FOI Request 18/00100

Document 1 – SRO SNSA Project Meeting – Standardisation – 18 May 2017

NATIONAL IMPROVEMENT FRAMEWORK –SCOTTISH NATIONAL STANDARDISED ASSESSMENTS PROJECT – STANDARDISATION FOR THE SNSA
SRO MEETING – 18-05-17

Purpose
- To provide an overview of the Standardisation process required for the SNSA. The overview will highlight key options and make recommendations on the way forward.
- To seek approval on the recommended options.

Proposals for consideration

SRO's are asked to agree which of the Standardisation options we should proceed with in session 2017/18

The Standardisation options all involve controlled sampling of pupils at set times within the academic year which could be perceived to be at odds the policy position of teachers deciding when a child or young person does the assessment. Therefore we need assess how to secure standardisation which is at the core of the SNSA, while still maintaining the integrity of the policy position.

Background

ACER have provided, through a workshop and subsequent paper (attached), a comprehensive overview of the various standardisation processes required in the first year of implementation - 2017/18. They have proposed a number of options and are seeking our response so that they can begin their planning processes. These options are grouped under three distinct but related standardisation processes that they identify as:

a) Norming - Norm referencing is an alignment of outcomes against distributions of achievement for population reference groups.

b) Calibration - Calibration of items is the mapping of test questions (items) onto a scale that indicates the difficulty of the items in relation to the capacity of learners. This scale is used to report performance of pupils in both quantitative and qualitative terms.

c) Equating - equating studies in the year groups in between the assessed years with students completing assessments with items from the upper and lower year levels.

These processes are necessary if in the second year of implementation ACER are able to provide a national standard against which all children and young people can be placed in a single long scale. This is a core part of the ACER package and provides the key data that is currently not available in any standardised assessments currently available in schools.

Key issues / decisions arising from the Standardisation process:

1. Agreeing which of the options to approve

The ACER paper outlines the various options with some analysis of the strengths and weaknesses of each option.

In considering the options carefully within the SNSA project team, our guiding principle has been that despite the complex nature of these options, we should be aspiring to the very best standardisation data possible. This will ensure that we get the best information to support our children and young people and teachers.

We recommend the following options:

a) Norming – we recommend Option 1a.1 or 2 (pg6/7 of the ACER paper)

This would involve the maximum number of children being sampled (8800) to give us the most reliable standardisation

Summarising the burden across schools, teachers and pupils:

- Around 4% of pupils to spend up to one hour completing one test
- Around 1 pupil per teacher
- On average, 3-4 pupils per primary school and 5-6 per secondary school
- At most approximately 15 pupils from any one school
- A small number of schools have no pupils sampled

Key features of this design:

- Similar to SSLN
- Schools and community familiar with the approach
- As with SSLN, school comparisons not possible
- Most or all schools involved, but burden is minimal
- Minimal burden on teachers and students
- Lower precision than achieved on SSLN outcomes, but equivalent to international standards for educational assessments such as TIMSS and PISA (so very defensible)
- Stratification by agreed variables – e.g. Local Authority/Deprivation/ Sex /Age – to ensure that sample distribution matches with population distribution on those characteristics

We also recommend that two norming studies are carried out in year 2017/18. These could be November and April (para 340 pg 11). These would provide us with time referenced standardisation at the end of the first year, a significant achievement.

b) Calibration – we recommend Option 4 (pg14 of ACER report)

This is a combined approach with collection of data throughout the school year.

This option would use three strategies:

- analysis of data from the norming study (or studies) (similar to Calibration Option 1)

- analysis of data from "slices of time" (eg three or four defined months) using all data collected in those slices. (similar to Calibration Option 2)
- analysis of (most of) the full year's data (similar to Calibration Option 3)

c) Equating – we recommend Option 2 (pg 18/19 of ACER report)

This would include ASN children in the equating studies meeting our inclusive approach.

2. Organising the sampling of children and young people

The timeline provided by ACER details the complex nature of this standardisation process. We will need to be very careful how we organise this with schools and local authorities. This is a considerable logistical exercise and resource will have to be allocated from within the SG team.

3. Providing a clear message on why this is necessary but does not contravene the policy position on teacher choice.

Standardisation is a necessary process for the implementation of the SNSA. However it involves controlled sampling of a small but significant number of children and young people (approximately 4%).

