on Taking Business in Private

10:00

**The Convener:** Agenda item 2 is a decision on whether to take agenda item 4 in private and whether to take future consideration of evidence on the Scottish national standardised assessments inquiry in private. Are members content to do so?

**Members** *indicated agreement.*

**The Convener:** Thank you.

## Scottish National Standardised Assessments Inquiry

10:01

**The Convener:** Agenda item 3 is the Scottish national standardised assessments inquiry. This is the first week of our inquiry and we will hear from a panel of witnesses that includes people who are involved in designing and delivering standardised assessments. I welcome Mhairi Shaw, director of education, East Renfrewshire Council, on behalf of the Association of Directors of Education in Scotland; Juliette Mendelovits, director of assessment and reporting for the Australian Council for Educational Research; Professor Sue Ellis, professor of education at the University of Strathclyde; and Professor Christine Merrell, professor in the school of education and deputy head of the faculty of social sciences and health at Durham University.

Please will you state briefly your involvement in SNSAs? We will then move to detailed questions from members. We will start with Mhairi Shaw.

**Mhairi Shaw (Association of Directors of Education in Scotland):** Good morning, convener. I am a director of education and oversee SNSAs in East Renfrewshire; I am not directly involved in them. As a member of ADES, I continue to monitor the implementation of SNSAs to see how we can make best use of the assessments on a local and national level.

**Juliette Mendelovits (Australian Council for Educational Research):** Until October 2018, I was the research director and general manager of ACER UK, which is a registered company in the United Kingdom and a wholly-owned subsidiary of ACER Ltd. In that capacity, I put together our bid for SNSA, as the system became, and led the implementation for two years until I returned to Melbourne where I am now based as research director for the ACER Group. I still have a strong and recent connection with SNSA.

**Professor Sue Ellis (University of Strathclyde):** I was involved in early meetings about assessment as a result of the Joseph Rowntree Foundation report on "Closing the attainment gap in Scottish schools". In writing that report, we noticed that there was no way to assess whether an initiative had closed or widened the gap, because there was no data. The report called for better data in schools; as a result, I attended three meetings on assessment with the Scottish Government.

**Professor Christine Merrell (Durham University):** Until 1 July 2018, I was director of research in Durham University's centre for

evaluation and monitoring, which provides standardised assessments as part of monitoring systems for schools. Many hundreds of schools in Scotland have used those assessments since about 1996.

**The Convener:** Thank you.

**Liz Smith (Mid Scotland and Fife) (Con):** Before we get into the detail of the evidence that the witnesses have submitted, I note that the Parliament is generally interested in the criteria that make for good-quality assessment. Professor Ellis flagged up the importance of good-quality data and said that it did not exist previously. Is the data getting better? What else will have to be done to ensure that parents, pupils and teachers understand exactly what makes for good-quality assessment? We are interested in the general parameters.

**Professor Merrell:** First and foremost, we have to consider the needs of stakeholders and establish the primary purpose of the assessments, before we get into the technical details of reliability, validity, content and so on; those need to be very clear from the outset. Different stakeholders have different needs: the learner may want to know about their current level of understanding and the next steps to aim towards; parents and carers need information; teachers are looking for various levels of information; and headteachers need management information, as do authorities at a national level. We need to be really clear in the first place about what we are conducting the assessment for, and move on to the quality and how we might best assess to get the information that we want.

**Professor Ellis:** We have to start with the idea that any assessment is a tool and that it takes time for professionals to learn how to use it; how to use it well; what it can do and cannot do; and what we can and cannot do with it. University of Strathclyde staff want any conversation to be rooted not just in ideological arguments about what stakeholders would like and need, but in an understanding of where Scottish education and educators are coming from in their current use of data and how it is seen.

We argue strongly that good standardised data is needed, as is a robust set of ethics around it to highlight to teachers, local authorities, inspectors, parents, the media and politicians what can and cannot be done with the data. A good ethics policy would educate teachers to use the data well and create a system that would work for the children of Scotland.

**Mhairi Shaw:** The criteria have to be based on what will make learning and teaching better; any information should help to inform teachers and young people and children and their parents of the

progress that they are making against national benchmarks or any other curricular element that would measure that progress. In essence, the information has to bring that about and eventually attainment will be raised.

Professor Ellis is right to say that there needs to be a sense of ethics around how we use the data and we all have responsibilities for that, including parents, schools, local authorities and the media. It has to be about improving the experiences of children and young people and making sure that they reach their potential; that has to be borne out in the criteria.

The data is getting better. The publication of curriculum for excellence teacher judgments is improving; the data is still experimental, but the SNSA will help to moderate the teacher judgments, as will other activities to make sure that there is professional dialogue around the judgments. As Professor Ellis said, the assessment will be one tool—not the be-all and end-all—that teachers will be able to use to measure progress and in their judgments about whether they are on the same page as their colleagues. Criteria around those aspects would be most useful for the whole system.

**Juliette Mendelovits:** I suppose that the focus of Liz Smith's question was on assuring the community of the quality of the assessments. In that respect, the quality of the assessment instruments is fundamental. We are at pains and we take a lot of trouble, using our expertise, to ensure that the instruments are sound, robust and valid from a number of perspectives; that includes consulting carefully and widely with people in the education community—the stakeholders—to make sure that what we are assessing is what is important.

We need to know that what we are measuring is what we intended to measure. We ensure that it is by getting qualitative feedback from stakeholders—learners, teachers, people in Education Scotland and the Scottish Government, for instance—but we also have statistical tools to ensure that the assessments are measuring something coherent and meaningful and that it is not just a random form-filling exercise. We have a lot of quality assurance measures in place and we have tried to make their implementation transparent to the public.

As a number of the written submissions state, including mine, if the results of an assessment are not understood and not used, it does not matter how good the assessment is—it will be pointless. The reporting is a key element of that, and we have worked hard with the Scottish Government to ensure that the reports are clear, transparent and accessible.

There are different levels of reporting. Fundamentally, the school-level reports are designed to give teachers information about individual pupil performance. There are school-level reports, which aggregate some of the data in a way that we hope is transparent and useful for schools. There are also local authority reports, which have a wider aggregative purpose but also give local authorities a lot of detail so that they can analyse the results in their own ways, with support.

The third key element in making assessments valid, useful and quality assured is to ensure that there is a good mechanism for providing professional learning to schools, teachers and local authorities so that they can interpret the results clearly, intelligently and effectively. The training programme that SCHOLAR at Heriot-Watt University is running, which has been implemented alongside the assessments from the beginning as part of our contract, is therefore a key element. It is unusual internationally that there was the foresight to bring forward a professional development and training programme from the inception of the national assessments to ensure that they were used in the way that was intended.

**Liz Smith:** Thank you. My colleagues will go into some of those specific aspects.

My second question has two parts. First, do you feel that there is a set of data that we do not currently have that would be helpful in informing the process of assessment? Secondly, I am interested in what Professor Ellis said about the ethical strand. Is it your view that the results are not being interpreted effectively and are not helping us to do what we are trying to do, namely to raise attainment? Will you expand on that?

**Professor Ellis:** I go in and out of local authorities and schools a lot and I talk to teachers, headteachers, local authority improvement officers and directors of education. Research is emerging that suggests that children's progression pathways are so variable that it is not appropriate to use a one-off standardised assessment for target setting, tracking or whole-scale interventions. There are examples in Scotland in which local authorities will test all the children in their area at a particular time and then automatically put the bottom 20 per cent into a fairly rigid and, for some, inappropriate set of work, and local authorities are sometimes using the information for streaming and setting. That is not just to do with standardised data that they get. Some local authorities do that with the formative data that they get from nurseries.

10:15

In the work that I do, I go into schools. Where schools have a two-form entry, for example, I will look at whatever data they have. That might be book-level data or standardised data. If I see a difference between the two classes, the heads might explain that they set them on entry to primary 1 on the basis of the formative data that they get from nurseries. If I explain to them how that enshrines disadvantage and that that is not an ethical use of data, they will often change their policy. I can think of two directors of education who have just sent emails to schools that say that they are not to do that.

At the moment, there is a very poor understanding of the research on how an assessment score can reliably predict results. I am looking at the research by Becky Allen when she was at Education Datalab. Only 9 per cent of children followed the projected pathway from their first standardised assessment to their fourth one. Ninety-one per cent either overshot or undershot. If there is that much variability in the system, it is unethical if somebody comes in at any point and says, "Right. We're going to group these kids on the basis of a single assessment score."

In learning to use standardised assessments well, we will need a really big professional mind shift in how staff think about and use assessments. That shift is probably not helped by high-level ideological debates. A shift probably needs to be made in how people respond to the data that they get. We see similar ethical difficulties when the media look at data from schools and try to pitch one school against another, because the sample sizes in primary schools are very often not big enough to enable such judgments to be made.

I would argue for a very grounded view of standardised assessments.

**Liz Smith:** Would you equate the term "unethical" with "misuse"? Is that what you are saying?

**Professor Ellis:** Yes. Some of the uses of standardised tests and non-standardised tests—local authority-devised tests—that I see happening in schools are not ethical. I see the introduction of national assessment as an opportunity to open that up for debate and to get a much better use of assessment that works for children and parents.