We will have to be clear on why it is necessary, how we are going to organise it and why it does not contravene our policy of teacher choice on the timing of the assessments.

To achieve this we will develop a communication strategy on the Standardisation process.

Recommendations

SROs are asked to:

a) Agree to the recommended options for Standardisation, which will then be communicated to ACER

b) Note the SG resource implication to oversee the sampling process

c) Agree to a communication strategy being developed to support the standardisation process

<mark>Document 2 – Options Paper from ACER – Standardisation – 18 May 2017</mark>

exemption applied S.33(1)(b) - withheld

NIF Programme - PMO - Meetings - SRC

3

NATIONAL IMPROVEMENT FRAMEWORK – SCOTTISH NATIONAL STANDARDISED ASSESSMENTS PROJECT – PROTOCOL FOR NORMING STANDARDISATION FOR THE SNSA SRO MEETING – 07-09-17

Purpose

- To provide an overview of the protocols to govern approaches to the norming standardisation process required for the SNSA. The overview will highlight the key options and make recommendations on the way forward and will inform the Communications Strategy signalled in PMO 18.05.17 on Standardisation for the SNSA.
- To seek approval on the recommended option.

Proposals for consideration

SROs are asked to agree the protocol governing approaches to the standardisation process which should be taken forward in session 2017/18

The standardisation options chosen on 18.05.17 each involve controlled sampling of pupils at set times within the academic year which could be perceived to be at odds with the policy position of teachers deciding when a child or young person undertakes the assessment. Therefore, we need to assess how to secure standardisation which is at the core of the SNSA, while still maintaining the integrity of the policy position.

Background

ACER provided a comprehensive overview of the various standardisation processes required in the first year of implementation - 2017/18, grouped under three distinct but related standardisation processes identified as:

d) Norming - Norm referencing is an alignment of outcomes against distributions of achievement for population reference groups.

e) Calibration - Calibration of items is the mapping of test questions (items) onto a scale that indicates the difficulty of the items in relation to the capacity of learners. This scale is used to report performance of pupils in both quantitative and qualitative terms.

f) Equating - equating studies in the year groups in between the assessed years with students completing assessments with items from the upper and lower year levels.

Following SRO decisions at 18.05.17, the following approach to norming was agreed:

b) Norming

Summarising the burden across schools, teachers and pupils:

- Around 4% of pupils to spend up to one hour completing one test
- Around 1 pupil per teacher
- On average, 3-4 pupils per primary school and 5-6 per secondary school
- At most, approximately 15 pupils from any one school

- A small number of schools will have no pupils sampled

Key features of this design:
- Similar to SSLN
- Schools and community familiar with the approach
- As with SSLN, school comparisons not possible
- Most or all schools involved, but burden is minimal
- Minimal burden on teachers and students
- Lower precision than achieved on SSLN outcomes, but equivalent to international standards for educational assessments such as TIMSS and PISA
- Stratification by agreed variables – e.g. Local Authority/Deprivation/ Sex /Age – to ensure that sample distribution matches with population distribution on those characteristics

In addition, it was agreed two norming studies will be carried out in year 2017/18. These would be in November and April to provide us with time-referenced standardisation at the end of the first year.

The process of time-referenced standardisation in session 2017-18 presents a difficulty where teachers and/or local authorities assess that individuals, in the normal course of learning and teaching, be presented at the SNSA within the same timeframe as the norming standardisation process i.e. in November or April in session 2017/18.

For those individuals not selected under the agreed stratification of participants' in the norming standardisation process, this presents no issue.

For those individuals chosen to be particpants in the norming standardisation process, teachers might have to delay or advance presentation of the SNSA for normal assessment purposes to avoid the 'practice effect' where individuals can be advantaged by having had exposure to the assessment. Normally, a gap of one or two months between assessments would be needed to minimise this effect. This approach could be characterised as a 'separation protocol' where standardisation and assessment processes are separated to provide distinct data sets for each. Unlike the normal assessment process, individuals participating in the standardisation process would not have reports generated for teachers.

This approach would involve individuals being presented with the SNSA twice in a session, could have cost implications and would interfere with policy allowing teachers to decide the timing of standardised assessments.

However, an alternative method, characterised as the 'accommodation protocol', could allow teachers to retain control of the timing of the assessments and allow the norming process to proceed. This is achieved by participants sitting an assessment coincidentally with the norming standardisation process, teachers receiving reports in the normal manner and ACER using the same information for standardisation purposes.

Protocols for norming standardisation for the SNSA

Scenario one: Separation protocol

Where there is no coincidence between teacher-led assessment timing and the norming standardisation process, an individual selected for the norming process would be presented with the SNSA twice, with no reports being generated for teachers for the norming standardisation assessment, but reports being generated in the normal way for the teacher-

led assessment held at a different time.   This would require the child to do two assessments but with a significant time period between them. This would require careful communications.

In this way, we separate the standardisation process from the professional judgments of teachers choosing to present individuals at a different time of year and provide the means to gather information for standardisation purposes. This method will have cost implications.

Scenario Two: Accommodation protocol

Where there is coincidence between teacher-led assessment timing and the norming standardisation process, an individual would be presented with the SNSA once only, with reports being generated for teachers in the normal way whilst information critical for standardisation purposes would be used by ACER at the same time.

In this way, we accommodate the professional judgments of teachers regarding timing of the assessments, provide the means to gather information for standardisation purposes and reduce the possible cost implications.

Recommended option for the norming standardisation process:

We recommend that, where no-coincidence between the norming process and the professional judgement of teachers exists, the separation protocol is adopted and, where coincidence between the norming process and the professional judgement of teachers does exist, the accommodation protocol is adopted.

This would ensure teacher judgements are unaffected, additional costs are minimised and the combination of approaches to gather standardisation data is more likely to be viewed favourably by the profession.

Recommendation

SROs are asked to:

   a) Agree to the recommended option for standardisation, which will then be communicated to ACER

   b) Agree to a communication strategy being developed to support the standardisation process using the agreed protocol

Document 4 – SRO SNSA Project – Long Scale Equating Study (LSES) Planning Report – 8 December 2017

NATIONAL IMPROVEMENT FRAMEWORK – NATIONAL STANDARDISED ASSESSMENTS PROJECT – Long Scale Equating Study (LSES) Planning Report

NIF SRO MEETING – 08 12 2017

Purpose

- To provide an overview of the Long Scale Equating Study planned for February 2018

Proposals for consideration

SROs are asked to note and comment on the Long Scale Equating Study planned for February 2018

Background

Long Scale Equating Study for Calibration

The standardisation options chosen on 18.05.17 involved sampling of a number of children in the year groups between the regular assessment (P1, 3,7 and S3) in order to create a robust long scale. In order to obtain more robust equating parameters, it is proposed to conduct equating studies in the year groups in between the assessed years with students completing assessments with items from the upper and lower year levels.

Vertical equating across four year stages is a serious undertaking. We will have more confidence in the reporting on long scales if we do it properly and we probably only need to do it once in several years.

Key features of the proposed equating study.

Fixed forms (not adaptive)
Each child will undertake a single form of the assessment that combines items from the assessments on either side (eg P2 pupils complete an assessment comprising items from P1 and P4). This is called "common item equating". This will entail ACER constructing special forms for the equating study.

The sample
A range from 400-600 children would be required for each year group by subject area. This equates to 2400-9600 children depending on whether a child does one or more assessments. A 'convenience sample' is sufficient for this study: that is, Local Authorities and/or schools could be invited to participate (in a manner similar to that used successfully to engage Local Authorities in the LA Trialling Phase). A spread of demographic characteristics is preferred but not essential as it is the questions that are being calibrated and not the child's performance.

We recommend that we invite volunteers while also targeting LAs of different types to ensure a good spread. We also recommend encouraging whole class/school participation to reduce the number of schools directly involved.

We will work up an offer to LAs and encourage participation formally and informally.

Timing

ACER recommend that the middle of the school year (February) would be the optimal time to conduct the equating studies. P1 items would have an increasing tendency to become too easy for P3 by the end of the year and, conversely, P4 items may be too difficult for P2 pupils at the beginning of the year. The same reasoning (though to a lesser extent) would apply to the other year groups. We propose a four week time frame, running from 5th February to the 2nd of March 2018, recognising that there are midterm school holidays in February.

Reporting to schools

Reports to schools on the equating studies cannot replicate the same formats to those provided for P1, P4, P7 and S3, for individuals and groups (classes). However they will have some of the same features.