**Johann Lamont (Glasgow) (Lab):** May I ask a question?

**The Convener:** I want to let the other panel members respond before I bring you in, Ms Lamont. Does anyone else want to respond to that point?

**Liz Smith:** Is there any more data that we need that we do not have?

**Professor Merrell:** Yes, there is a lot. Are you thinking about what schools need or what could be done in the Government?

**Liz Smith:** I am thinking of what the schools need, because they deliver the assessments. Is there any data that you think is missing when it comes to our ability to produce good-quality assessments?

**Professor Merrell:** Assessments can be done at time points that are different from those for the national assessments. I have examples of schools that do that. They collect information from multiple sources to inform their practice. Assessments from the CEM are one example. I have the example from the current year of a primary 1 teacher who assesses her children with CEM assessments at the start of primary 1 because she wants some information about what they know and what they can do to inform her practice. Later in the year, she uses the standardised national assessments to confirm her judgments about where the children are. That is a nice blend of both assessments.

**Liz Smith:** That will give good results, which is the key thing. The nub of the issue is which assessment process is giving the best results.

**Professor Merrell:** Yes. That system is giving the best results, and it is not too onerous on the child or the teacher.

That was one nice example. As we go up through the primary school, assessments are used—maybe in alternative years—to give a bit more information, so that people are not waiting more than one year for information to come in.

I will add to Professor Ellis's comments about predictions. One study that we have done in England looked at children—about 45,000 of them—at the start of school and followed them up to the end of secondary school. The correlation between attainment at the start of school and age 16 is 0.5. As has been said, there is a lot of variation. Children do not necessarily follow a linear trajectory. There could be a burst in activity followed by a consolidation phase, so it is really important to look at attainment holistically. Between two time points, a child could be consolidating their learning, or maybe they might have just learned something new and they are going on from there. That is an interesting study to look at and to bear in mind. There is a relationship, but it is not fixed and set in stone.

**Professor Ellis:** That matters when schools respond to data in ways that are not appropriate.

**Juliette Mendelovits:** One piece of information that will be really useful to schools and the wider education community is the mapping of progress



over time. That work has been initiated, but it is not yet in force.

The methodology that is being used for the SNSA allows that to happen in a quite transparent way, because there is a long scale, which is being implemented in this year's assessment—it was not available in the first year. It will allow tracking, over time, of pupils as they go from primary 1 through to secondary 3 in each of the subject areas and equating as a year group. Schools will be able to look, for example, at how primary 4 results for this year compare with last year and next year and so on. The methodology that we are implementing in the SNSA allows that to happen. That area of data, which has already been instigated, will be improved.

Another important area that could be developed alongside or in the SNSA is qualitative explanatory information about, for example, how children engage with their learning, what their attitudes are to learning and what the school atmosphere is like. Currently, there is no instrument or survey mechanism in the SNSA, or alongside the SNSA, that captures such information. If ways of managing that could be integrated with the SNSA, that would be really helpful in trying to work out why things are happening in the way that they are.

**Johann Lamont:** On the question of whether something is ethical, if it is not ethical to get information about a child and decide how you then support a child, or presume how that child might be supported in terms of what work they would get, why would it be ethical to make judgments about an individual or a school against national benchmarks? Are you saying that the only way we should use the data is to support the individual child but that we cannot make any presumptions or assumptions about the child's learning from it, because it might lock them into particular forms of support? We heard from a colleague here that, in fact, judging a school or an individual against national benchmarks is seen as a way of pushing up attainment. Where does the balance lie?

**Professor Ellis:** Any short assessment can give you only a snapshot of where a child is at that one time. If you use that snapshot to make systemic changes to how that child is educated, such as putting them in the bottom set or into a catch-up learning group from which they find it difficult to escape, that is—

**Johann Lamont:** The concern is therefore not about the support but about the child not being able to move on from the support.

**Professor Ellis:** The school that I have seen making the best use of standardised assessment is in Woodlands primary school in Linwood. It makes hard use of assessment data to have hard conversations with staff but keeps the children

completely as part of the class and the learning community. The school recognises that learning is not just about the programme that the child is provided with but about the whole environment that the child is in.

Schools sometimes overplay their hand in the way they use some of the data—formative and standardised—at the moment. They are not doing it because they want to be bad; they are doing so because they have not realised that the data does not have predictivity.

**Johann Lamont:** Would that not be true of standardised assessments as well?

**Professor Ellis:** It is true of everything, yes.

**Johann Lamont:** The assessment is a snapshot and we should not presume from the data how a child should be supported. We should not assume that the assessment will be predictive. Can you explain why it is given such priority in education policy at the Scottish level if it neither predicts the child's abilities in the future nor determines the support that the child should have?

**Professor Ellis:** It gives useful information to schools and local authorities about how individual children are getting on. It could be useful for a class teacher.

There are two different ideas of what the assessment is good for. Curriculum for excellence is a complex curriculum with many layers and it is very responsive. The emphasis is on teachers getting the right learning mix for children. That is different from the five-to-14 curriculum, which was much more rigid. Kids progressed through that curriculum at different rates, but the curriculum was not greatly changed. Tasks might have been made a bit easier or harder, but the learning was not changed. Indeed, it was difficult in the five-to-14 curriculum to change the learning mix. Curriculum for excellence is premised on the idea that the learning mix matters. There is therefore a need for points at which to check that the learning mix is right, who the learning mix is serving well and who it is not.

**Johann Lamont:** The standardised assessment is therefore not a snapshot but shows whether an individual child is getting the right learning mix. Is that what it is for?

**Professor Ellis:** It can act as an opportunity to reflect on that. It can be diagnostic. The research on the predictivity of assessments is based on English data. It may be that, because the Scottish assessments are broader and better linked, they have a better predictive capacity. We would need to use them for 12 to 15 years to work that out. Until that point, the ethical position has to be to do no harm—so no sets, streams or catch-up programmes that remove children from the main

body of the class and put them in a different category from others on the basis of one snapshot.

**Johann Lamont:** Before I let my colleagues ask their questions, I have a specific question about the process around standardised assessment. In the briefing that we were given by Scottish Government officials on how the primary 1 test would be run, we were told that it could be done at any point that the child was in primary 1, so at any time between the ages of four and a half and six. If there is such a range, to what extent can that be a standardised assessment, given the gap in capability between a four-and-a-half-year-old and a six-year-old?

The second thing that we were told was that the test was multiple choice, with the option of answering A, B or C. A question might ask which of three words sounds the same as another word, and there is a button that the child can press to hear the answer options being said. I asked whether any distinction would be made in the assessment by the teacher or whoever between the children who needed to press the wee button to hear the words and the children who did not. I was told that no distinction would be made. Does the panel not think that, if a child is able to go through the process without needing the words to be spoken to them—because they would read and hear the words themselves—it should at least be reflected in the test?

10:30

There is the question of age range and then there is the question of functionality—how much more information are we getting than a teacher might get by working with the child in the class?

**Juliette Mendelovits:** The standardisation of the assessment resides in the fact that there is a single pool of questions from which an assessment is selected for each child who is taking the assessment in the year groups that have been identified. As the committee will know, it is an adaptive assessment, which means that, depending on how the child is performing in the assessments, they will get more difficult or easier questions according to the capacity that they have shown. Therefore, the assessment is pitched at the appropriate level for the child in order to get the maximum amount of information about what they know and what they do not yet know. The standardisation is in the pool of items being common to all children in the year group, the fact that there are some limitations around the administration of the assessments and that the results are processed in the same way for all children. Within that, there is some flexibility that is appropriate for an assessment that has low stakes—no individual child's future depends on

the results—and which takes into account the different equipment that the child might have at their disposal, because of the availability of hardware at their school, and also the child's way of approaching the assessment. There is flexibility in the way in which children approach an item, depending on their capacity.

Ms Lamont asked whether the children hear audio. There are some items for which the child would need to hear the audio in order to answer the question. There are other items for which, if they can already decode what is being asked, they do not need the audio support.

**Johann Lamont:** The point is whether the assessment that the teacher got would reflect that difference. I would have thought that that was a basic thing. The other question is how valid the group is if a child could be four and a half or six.

**Juliette Mendelovits:** Due to the way in which Scottish education works, children can start school at different ages.

**Johann Lamont:** Would it be better if the assessments were done by age rather than stage?

**Juliette Mendelovits:** Our approach is that children are in a particular year group, a curriculum is established for that year group, and we are assessing where children are in their stage of learning, as Professor Ellis said. Any assessment will only take a measure of a child's capacity at a particular stage. When the teacher receives a report on the child, it has the child's age—in case the teacher does not know, although they probably will—which is one of the factors that the teacher will take into account when interpreting the results of the assessment. As Ms Lamont pointed out, the child can take the assessment at any time in the school year, so it is not standardised in the sense that there is a particular day on which the child must take the assessment.

One of the key elements of the assessment is that the children are able to take it when the school deems that they are ready. It is designed to provide information to the teacher about where the child is in their learning, given that there are benchmarks for learning for the stages of schooling in Scotland. The other fact is that teachers know the child's age, how they are faring at school, what their attitude to school is and so on when they interpret the results, as well as other types of formative assessment.

**Johann Lamont:** So they are having to assess whether the child is ready to be assessed before they assess them?

**Juliette Mendelovits:** I am sorry, I did not catch that.

**Johann Lamont:** They have to assess whether the child is ready to be assessed. I understand

that that includes practice of this kind of test with the child, so that they know what to do.

**Juliette Mendelovits:** We are assuming that teachers will take into account when the child is ready to do the assessment. That does not mean when the child is going to be able to answer all the questions correctly; it means when they think that the child is emotionally, psychologically or intellectually ready to take the assessment.

**Professor Merrell:** Can I pick up on standardisation? Children learn through maturation—through their environment and so on—and they also learn at school. We have been looking at the impact of schooling on children at different ages. We have done that only through primary school, but we know that a huge amount of learning takes place in primary 1. Children tend to go to school knowing a few letters and by the end of primary 1 many of them can read a lot of words and do some comprehension and mathematics. As they go up through the year groups, the progress gets less. At secondary school, it has started to flatten out and at adulthood it is probably a level line and then declines. Both the age and the stage of the school year need to be taken into account.

It is problematic to have a standardisation that covers a large stage as well as the age, especially for the younger year groups. It is not so problematic for older ages—it is not a problem at the top end of primary and in secondary. We have quantified the amount of learning that takes place in a school year and that needs to be accounted for in any standardisation.

**Mhairi Shaw:** I have some points relating to ethics in the use of assessments. I emphasise that the SNSA is only one piece of assessment. Leadership is key in all this. There needs to be leadership at all levels—from the directorate and, especially, from headteachers—about the ethical use of data. The data has to be thought of as being about one point in time and reflecting how a child has performed in that assessment. It should be considered in the midst of all the other assessment information that a teacher will use daily about a child's performance against the curriculum and the activities that have been set for that child to make progress with the curriculum.

On the earlier point about the survey of pupils' attitudes to learning, one of the good things about the Scottish survey of literacy and numeracy was that there were both pupil and teacher questionnaires that measured confidence levels and gave good information to local and national authorities about how confidence levels could be improved for particular aspects of the curriculum. It was very valuable information, and we should look at building that into the SNSA. It was certainly

something that we used in East Renfrewshire when we reviewed areas of the curriculum.

**Professor Ellis:** Part of the difficulty is that, when we ask what teachers can do with the SNSA and whether it is to determine whether a child has achieved a level, we are looking at it as if it is a summative assessment. I found the Educational Institute of Scotland submission interesting in that it contained a debate about whether the SNSA was about confirming teachers' assessments or informing teachers' assessments. If it is confirming teachers' assessments, it almost uses the national assessment as a summative tool—has a child reached the level or not?

With curriculum for excellence, we need a shift in mindset. We need teachers to look at how they can use the assessments in a more diagnostic way. That diagnosis could be for particular items. There might be a whole class in which the comprehension levels are low and that would give the headteacher or the teacher herself the opportunity to say that they have not got the mix right.

There are also opportunities in the new assessments to look across items and take a diagnostic view. For example, in one item children have to listen to a story that is read to them and answer comprehension questions on it. There is another item in which they have to read a few sentences and answer comprehension questions. It can be easy as a class teacher to identify a child who is not comprehending when they read, as they cannot retell a story that they read two minutes before. If a class teacher has two children, both of whom do badly on the reading comprehension but only one of whom does badly on the listening comprehension, the teacher's point of intervention for each child will be different. In a class of 25 or 30, it can be very easy to miss poor oral story comprehension.

The assessments are a tool that teachers need to learn to use well, and they need time and space to learn to use them well. We could explore such uses. There is a lot of opportunity to provide good case studies of how the assessment items are being used well and ethically with lots of explanations about why that use is good. There is a danger that schools will look at the assessments and think that they are a predictive measure—that is a common strand that goes all the way up. Therefore, there needs to be a lot of education around that. However, teachers in Scotland want to do their best for the children whom they care for, so giving them opportunities to explore that is important.

If we look at the different views on assessment in all the submissions that the committee has received, we find that teachers are thinking about assessment in very different ways.

**Iain Gray (East Lothian) (Lab):** I have a couple of specific questions for Juliette Mendelovits on the design of the SNSAs, but I want to start with something that Christine Merrell said. She said that the most important thing in designing assessments was to have the primary purpose clear from the outset. That has been part of the debate on SNSAs. Are those tests designed primarily to provide information that teachers can use diagnostically in their learning strategies with pupils, or are they a way of measuring standards in schools and progress on addressing the attainment gap? In designing SNSAs, were you clear about what the primary purpose was? What was it? Those are questions for Juliette Mendelovits.

**Juliette Mendelovits:** There are dual purposes. I do not think that we would subscribe to the view that there is a single primary purpose.

**Iain Gray:** So there is not a primary purpose.

**Juliette Mendelovits:** That is not to say that there is no purpose. There are two really important purposes.

**Iain Gray:** But Christine Merrell's point was that the most important thing was to establish the primary purpose. You are saying that, in designing SNSAs, you did not know what the primary purpose was.

**Juliette Mendelovits:** No, I did not say that. I said that there was more than one very important purpose. A very important purpose is to give teachers good information about where children are in their stage of learning, which allows them to reflect on where those children are and to find out something new about them to help them to take the next steps. It allows them to reflect on whether children are showing challenges in their state of learning or are going great guns, so that something could be done to help them to extend.

**Iain Gray:** Do you mean children as individuals?

**Juliette Mendelovits:** Yes. There is also class-level information, so that people can look at where children are performing very well as a group, or not so well, so that action might be taken to support them. That is one very important purpose.

Another important purpose is to help the Scottish Government and the education community to improve the overall capacity of children in literacy and numeracy and to close the attainment gap. In order to have information about what the gap is and whether it is being widened or narrowed, one needs national-level data as well as data at the individual school level.

Both are very important primary purposes for the assessments. We are working towards meeting those goals through the way in which the assessment has been designed and reported on.

**Iain Gray:** In the first report on SNSAs, which was published back in December, ACER said that the national level results had to be treated with caution. I think that that was because the tests took place at different times in the year in different areas. The ACER report says:

“results from all learners should be interpreted with some caution when making any comparative judgements.”

Will you elaborate on that? You said that national monitoring is one of the purposes, but the report seems to imply that that will work only if all the children take the tests at the same time, which they do not.

10:45

**Juliette Mendelovits:** The comment that you quoted refers to interpretation of the results of smaller groups of individual children, class groups, school groups and local authority groups. When people interpret the results against the national norms, they should take into account when the assessments were administered.

The norming studies were conducted at two points in time—in November 2017 and March 2018. A scientifically drawn random sample of pupils across Scotland was stratified so that it took into account local authorities, gender, the Scottish index of multiple deprivation and age. We are confident that the measures of children's performance from those two norming studies are robust and reliable. What the report points out is that, when we are looking at the results of smaller groups of children—at the school level or whatever—in relation to the national norms, we need to take into account when the assessments were done.

We have achieved in the first year of implementation robust national standards across the country that are scientifically defensible.

**Iain Gray:** The norms are benchmarks, but if in future years you compare against them results that have come from tests that were taken at different times—

**Juliette Mendelovits:** As has been mentioned, there is flexibility in the design of the programme for assessments to be administered at a time when the school judges it to be appropriate. When the results are interpreted, people will need to take into account the point at which the assessments were administered if they are looking at the national norms as points of comparison.

Both things are true. We need to be cautious in making comparisons, but there is a set of statistics that allows us to look at what is happening nationally.

**Iain Gray:** I am really asking about year-on-year comparisons of the performance of the system.

Does the report not imply that results have to be treated cautiously? We might look at one year and then at another year and say that there has been an improvement or that the attainment gap has closed, but that would surely be affected by when the children took the tests.

**Juliette Mendelovits:** That is true, but we are recommending to the Scottish Government—I think it is enthusiastic about the idea—that national norming studies are conducted regularly, perhaps every couple of years, so that we can track how the nation is performing over time.

**Iain Gray:** Okay. Can I ask a slightly different question? The design—

**The Convener:** I think that Mhairi Shaw wants to come in on the previous question, Mr Gray.

**Iain Gray:** Surely—sorry.

**Mhairi Shaw:** I reiterate that the purpose of the SNSA is to confirm and verify or moderate teachers' assessments. Primacy in measuring performance and progress and whether the country is improving will lie with those teacher judgments and not necessarily with the SNSA.

**Iain Gray:** Yes. I get that.

ACER's submission says that, when the tests were being designed, the questions in the pool

"were reviewed and critiqued by panels of experts from Education Scotland and the Scottish Government"

and that, later, those panels were consulted again. What was the involvement of practising teachers in the design?