Single report: Group Diagnostic
Features:
- Raw score for each learner
- List of items presented and learner result: correct, incorrect or not attempted
- Percentage correct for group
- Table and summary modes
- No capacity rating or item difficulty
- Generated on demand through teachers' SNSA account

Technical support

As the teacher at these stages would not normally have been trained we need to assume minimal knowledge of the SNSA platform. We will therefore provide a similar level of documentation and Service Desk support as with the first norming study.
We will also make all the online help guidance available.

It is also more difficult to preload these assessments as we are not preselecting them as in the norming studies.

A number of technical options have been discussed and risk assessed. We recommend the following with the pros and cons

Option 2 - Assigning assessments to all learners in selected LAs

Pros	Cons
• Reduced workload for schools and teachers • Less exposure of LSES to the wider school population • Less scope for learners to be assigned incorrect assessment type • More control over participation • Likely to be less traffic for Service Desk	• Learners may be assigned assessments that teacher does not wish to administer • Less flexibility and autonomy • If additional recruitment required, scripts would need to be modified and/or re-run • More technical processes to be completed within timeframe with concomitant risk • Learners may complete unexpected assessment

Communications

As with previous norming study timely communications will be crucial and multi layered.

- SG contact LAs inviting participation prior to Christmas
- SG contacts partners such as professional associations prior to Christmas
- SG contacts LAs to confirm level of participation in January
- Service desk contacts all participating schools explaining process and providing support documentation
- Service desk reminder on first day of window (all sampled schools)
- Service desk reminder on wed. of last week (just uncompleted schools)
- Message to all participating schools week after completion: thanks

Time-line of activity

Activity	Timing
Workshop with SG	6 December
Decision required on recruitment and assigning options	12 December
SG sign off of LSES assessments and reports	21 Dec - 5 Jan
Release of LSES assessments and reports onto SNSA platform once LAs confirm participation	15 January
Service Desk contacts Account Managers at participating schools with administration information	From 22 January
LSES window	*5 Feb - 2 March*
Psychometric analysis	March & April
Construction of long scales	April-June
Adaptation of reports to include long scales	May-June

Recommendation

SROs are asked to:

- note and comment on the Long Scale Equating Study planned for February 2018

From: [redacted] [mailto:redacted]
Sent: 07 August 2017 09:01
To: [redacted]
Cc: [redacted]; [redacted]

Subject: RE: SNSA - Sampling Teleconference call

Thanks again [redacted]

I had the thought over the weekend that it might be a good idea to have the student age as the final stratification variable, and use this as the final sort variable. That would give us the best chance of capturing the distribution of ages for each stage in the sample, and any groupings (e.g. for analysis and reporting) can be decided from there.

I do think this would be slightly better controlled if we drop the urban/rural variable, and then within each grouping defined by LA x Deprivation quintile x Gender, we sort ('continuously') by student age.

Just to emphasize, it's an equal probability sample. The sort order does not change the selection probability, nor the expected frequencies of outcomes. For a subgroup (e.g. of a certain age) that contains 5% of the population, we expect that about 5% of the sample will have representatives from that subgroup. The sort order (and systematic sampling) just puts some controls around the key variables so that we don't by chance get something a bit more out of proportion than we'd like.

How does that sound?

[redacted]

From: [redacted], [[redacted]

Sent: Saturday, 5 August 2017 1:30 AM
To: [redacted]
Cc: [redacted]; [redacted]

Subject: RE: SNSA - Sampling Teleconference call

I've produced some frequencies of month of birth by stage as that was quick and easy. This should give you a good enough idea of how the pupils will be distributed across the 4 age categories but let me know if you need something more detailed.

Ignore the day value in the birth dates – we only hold month and year of birth so all the days default to 15. I've suppressed very small numbers with *.

[redacted]

From: [redacted] [[redacted]
Sent: 04 August 2017 11:17
To: [redacted] [redacted] [redacted]

Subject: RE: SNSA - Sampling Teleconference call

Thanks for this.

Nationally, or nationally by stage will be fine.
[redacted]

From: [redacted] [[redacted]

Sent: Friday, 4 August 2017 7:01 PM
To: [redacted]; [redacted]; [redacted]

Subject: RE: SNSA - Sampling Teleconference call

Hi [redacted]

[redacted], can you give a rough estimate of the distribution of pupils across these age groups?

Is this something you could do for [redacted] using September 2016 data for P1, P4, P7 and S3 pupils who are on the roll of publicly funded schools in Scotland. Obviously, the date of births will need to be a year earlier than those provided to [redacted], such that for September 2016 they need to be as follows:

P1 Sample
Group 1: Children born November 2011 to February 2012 (i.e. the youngest children)
Group 2: Children born July 2011 to October 2011 (i.e. 2nd youngest group)
Group 3: Children born March 2011 to June 2011 (i.e. 2nd oldest group)
Group 4: Children born on or before February 2011 (i.e. the oldest group) – these are all children who deferred entry to Primary 1 (most of which will be born in Jan/Feb 2011).

P4 Sample
Group 1: Children born November 2008 to February 2009 (i.e. the youngest children)
Group 2: Children born July 2008 to October 2008 (i.e. 2nd youngest group)
Group 3: Children born March 2008 to June 2008 (i.e. 2nd oldest group)
Group 4: Children born on or before February 2008 (i.e. the oldest group) – these are all children who deferred entry to Primary 1 (most of which will be born in Jan/Feb 2008).

P7 Sample
Group 1: Children born November 2005 to February 2006 (i.e. the youngest children)
Group 2: Children born July 2005 to October 2005 (i.e. 2nd youngest group)
Group 3: Children born March 2005 to June 2005 (i.e. 2nd oldest group)
Group 4: Children born on or before February 2005 (i.e. the oldest group) – these are all children who deferred entry to Primary 1 (most of which will be born in Jan/Feb 2005).

S3 Sample
Group 1: Children born November 2002 to February 2003 (i.e. the youngest children)
Group 2: Children born July 2002 to October 2002 (i.e. 2nd youngest group)
Group 3: Children born March 2002 to June 2002 (i.e. 2nd oldest group)
Group 4: Children born on or before February 2002 (i.e. the oldest group) – these are all children who deferred entry to Primary 1 (most of which will be born in Jan/Feb 2002).

I presume [redacted] is just needing an overall 'National' picture (i.e. not broken down by school/LA).

Thanks in advance.

[redacted]

****
****

gov.scot

[redacted]
Statistician
Education Analytical Services
Learning Analysis Unit

T: [redacted]
M: [redacted] (if required)

Scottish Government
Victoria Quay
Commercial Street
Edinburgh EH6 6QQ

From: [redacted] [[redacted]
Sent: 04 August 2017 01:48
To: [redacted]; [redacted]; [redacted]

Subject: RE: SNSA - Sampling Teleconference call

[redacted], can you give a rough estimate of the distribution of pupils across these age groups?
I presume that the two middle 4-month groupings will be roughly the same size. How will the oldest and youngest groups compare?
Thanks,
[redacted]

From: [redacted] [mailto: [redacted]]
Sent: 04 August 2017 11:19
To: [redacted] [redacted] [redacted]
Subject: RE: SNSA - Sampling Teleconference call

Hi [redacted]
Thanks for the clarification re location.
Yes I will make population sample comparisons for us to check over before we finalise the sample.
[redacted]

From: [redacted] [redacted]
Sent: Friday, 4 August 2017 7:17 PM
To: [redacted]; [redacted]; [redacted]
Subject: RE: SNSA - Sampling Teleconference call

Hi [redacted]

If the SEEMIS extract being used to draw the sample will not include either the child's home postcode or their Urban/Rural 6-fold indicator (which is assigned using the child's home postcode) then I'm happy to say that this variable is not included in the stratification (as including the LA should pretty much take this into account).

There are obviously some urban/rural spread within LAs (e.g. the Highlands will have many small rural areas, but also includes the city of Inverness which I suspect will be an Urban area). Similarly, Edinburgh will largely be very Urban (but areas of the LA on the outskirts will probably be classed as "accessible town" areas). However, I don't this should be a major issue to worry about.

I think once the sample is drawn, it would be useful to check that the sample is 'representative' of the 'population' (by providing some summary of the number of children who have been drawn from each strata variable, etc.). Would that be OK to do?