**Juliette Mendelovits:** There was a little involvement; I would not say that there was a great deal. We did some piloting in schools in February 2017 and we invited teachers to give feedback on the assessments as they saw them, so we took that into account in going further with the assessments. The nominated representatives from Education Scotland came from schools originally, so in that sense teachers were consulted, although not teachers who were working in classrooms at that point.

At the moment, we are implementing a questionnaire for teachers, which will be distributed widely in February, to ask for their responses to several dimensions of the SNSA, such as ease of administration, the usefulness of the reports, and the behaviour of the children and their attitude to the assessment. We are gathering systematic data from teachers during the current year.

**Iain Gray:** But practising teachers were not involved in the design of the tests.

**Juliette Mendelovits:** We were given a brief in the contract, but I do not know how much teacher

input there was at that end. During the development of the instrument, there was only a small amount of direct teacher consultation.

**Tavish Scott (Shetland Islands) (LD):** I would like to try to understand what you said to Mr Gray about norms. I will try to interpret what you said. A norm is a benchmark. Are you suggesting that the norming studies on the tests—I think that you suggested to Mr Gray that the norming studies are now going to be done every two years—will be used by the Government or policy makers to assess what was happening nationally?

**Juliette Mendelovits:** That is one way in which they can assess what is happening nationally, via the SNSA.

**Tavish Scott:** What is the other way?

**Juliette Mendelovits:** As Mhairi Shaw pointed out, ACEL data collection is the primary means of measuring whether children are attaining the standard.

**Tavish Scott:** Which one will be used, then?

**Juliette Mendelovits:** I do not think that it is a dichotomy. I think—

**Tavish Scott:** I am sorry, but I just do not understand what the Government is trying to achieve. I think that that is why we have all been asking questions about the purpose. Is it about teacher judgment, or is it about the national performance of schools?

**Juliette Mendelovits:** The SNSA is one contribution to the overall assessment picture and it is taken into account along with all the other kinds of assessment that teachers do daily in their classroom practice. I do not think that we would place the two things in opposition to each other in the way that we see the assessment profile developing.

**Tavish Scott:** No, but you said there was a proposal to the Government on producing information to allow it to make a national assessment of what is happening in education, which will have to happen every couple of years because—as Iain Gray said—schools are not doing the tests at the same time, so the data cannot be perfect and cannot be comparable year on year.

**Juliette Mendelovits:** Schools will get information from the SNSA annually, which they will take into account along with other assessments that they do. The SNSA can contribute to the information at the national level by conducting norming studies at regular intervals to track whether, for instance, the attainment gap is closing, which is one of the primary focuses of Scottish education.

**Tavish Scott:** I read ACER's view and, from a statistical point of view, my understanding is that, when you were first asked to do this work, you suggested to the Government that the tests should all be done at the same time so as to be capable of being compared. I mean during the school year, so that they are all done in May, for example.

**Juliette Mendelovits:** If you want to have a strict comparison of results from one year to the next—

**Tavish Scott:** Is that not what the Government asked you to do?

**Juliette Mendelovits:** The Government asked us to help it to develop an assessment that would allow teachers to understand where children were in their development of literacy and numeracy. One of the purposes of having a national assessment is that there is consistent data—

**Tavish Scott:** Indeed—consistent.

**Juliette Mendelovits:** There is no doubt about whether different instruments are being used. There are a number of different factors that need to be taken into account in a developing such a programme. The SNSA has a lot of wonderful features from ACER's point of view. We are a not-for-profit organisation and our mission is to improve learning. We are keen to promote programmes that honour teacher judgment and which respect the fact that teachers are in the best position to make decisions about an individual child's learning at school. I am talking about programmes that combine that with the ability to generate useful, larger-scale data sets that can be used to work out whether things are working well and where there may be a need to reflect on what is not going so well.

**Tavish Scott:** That is all entirely fair. I am trying to establish whether you think that that is best achieved if the tests are all taken at the same time during the school year.

**Juliette Mendelovits:** One of the aims might be best achieved in that way. If you want very strict comparisons of how a child is doing from one year to the next, taking a measure at the same time in every year—and this goes for large groups, too—is important. That is why we included a caveat to the national report about the caution that must be taken when making comparisons. That does not mean that no comparisons can be made; it means that people have to reflect on and appreciate, in a nuanced and intelligent way, the results that come out.

**Tavish Scott:** That is entirely fair. You sought to achieve that consistency of data; I guess that that was the whole purpose of the work that you have been trying to do for the Government in establishing this testing regime across Scotland. Is

it fair for me to assume, therefore, that your preferred approach to consistency of data would be to have the tests done at the same time?

**Juliette Mendelovits:** If that were the sole purpose of the programme, yes. Given that there are other purposes that are at least as important—namely, providing formative information to schools, teachers and individual learners of the kind that Sue Ellis has outlined—combining those desiderata is the way to move forward.

**Tavish Scott:** That is fine. Standardised tests have a number of other purposes.

**Juliette Mendelovits:** Yes. As I said in answer to Iain Gray's question, there is not one single purpose that supersedes all the others. We are looking for an assessment programme that combines the best features to serve a number of purposes.

**Tavish Scott:** That is helpful, thank you.

My next question is for Mhairi Shaw. As a director of education, you were clear in your answer to Liz Smith that there is one purpose. Do not let me misinterpret you, but you said clearly that that one purpose was to assist teacher judgment of the pupil's learning journey. Am I being fair in saying that?

**Mhairi Shaw:** That is the primary purpose, and it would bring about improvement in learning and teaching, which must generate data that can be used at many different levels. It can be used at individual teacher level and individual pupil level, but it can also be used at the whole-school, local authority and national levels, so that can we put in the right supports to improve learners' experiences. It is about the multiple use of the same data.

In East Renfrewshire, we are very experienced at using that data—I am not saying that Sue Ellis always agrees with what we do, but I think that our results stand for themselves.

**Tavish Scott:** Do your schools test in May?

**Mhairi Shaw:** Our schools have a six-week window for the SNSA. We continue to use our own standardised assessments to bridge the gap between those tailing off and the SNSA giving us more robust information.

**Iain Gray:** Is that a transitional measure?

**Mhairi Shaw:** Yes.

**Tavish Scott:** Now that those tests are in place, what information will they add that teachers did not have already, especially for P1?

**Mhairi Shaw:** They measure against a national benchmark so teachers can say whether children are performing well, or whether they are

performing as the teacher expects them to perform, given their performance in the classroom.

In P1, we do something similar to what Christine Merrell outlined. We use baseline information on entry to primary school, and the SNSA comes along later in the school year once the teachers think that the children are ready to take the assessments against the national benchmarks.

We would expect teachers to look at curricular advice and to think about how children are performing in that regard. They are not there yet, and there is still a bit of confusion. National assessments for 5 to 14-year-olds were taken when teachers deemed a child to have completed the work. The assessments do not have to be completed then, but they do have to be taken within that six-week window. We do them in May, along with the majority of the rest of the country, and that also helps with the norming exercises.

I point out that your papers indicate that the EIS was instrumental in taking the opportunity to make the tests more high stakes, if you like, if they were taken at the same time. I do not want—and I am sure that parents do not want—to end up with children having tutorials in the lead-up to the assessment in the way that that happens in England. Teacher judgment should be primary in all this, and in making sure that it all helps to moderate—

**Tavish Scott:** I agree, but do you not think that that is a danger, simply because of the pressure to close the attainment gap and all the things about the national picture that have been said by education secretaries and so on? Do you not think that it is inevitable that the pressure will be on schools at all levels, from P1 up, to make sure that assessments are all done at the same time so that national results can be produced and things can be said nationally about what is happening in Scottish education?

11:00

**Mhairi Shaw:** The curriculum for excellence teacher judgments are gathered at the same time. Therefore, the timing of the assessments is important so that they can inform those teacher judgments, which are gathered nationally. My answer is therefore no—I think that it goes back to Sue Ellis's point about the ethics and my point about leadership. Everyone has responsibilities in leading and making sure that the data is used appropriately when it is sufficiently robust to be used in that way.

I have spoken to class teachers who have used the raw data to look at how children have performed against particular skills or questions in the SNSA. That data has made them question pupil progress or have dialogue about that issue,

whether that is done to confirm the teacher's view or to consider whether children are making progress more quickly or more slowly than the teacher expected from class work.

The SNSA is just one piece of information, and we would never say that it is the only piece of information. It has to be in the mix with everything else. In that sense, the EIS was right. If we go down that road, it will become high stakes.

**Tavish Scott:** You are saying that we should stop obsessing about standardised tests, are you not?

**Mhairi Shaw:** My advice would be that the profession has welcomed them—

**Tavish Scott:** Some in the profession have welcomed them.

**Mhairi Shaw:** A lot of the profession—certainly those in East Renfrewshire—have welcomed them—

**Tavish Scott:** What about the EIS submission?

**Mhairi Shaw:** The EIS agenda is slightly different, and Mr Scott would need to ask the EIS about that.

**Tavish Scott:** Do not worry—we are asking the EIS.

**Mhairi Shaw:** Five-to-14 assessments were always rubbish as not being robust enough, but when they were taken away, everyone thought that they had been the best thing since sliced bread, because at least they gave some information. As indicated in the ADES submission, that is why a lot of people went on to use standardised assessments and the Durham approach. That gives people back the opportunity to measure their children's progress against national benchmarks and assessments that are taken on a national basis.

**Tavish Scott:** I will ask Professor Ellis a question, but I do not want her to go on about East Renfrewshire.

**The Convener:** Professor Ellis has wanted to come in on a couple of points, so she could answer the question—

**Tavish Scott:** Indeed, so can I ask a question before she answers the questions that I do not want her to answer?

My question is about your response to Johann Lamont, Professor Ellis. If I have got this wrong, do correct me, but I think that you said something along the lines that we would need 12 to 15 years of data before we could fully understand what was happening. I cannot exactly remember the context of the answer to Johann Lamont, but that strikes

me as one heck of a long time to find out what is happening.

**Professor Ellis:** It would perhaps not be that long, but data would be needed on children moving all the way through the school system before it could be seen whether a score that they received in primary 1 determined what university degree they got.

If taxpayers' money is being spent on assessment, the assessment has to be useful and make an impact. I have worked in Scotland for 30 years. As the committee knows, the results from the last few SSLNs have gone down. The only people I heard talking about that were politicians and the odd academic. I did not hear directors of education saying that they would look at their system because obviously something that they were doing was not working and that they would reassess their teaching; I did not hear class teachers or headteachers talking in that way.

If we want something that works and benefits the children of Scotland, we need something that has purchase with the practitioners who can make a difference. It has to speak to the teaching and learning that goes on in classrooms and how teachers think about the children sitting in front of them.

None of this will be perfect. It has to be good enough, and we have to interpret the results as being good enough, rather than as a truth—there are lots of different sorts of truths.

**Tavish Scott:** That is a Donald Rumsfeld answer if ever I heard one. Thank you.

**Ross Greer (West Scotland) (Green):** I have a couple of questions on some interesting points that have come up. Sue Ellis talked about Woodlands primary school in Linwood, which was an interesting example. She said that the school uses the results to have what she described as “hard conversations” with teachers, but that it is not jumping straight into ability sets or anything like that. Are those hard conversations about the needs of individual children, or are judgments being made about teachers on the basis of their class results?

**Professor Ellis:** Those conversations are not about judging teachers. The hard conversations with teachers are about the children who they are teaching and what those children need. The data set from any one primary classroom would not be enough to say whether a teacher is doing a good job. It is not a robust sample. The head in the school asks teachers how a child feels about their reading, the sorts of things they enjoy reading and what they are finding difficult. She asks the teacher what they have noticed about that and whether they can get more information on it. She will ask, “What can the child be introduced to?”

Who are their friends? What are their friends doing? Can we network them all?” It is a very inclusive approach.

We carried out a small study in Renfrewshire that looked at really hard-to-teach children—children whom the school system is not serving well at the moment. We adopted a small case study approach, but we found that the children who made the most progress were those whose health and wellbeing data—on how they felt about school and about themselves as learners—was integrated with their literacy data in professional conversations at a school system level. It is not rocket science: if a child is happy, they will learn better. However, when those conversations are actively integrated at a school level by the headteacher, you get children who are more relaxed and happy in school and who learn more effectively, too.

The hard conversations are specific conversations about how the child feels in their class, what opportunities they are getting and how those opportunities can be maximised.

**Ross Greer:** I want to ask a bit more about the class-level data. I completely take your point that the results of the SNSAs on their own would not constitute enough evidence on which to start judging a teacher's ability or performance. Should class-level results ever be used as part of a wider judgment of a teacher's performance? A concern that has been raised by a number of teachers is that the data will be used by management or the local authority in considering their performance. Should the data ever be used as a contributing factor?

**Professor Ellis:** No. I do not think that there is a robust research base for that. In fact, the British Education Research Association recently produced a publication on baseline assessments in which it made the point that it is not a robust approach. Academic papers published by Becky Allen from Education Datalab make the same point.

**Professor Merrell:** There are other means of finding out about a teacher's performance, and the assessment results will reflect other existing information. There will be classroom observations and a whole host of indicators in a school that a teacher is not bringing the best out of their children. Rather than the results being the motivation to investigate a teacher's performance, they should reflect what is already known about the teacher. If the results are used as the motivation, that brings in the risks of children being coached or whatever else might happen because teachers are fearful that that is the primary purpose of those data.



**Ross Greer:** I will move on to my substantial question, which is about the use of the data at local authority level. We broadly understand what the purpose would be in using it at the class level with an individual pupil and at the school level, but I ask Mhairi Shaw to elaborate a little on what local authorities are using the data for.

**Mhairi Shaw:** I do not think that we are using it for anything yet—we are certainly not using it in East Renfrewshire. I get results on how well children are doing. We have developed a tracking database that tracks all children individually, and we are speaking to ACER about that. The database has lots of information. It includes details not just on standardised assessments but on teacher judgments, for example, and we want the SNSA information to go into it, too.

We can use and cut that information in lots of different ways to have conversations. In essence, it is just about asking questions through the analysis of the data that it generates. For example, it might show that, at a school level, particular components, such as addition and subtraction, are not being taught particularly well. If we looked at that as a local authority and found it to be an issue across the authority, it would be incumbent on us to do something to bring about improvement, including by helping teachers to improve the learning experiences of youngsters. That is how the data is used formatively, and that summative information that we will get will allow us to do that.

**Ross Greer:** From what you have seen through ADES, is there a consistent approach across the 32 councils or are local authorities taking different approaches?

**Mhairi Shaw:** Local authorities will all be at different stages of development. In East Renfrewshire, we have used standardised assessments and the information that they provide for more than 20 years. We have not always got it right, but we are getting better now. We use the assessments to ask the questions. As with any analysis of data, all that it does is point the finger and, if you want to shine your torch on a particular area, it lets you ask how you would bring about improvement.

Certainly through regional improvement collaboratives, we would expect sharing of practice in relation to system leadership or system improvement. I cannot speak for all local authorities, but I can speak about what we are doing in the west partnership to bring about improvement through the analysis of data. However, I do not want to take us down that road.

The work that is going on through SCHOLAR will lead to real improvement and understanding of how the analysis of data can take place and what teachers should be extrapolating from their pupils'

results. That is to be welcomed, but we are at different stages and there is variability at all levels in the system.

**Ross Greer:** Professor Ellis, you mentioned that, as one would expect, you have recently spent a lot of time in a number of schools. What is your experience of the consistency between local authorities in their approaches to the data?

**Professor Ellis:** Local authorities are taking slightly different approaches, but we often do not know what those approaches are. Also, things happen at school level that directors of education do not always know about. There are different pressures on teachers and headteachers to do different things. One very popular literacy scheme recommends that, if children are not doing well in their literacy in primary 4, they should sit in with the primary 2 children for their literacy lessons. That daily walk of shame must do terrible things to how children feel about themselves as learners and be positively detrimental to their health and wellbeing. A director of education would not necessarily know that that is happening in their schools, unless someone like me notices it or Her Majesty's Inspectorate of Education calls it out.

When talking about making low-stakes assessments, we need to look carefully at the checks and balances in the system. We should ask HMIE and Education Scotland, when they inspect schools, to ask parents about things such as teaching to the test and repetitive testing. That monitoring has to be built in. We also need to consider how the inspectorate thinks about, uses and talks about data and look at the language that we use in that regard.

Parliamentarians could be really useful in that respect, because people often talk about the data being about the ability of the child, but it just tells us about the attainment on that day for that particular child; it does not have any capacity or ability implications. Therefore, the language that we use is important.

11:15

One thing that could be done is getting the unions and local authorities to have robust whistleblowing processes for teachers who feel that they are being pressurised to use data in inappropriate ways. There are a lot of practical things that we could do to move away from assessment debates being simply about ideological differences and start to look at the grounded picture of how assessments are used in Scotland and how we can get them to be used well. That is the practical problem that needs to be solved. If Scotland started to do that, it would probably be the only nation that I have heard of that has such checks and balances in place. That

is not impossible to do, but quite a hard and collaborative debate is required about that with all stakeholders involved. That is possible, and it would serve the children of Scotland well.

**Oliver Mundell (Dumfriesshire) (Con):** I want to go back to the point about whether the tests are robust enough to give a snapshot. I have heard from teachers in my constituency who are concerned about how the tests are formatted; whether they work for young people with additional support needs, such as those with dyslexia, dyspraxia or autism; whether the adaptive approach works or just means that people lose interest; whether people have the fine motor skills to manoeuvre the mouse; whether the difference in time limits for people to complete the test masks other things that are going on; and, to go back to Johann Lamont's point, whether different skills are tested at the same time in the same questions. Do you recognise any of those concerns?

**Juliette Mendelovits:** The question about accessibility for children with additional support needs has loomed very large in our development of the assessment. We have implemented a lot of affordances in the programme to help children who have a visual impairment or motor skill needs to allow them to do the assessment. That was clearly a very high priority in the Scottish Government's requests. We have had a lot of workshops and consultations with accessibility experts in and beyond Scotland on how to make the assessments accessible for children, and we have used the AA measure in the web content accessibility guidelines, which are a world standard of accessibility.

**Oliver Mundell:** Do you recognise that, in making the tests more accessible, you could end up masking other difficulties that a child faces? Does making adjustments to the test make it more difficult for the teacher, particularly in primary 1, to pick up nuances? Professor Ellis talked about differences in comprehension and other things. Does allowing more variables, differences in how the questions are answered and other things make it more difficult for the teacher to identify some things?

**Juliette Mendelovits:** When we introduce affordances in the assessments to allow children with additional support needs to take the assessments, we are always conscious of the key intent of the question, and we do not adjust it in such a way as to obliterate what it tries to measure.

**Oliver Mundell:** Having looked at the tests and having spoken to teachers, I cannot see how that can be the case. For example, if we go back to the point about decoding, there is a huge difference between listening to a question and reading it.

Those are two completely different things, are they not? Is it not possible for them to get muddled up?

**Juliette Mendelovits:** If the point of the question is to know whether the child can hear rhyming words, for instance, they will have to press the button to hear the word in order to answer the question, regardless of whether they have additional support needs. If we cannot measure that skill because the child does not have hearing capacity, they will not be able to answer that question, so some questions will not be available for all children. However, on the whole, as far as possible, the questions are available for all children.

What we have said in our guidance to teachers is that they should give children the classroom support that they would normally give them. As far as possible, we have made the assessments available to children who have additional support needs without teacher support but, if the child has an aid to assist them in normal classroom practice, it should be available to them. It is about striking a balance between making the assessments available to as many children as possible—to the vast majority of children—and preserving the integrity of what the assessment is trying to measure.

**Oliver Mundell:** Do you believe that you have got that balance right?

**Juliette Mendelovits:** I think that we have done a lot better than many other assessments do. It is not perfect, of course, but the fact that 95 per cent of the available assessments were taken while, I think, about 10 per cent of children are tagged as having additional support needs in the SEEMiS database suggests that many of the children with additional support needs have been able to take the assessments. When teachers receive the reports for such children and reflect on them, they will take into account the child's additional support needs. Again, it is a matter of interpretation and nuance.

**Oliver Mundell:** However, it is in theory possible that children with additional support needs will perform better, because of the adaptations in the test, compared with their actual ability. Is it possible that, in some cases, issues that children have are masked because of adaptations that have been made, and perhaps even adaptations for other children?

**Juliette Mendelovits:** When adjustments to items are made for children with additional support needs, the mantra that I use as a test developer—my background is in test development—is, “Would the affordance that is being added help a child who does not have additional support needs to do better in the assessment?” If it would, it is not a good affordance. We try to create a level playing

field so that a child with additional support needs can approach the item in a similar way to a child without additional support needs. Does that make sense to you?

**Oliver Mundell:** It does, but it does not seem to match up with what teachers are saying about the tests. Even for me as an adult, it is easier to identify words that rhyme by hearing them than by seeing them. If someone sees and hears them together, it will be easier to identify whether they rhyme than if they have just one option or the other. A bright young person might well take the opportunity to listen as well as read in order to maximise their chance of getting the question right. I know from speaking to teachers that even those who are positive about the assessments have questions about how they have been configured and road tested and how they compare with what is done elsewhere. However, I will leave that there.

My other question goes back to what we have heard about the rapid change in pupils' ability and how much knowledge they pick up in the early years of primary school. Does that mean that standardised assessment is more useful at some stages than at others? Is it better to let some things even out before we start to make judgments?

**Professor Merrell:** Standardised assessments should be able to give us useful information at all ages. As I said earlier, the warning about when we do standardisation is more important in the earlier years, when there is a period of rapid change. If we assess children within, say, a month, base our standardisation on that and then compare them with other children who were assessed at that time of the school year, we will get a more reliable result than if we have a standardisation that spans, say, six months, because in that case how would we control for the amount of learning that the children have done in that time, as well as their increasing maturity through age? That was a warning. It comes back to the point that a standardised assessment can give useful information throughout an individual's education career.

**Oliver Mundell:** My question was whether there are fewer risks in taking the baseline once some of the initial variables have settled down after that period of rapid change. Some children, because of home circumstances or other issues, start off with less knowledge. They might not be familiar with particular animals or might not have done a lot of reading at home, but within a year or two at school some of those things, particularly for more able children, settle down quickly.

**Professor Merrell:** Absolutely. That is a reflection of their learning.

**Oliver Mundell:** Therefore, is it useful to have a snapshot of individual pupils' knowledge before they have had the chance to start their formal learning?

**Professor Merrell:** That comes back to what we want to use the assessment for. If it is to be used to inform practice and the way that activities are tailored towards the level of development of that child, having a baseline at the start of the phase is really helpful. The teacher can then see the progress that the child has made during school time. If assessment is left too late, the amount of progress made is not captured.

**Juliette Mendelovits:** I agree with everything that Christine Merrell has said. The suggestion that it would be better to wait until later, when children's knowledge, understanding and skills will have evened out a bit, is not supported by the data. What we see in the data from the first year of implementation, which is consistent with what was found in the SSLN, is that the gaps between children's knowledge increase over time, rather than decline. Getting a good measure of where children are early on—

**Oliver Mundell:** Individual pupils shift around and there is huge variation in individual performance in that time. The issue is whether something of particular interest happens in that period of change for individuals. Or does it follow enough of a pattern to make that a useful measure?

**Juliette Mendelovits:** Overall, looking at aggregates, we see that the gaps in skills, understanding, capacity and attainment increase over time. For the individual child, there are many different trajectories of growth. Work that ACER has done indicates that there are up to six years of difference in attainment within any one year group. We would like to minimise that as far as possible, but we must recognise that children are at different stages and develop in different ways.

**Professor Merrell:** On the last point in Mr Mundell's line of questioning, we do not want to assess children on their first day in primary 1. That is not what I am saying. We want to give them a little time to acclimatise to the new school and classroom and settle down in that respect, but not to wait too long for assessment to happen.

**The Convener:** Mr Greer has what must be a very quick supplementary question on additional support needs.

**Ross Greer:** It should be very quick. I want to ask Juliette Mendelovits what I hope will be a yes/no question.

Are the tests designed to be a diagnostic tool for additional support needs?

**Juliette Mendelovits:** No, they are not.

**Dr Alasdair Allan (Na h-Eileanan an Iar) (SNP):** I do not want to put words in the panel members' mouths, but one of the messages that comes across loud and clear to me is that assessment is nothing new, and neither are some of the temptations or risks that others have identified that are to do with problems associated with assessment.

One thing that we have not talked about much yet is where the assessments fit in with curriculum for excellence. The panel has described how the curriculum is multi-layered and not statutory. Can you say anything about how the content of the assessments measures up against what we are trying to teach and measure in curriculum for excellence?

11:30

**Juliette Mendelovits:** The brief for the SNSA covers literacy and numeracy only; it does not cover the whole of curriculum for excellence, which has many other facets. Even within literacy and numeracy, there is no attempt to cover every aspect of curriculum for excellence. We must be perfectly frank and acknowledge that. For example, we cannot hope to assess engagement in reading through the kind of assessment that the SNSA is.

Given that, the benchmarks that were published in draft form in June 2016 and in finalised form in August 2017 are the basis for the development of the framework or blueprint for the assessment. In consultation with the Scottish Government and Education Scotland, we have shaped the assessment around key organisers in numeracy and in reading and writing, and every item in the assessment has been aligned with one of the benchmark statements. It is a Scottish assessment, and it is designed for Scotland. As members know, the original items came from an international pool, but they have been reviewed and, in some cases, modified or rejected because they did not align well with the benchmarks.

The assessment addresses aspects of curriculum for excellence literacy and numeracy benchmarks, but there is no attempt to say that it covers every aspect. It is, of course, only one ingredient in teachers' evaluation of how children are coping with the curriculum. It has a particular focus, but that focus is matched to curriculum for excellence.

**Dr Allan:** I am interested in hearing from ADES about that. I know that ADES made a submission on the benchmarks and how it thinks the assessments will measure up against them in the future. Is there anything that ADES wants to add about that?

**Mhairi Shaw:** We take as read what has just been said about the design of the assessments. As I understand them, they are designed to be reflective of the benchmarks and the experiences and outcomes in curriculum for excellence. Teachers can therefore use them to confirm—or not—their own judgments about children's progress with the curriculum.

I do not know that I have much to add to what Juliette Mendelovits has said.

**Dr Allan:** Is it your experience that, as the assessments have been developed and devised, the expertise and the views of teachers have been fed into the process? Are you content that that has taken place?

**Mhairi Shaw:** I am very content that that has taken place. There was quite a bit of evidence gathering, and other officers in East Renfrewshire were heavily involved in supporting ADES and the Scottish Government with the brief before the tendering document went out. ACER won that tender. I am content that we had input to all of that and continued dialogue on improving aspects. We have found ACER to be very open to that and, indeed, willing to work with us and listen. Juliette Mendelovits spoke about taking feedback from teachers about their and youngsters' experiences, and that is to be welcomed.

**Dr Allan:** Are the standardised assessments a better fit with Scotland's curriculum for excellence than the assessments that took place before? I see Professor Ellis nodding. I wonder whether she has a view on that.

**Professor Ellis:** I think that they are. They measure a broader range of skills. If we get the right ethical debates on them, we can help teachers, politicians, the media and parents to understand that assessment scores are not necessarily about some children being more able than others but are simply about the experience that they bring. Curriculum for excellence is very much about working to the needs of children in a rich and inclusive way, and the assessments are better than most that I saw happening—both local authority internally devised assessments and published ones.

**Mhairi Shaw:** In East Renfrewshire, we redesigned our internal standardised assessments a number of years ago to fit with experience of the outcomes as they were published. We also ask our teachers to make judgments about children's progress so that that judgment can also be benchmarked against the outcomes from the standardised assessments.

We took those steps, and at this point I am not sure whether the SNSAs are giving us any more information other than the ability to look at how children are doing against a national benchmark. I

cannot speak about what used to happen in around 24 local authorities that used the Durham assessments.

**Professor Ellis:** One of the issues is timing. Teachers want instant feedback, but they need a bit of time to get their heads around the assessment. I cannot remember which submission it was, but one of them said that the assessment gives a lot of information at a very granular level and that teachers do not have time to look at that level. Part of me wonders what people would think if their doctor said that they did not have time to look at the granular level of blood tests. Teachers might not want to look at that deep granular level for every child but, if my child is not being well served by the curriculum, I want the teacher to have data that she can go into—data that she can look at, interrogate and think about in lots of different ways.

One of the things that any assessment does is get teachers to look at lots of different kinds of data about what progress might mean for individual children. Making progress as a reader is not linear; it involves working to a broad horizon, and lots of different pathways can be taken. Teachers need the time to look at the data, think about it and learn how to use it well in the context of curriculum for excellence. The assessment has the potential to work if there is the professional and political will to let teachers do that.

**Dr Allan:** There have been a couple of coded—perhaps I mean polite—references to the way in which politicians talk about the assessments. Are there any lessons for the body politic in Scotland to learn about how we should talk about the assessments and promote public understanding of what they are and are not?

**Professor Ellis:** I would welcome a cross-party collaborative and professional consideration of how local authorities, schools, teachers, parent groups and the media can work together to design a system that works well for children. There are issues about making sure that it is not used to classify and grade teachers or schools, because such grades have negative effects on children. Ultimately, it has to be about teaching and learning, as Mhairi Shaw said, and about empowering teaching and learning.

I would like politicians to talk about children being more or less experienced rather than more or less able. I would also like a bit more focus on what is happening in the system at the moment that is not perfect or desirable. Discussions should be grounded in how we can make things better, not in saying, “This is right; this is wrong,” or, “This is good; this is bad.”

**Tavish Scott:** I would like Liverpool to win the league, but we do not get everything that we wish for.

I have a question on the topic that Alasdair Allan has been asking about. The EIS submission tells us what Scottish Government officials said when they introduced testing in Scottish education. According to the EIS evidence:

“The assessments were said to cover at a maximum around one tenth of the skills and knowledge expected at each CfE level in ... Literacy and Numeracy.”

My question is for a director of education. Do you recognise that as the reality?

**Mhairi Shaw:** Ha!

**Tavish Scott:** How many tenths does it cover?

**Mhairi Shaw:** I would not be able to answer that question. I do not work with the curriculum—I lead the curriculum.

**Tavish Scott:** Okay. It was an unfair question about the one tenth. The point is, how important are the assessments? That is the point that Alasdair Allan has been driving at. What the EIS says rather supports your contention that we are all getting too obsessed with the assessments. It says that only one tenth of the skills and knowledge that a pupil gains at each CFE level is helped by the assessments.

**Mhairi Shaw:** The EIS is making the point that we should not blow the issue out of all proportion. My advice is that we should ensure that the assessments are allowed to be used for their primary purpose, which is to monitor whether the system is working well. Yes, they will give you, as politicians, information about whether attainment is growing and whether the gap is being closed, but it is more important that they inform the professional judgments of teachers. We need to ensure that teachers have confidence that we will allow them to make those judgments and expect them to use that information in a professional way.

If that 10 per cent covers the most important skills, and if those skills have been put into the design because they are the key skills that a teacher would want to see children making progress on, I would say that that is enough. However, as I have said several times this morning, we all have a responsibility to allow the assessments to ensure that the public at large have more confidence in the system and that teachers are getting it right. That takes us back to the primary purpose of assessments, which is to inform teachers' judgments.

**Professor Merrell:** That is quite a helpful point for the EIS to make. Assessments will never assess every aspect of learning against one area of the curriculum, so that is a really healthy way to

look at it. If we say that openly, it might prevent a narrowing down of the curriculum solely to those aspects. Those aspects are important and we can use them to monitor progress, but we should not think that they are the be-all and end-all and should try to prevent a narrowing down that would result in children learning only those things.

**Juliette Mendelovits:** I want to go back to Alasdair Allan's question about what politicians can do. I would hope that Scottish parliamentarians would grow to feel proud of the assessment. It has many features that will be admired internationally. Scotland should be shouting about some of those excellent features, such as the way in which teacher professional judgment is valued in combining the results of the assessment with the teachers' own judgments of children's progress.

The fact that the assessment is online and adaptive is excellent—I do not know of any other national assessment that has those features. Attempts have been made in Australia, but they have not been as technically successful as the introduction of the assessment in Scotland. As my colleague pointed out, there are some caveats and questions about the accessibility features, but the fact that the assessment is designed to be as inclusive as it is and has tried to take into account accessibility is an important feature of which Scotland should be proud.

I would like the committee and the other members of the Scottish Parliament to take pride in what has been achieved so far. It is not that there is no room for improvement—there is—but we should acknowledge the achievement so far.

11:45

**Jenny Gilruth (Mid Fife and Glenrothes) (SNP):** I will pick up on Alasdair Allan's point about benchmarking the assessments against curriculum for excellence. Perhaps Professor Merrell will be able to help me with my first question, which is an historical one. Were the CEM assessments that were used previously benchmarked against the five-to-14 curriculum?

**Professor Merrell:** We made a prediction of the five-to-14 level on the basis of the CEM assessments, but it was a percentage prediction rather than a direct link to a CFE level. It was important to do that; we would not give a one-to-one mapping of how one assessment predicted that you would do on another one. Many years ago, we worked with Fife Council staff and with teachers on the content to ensure that we were aligned to the Scottish curriculum.

**Jenny Gilruth:** Were the CEM assessments done every year and at every stage?

**Professor Merrell:** They were available for use every year, but different authorities and different schools chose to use them with different year groups.

**Jenny Gilruth:** I ask the question because I am a Fife MSP, and you might be aware that Fife Council recently voted to scrap SNSAs and revert to CEM tests for every year group and at every stage. Therefore, Fife schools are assessing more than they would have been had they chosen to use SNSAs. That is a local point.

I want to pick up on Sue Ellis's interesting point about educating teachers to use data well. Historically in Scottish education, data has been used by the management in schools—principal teachers, deputy heads and headteachers. You also mentioned the SSLN. When I was teaching, kids would be taken out of my class. I had no idea where they were going, and then they would suddenly appear back in the classroom. To me, as a practitioner, that data was not useful in informing my understanding. I note that the Organisation for Economic Co-operation and Development's 2011 review says:

“without adequate training, teachers may not have the assessment literacy and ability to appropriately interpret results and to identify areas where curricular strategies may require adjustment”.

What kind of teacher training will be required to support their understanding of the data?

**Professor Ellis:** My experience, from working with teachers, is that the most useful education is when teachers work with real data from their children. There is work to be done so that teachers learn about assessments not only in an abstract sense, but by navigating through what is in front of them. They need to know when and how they can take a deep dive and look at the granular information that is provided. One of the advantages of the SNSA is that teachers get the results instantly, and they can click through to see the different ways in which children responded.

**Jenny Gilruth:** I would like to bring in Mhairi Shaw. Should there be a consistent national approach? Does ADES have a view on how the training that is given to teachers is monitored at local authority level, to ensure that there is parity of access to training and that teachers have a good understanding of what the data means? There is a bit of a gap between the data with which teachers will be provided and how that will help to inform their practice.

**Mhairi Shaw:** SCHOLAR is already providing training for headteachers and deputies. They can have a dialogue with a link worker from SCHOLAR and say, “This is where we're at.” I take your point about such training not always being available to class teachers, but the expectation is that it will be

cascaded down, particularly to the P1, P4 and P7 staff who will be using the data—although those staff will not always work with those age groups. Historically, secondary schools probably have more experience than primary schools of using attainment data.

In East Renfrewshire, we are very pleased, and I am sure that my colleagues in Asda—I mean ADES; that was a Freudian slip—are also appreciative of the information. We are able to bespoke it; our conversations involve people saying, “Our staff are already well versed in using blah, blah and blah. Let’s see whether we can use the data at a more granular level.” That is already in place, and it is up to local authorities to make use of it.

**Professor Ellis:** It might be worth going back to SCHOLAR and asking it to do some standalone things that teachers can download off school premises and outside school time to get the information that they need. Checklists for local authorities, headteachers and teachers about what they know and what they do not know could also identify where the gaps for assessment might be. The roll-out, in terms of both initial teacher education and continuing professional development, has not been as proactive as it could have been. However, it hit schools at a really busy time, with funding and a whole load of other things going on. Take two: now is our opportunity to improve it; that is a growth mindset.

**Juliette Mendelovits:** Another item on my list of things that Scotland should be proud of is, as I mentioned earlier, that a training programme was initiated at the beginning of the assessment programme. That was a really innovative move on the part of the Scottish Government.

SCHOLAR is developing the professional learning programme as the SNSA matures. In the first year, a lot of it was about the more technical dimensions, such as how teachers accessed the assessment, designed log-ins for the children and downloaded the reports. Increasingly, the emphasis will be on the interpretation of reports and what teachers do with the information that they have from those reports.

Those programmes are being developed and are available not just in the face-to-face meetings—which are extremely important and probably more fun than sitting and looking at a webinar—but also through webinars. There are also PowerPoint presentations and text guidance on the platform, to help teachers to become familiar with the assessment and how it might be used.

**Jenny Gilruth:** Back at the start of the meeting, Juliette Mendelovits mentioned the questionnaire that will go out to teachers. Will that questionnaire

consider their experiences of implementing the assessments?

An issue that has come up in conversation with a lot of my friends, who are still teachers, is the provision of information and communications technology in schools and the lack of opportunity to access appropriate ICT to deliver an assessment. That is a critique not of the assessments but of the provision of ICT. Will the questionnaire consider that?

**Juliette Mendelovits:** Yes. There is a section about the ease of implementation in the classroom that includes questions such as whether the school used the diagnostic assessment to ensure that it had the appropriate level of equipment before the kids took the assessment and how easy it was to get the children to log on. The questionnaire includes those types of questions, as well as questions about the quality of the reports and so on.

**Jenny Gilruth:** Is ADES looking at that in terms of equality of provision across the country, Mhairi Shaw?

**Mhairi Shaw:** ADES would not necessarily look at the provision of ICT in each individual local authority.

However, it is an important point that the assessments become more high stakes if children have to be taken to ICT suites to be able to undertake them, such as in schools where wi-fi is not available. We need to be mindful that, to make them as low stakes as we possibly can, the assessments would be best done on a tablet or in the classroom where possible. That is the advice that we would give to schools.

**The Convener:** Before I move on to my final colleague, I will ask a little question. I have a better understanding of the standardised tests that were used prior to these. We talked about the CEM tests, and when my son went through the five-to-14 curriculum I remember him talking about cognitive abilities tests in schools. To what extent are those being used by schools?

I understood Mhairi Shaw to say that East Renfrewshire has developed its own model. Has any other local authority developed its own model or is everyone else using a commercial model? Will the introduction of the new tests remove the requirement for CEM or cognitive abilities tests to take place?

**Mhairi Shaw:** My understanding was that 24 local authorities used the Durham assessments. I am not sure, but I think that 31 local authorities in total used some form of standardised assessment, although obviously they did not all use the same ones.

Certainly, the publicity and the advice around the introduction of SNSAs suggested that they would save local authorities money, as they would not need to continue with the other tests. I cannot say whether other local authorities have stopped using the assessments that they used before the introduction of SNSAs, but I would think that the intention is that people would not overassess children. We are involved in helping to shape SNSAs, and will continue to do so until we get them into a form that will be able to replace the assessments that we have.

**Rona Mackay (Strathkelvin and Bearsden) (SNP):** I am conscious of time, convener, so I will keep my question brief.

What can be done to maximise the potential of the tests? I am interested in something that Professor Ellis said earlier about the health and wellbeing of the child. Clearly, these tests do not provide that kind of data. Should that be done? Could it be done easily, perhaps by the addition of a few extra questions in the test?

**Professor Ellis:** Schools collect health and wellbeing data. A lot of them will use the SHANARRI wheels—safe, healthy, achieving, nurtured, active, respected, responsible and included—and they will ask children about their friendships, how they feel about the curriculum, how they feel about different aspects of learning in the curriculum, how they feel about coming to school and so on. That data exists.

**Rona Mackay:** Where does that data go? Who sees it?

**Professor Ellis:** That will be kept at a school level. However, even in local authorities in which all the schools are collecting that data, we have found that, in the schools in which there is a discussion of that data along with the other data as part of the progression meetings between teachers and headteachers—during which they discuss the planning for the class, what the school needs and what individual children need—the children seem to be happier and make better progress than the children in schools in which that data is quite separate and is discussed in separate meetings.

**Rona Mackay:** Should that data be collected as part of the assessment?

**Professor Ellis:** No. You need to keep it simple. Good enough is good enough. There are points where you just have to say to teachers, “This is really complicated—you are the professional; you pull it all together.” However, we need to be learning from schools that seem to do that really well and we should promote that sort of data collection and use as a good way forward for others.

**The Convener:** I am afraid to say it, but Johann Lamont has a final supplementary question.

**Johann Lamont:** You are quite right to be afraid to say that, convener. Thank you for allowing me to come back in.

There is a suggestion that this issue is a sort of political and ideological battle. However, would you accept that the debate is really about balancing the benefits of the tests against the consequences for and the costs to local authorities and individual schools of running them?

The evidence that the committee has gathered from teachers and the parents of people with additional learning needs or learning support needs suggests that those needs are not being met currently. A number of reports suggest that schools are under huge pressure and that there are fewer and fewer members of the support staff to assist teachers in doing their job. Does there come a point at which the consequence of doing these tests is not so much that there are more resources, as I think was suggested by Mhairi Shaw from ADES—with needs being identified and resource being brought in—but is actually that resources are being taken away from that side in order to deliver the tests? Do you accept that many people at a school level are saying that the consequence of running the tests is that support staff are being removed from schools? If you accept that that is the case, are there other policy choices that can be made? In other words, would it be reasonable for a headteacher to say that the consequence of running the tests, which might theoretically be good, is that there is less support for young people in the classroom and that, given that that is the case, they will exercise leadership and say that the tests are not a priority?

**Mhairi Shaw:** I think that that goes back to the question of ICT and whether there is an opportunity for the tests to be administered in a classroom setting without taking children to another setting, and who would do that. I have not heard comments about the tests being linked to issues with pupil support assistants who are allocated to schools for the purpose of ASN. There will be other pupil support assistants who might come under the category of classroom assistants as opposed to those who are there for particular children.

12:00

**Johann Lamont:** But, with respect, they are all now categorised as classroom assistants, as we have learned from the statisticians.

**Mhairi Shaw:** I am not sure that that is accurate, actually.



**Johann Lamont:** I have been told, anecdotally, by people who work in schools—in particular, primary teachers and additional support staff, who are under phenomenal pressure—that the tests are bringing added pressure and are taking them away from their core job of supporting young people in the classroom. If we could evidence that in a way that would satisfy you, would your view be that different policy choices should be made and that the first priority would not be managing these tests but ensuring that schools are properly resourced to support the learning of young people, particularly those with additional support needs, based on the reports that have come to this committee?

**Mhairi Shaw:** I have to say that I think that that is an unfair question. Gaining information from such assessments and supporting particular and individual children should not be an either/or question.

**Johann Lamont:** It should not be, but, if somebody tells you that people are making that choice in primary schools, would that cause you to reflect upon the priority that has been given to the policy of those assessments?

**Mhairi Shaw:** Being a solution-focused person, I would find a solution.

**Johann Lamont:** Would that involve further resource?

**Mhairi Shaw:** I would find a solution.

**Johann Lamont:** Which may include further resource.

**Mhairi Shaw:** I would find a solution.

**The Convener:** Does anyone else have any final thoughts? If not, I thank everyone for their attendance this morning. It has been quite a long session and we really appreciate everyone coming along.

12:02

*Meeting continued in private until 12:29.*



This is the final edition of the *Official Report* of this meeting. It is part of the Scottish Parliament *Official Report* archive and has been sent for legal deposit.

---

Published in Edinburgh by the Scottish Parliamentary Corporate Body, the Scottish Parliament, Edinburgh, EH99 1SP

---

All documents are available on  
the Scottish Parliament website at:

[www.parliament.scot](http://www.parliament.scot)

Information on non-endorsed print suppliers  
is available here:

[www.parliament.scot/documents](http://www.parliament.scot/documents)

For information on the Scottish Parliament contact  
Public Information on:

Telephone: 0131 348 5000

Textphone: 0800 092 7100

Email: [sp.info@parliament.scot](mailto:sp.info@parliament.scot)

---



The Scottish Parliament  
Pàrlamaid na h-Alba