[redacted]
****
****

![gov.scot logo]

[redacted]**Statistician**
Education Analytical Services
Learning Analysis Unit

T: [redacted]
M[redacted] (if required)

Scottish Government
Victoria Quay
Commercial Street
Edinburgh EH6 6QQ

From: [redacted] [mailto: [redacted]]
Sent: 04 August 2017 01:24
To: [redacted]; [redacted]); [redacted]
Subject: RE: SNSA - Sampling Teleconference call

Hi [redacted]

Yes, good point about the location variable. I must admit I had in mind that location would be school level as well (I had it stuck in my head that we'd be incorporating the school into the stratification). Given that we're not including the school, it makes better sense to use the student location if it can be collected. I remember in our discussion that any location effects were probably already captured within LA, so if it turns out that it is not straightforward to have a student level location level, could we agree to reduce the stratification to LA, Deprivation, Age/stage, Gender?

Re the number of strata, I think the main point is that it will systematic equal probability sampling down the list of records. If they were completely randomly sorted, then we'd EXPECT that an LA with 10% of the population would have about 10% of the sample, but we might get 'bad' samples where a higher proportion of records from that LA were included. By sorting the list by LA, and then systematically sampling, we'll get a sample of more or less exactly 10% from the LA. Within that LA, if the records are sorted by derpivation quintile and a systematic sample of (now 80 as 10% of the full sample) will capture more or less exactly 16 from each quintile. And so on....
Hope that makes some sense :)

Thanks for the clarification on the variables and the age ranges, happy with these.

[redacted]

___

From: [redacted]
Sent: Friday, 4 August 2017 10:40 AM
To: [redacted]; [redacted]; [redacted]

Subject: RE: SNSA - Sampling Teleconference call

Thanks **[redacted]**, noted

[redacted]

Sorry [redacted], just amend the date of birth ranges for the "age groupings"…

[redacted]

****
****

[redacted]
Statistician
Education Analytical Services
Learning Analysis Unit

T: [redacted]
M: [redacted] (if required)

Scottish Government
Victoria Quay
Commercial Street
Edinburgh EH6 6QQ

Hi [redacted]

This seems to nicely reflect our discussion yesterday.

At present, the list of pupil variables you have (as shown in the table below) does not have the 'location' (what we call the 6-fold urban/rural classification). You will need to see if SEEMIS can add that (similar to what they're doing for deprivation (what is called SIMD).

Further information about this classification can be found here: http://www.gov.scot/Topics/Statistics/About/Methodology/UrbanRuralClassification

Upon further thinking about the stratification variables, I'm still struggling to get my head around how the sample will be drawn from these. As discussed over the phone, using all of the below strata variables would equate to approx. 8,000 possible strata. However, we only propose to sample 800 pupils for each required sample. Therefore, I'm not fully sure how this squares up (800 doesn't feel a big enough sample to include/represent all of these strata variables)? Therefore, more information from you on how this would be done would be helpful here.

Also, slightly expanding on your Point 1 below, I thought it useful to clarify the current proposed list of strata variables:

1) Stratification
   a. Agreement to recommend the following stratification structure: LA(32); Location (Urban/Rural 6-fold) (6); Deprivation (SIMD Quintiles) (5); Age within stage (4); Gender (2)

In relation to the 4 groupings of age, can I suggest the following groups (based on the child's month/year of birth):

P1 Sample
Group 1: Children born November 2012 to February 2013 (i.e. the youngest children)
Group 2: Children born July 2012 to October 2012 (i.e. 2nd youngest group)
Group 3: Children born March 2012 to June 2012 (i.e. 2nd oldest group)
Group 4: Children born on or before February 2012 (i.e. the oldest group) – these are all children who deferred entry to Primary 1 (most of which will be born in Jan/Feb 2012).

P4 Sample
Group 1: Children born November 2009 to February 2010 (i.e. the youngest children)
Group 2: Children born July 2009 to October 2009 (i.e. 2nd youngest group)
Group 3: Children born March 2009 to June 2009 (i.e. 2nd oldest group)
Group 4: Children born on or before February 2009 (i.e. the oldest group) – these are all children who deferred entry to Primary 1 (most of which will be born in Jan/Feb 2009).

P7 Sample
Group 1: Children born November 2006 to February 2007 (i.e. the youngest children)
Group 2: Children born July 2006 to October 2006 (i.e. 2nd youngest group)
Group 3: Children born March 2006 to June 2006 (i.e. 2nd oldest group)
Group 4: Children born on or before February 2006 (i.e. the oldest group) – these are all children who deferred entry to Primary 1 (most of which will be born in Jan/Feb 2006).

S3 Sample
Group 1: Children born November 2003 to February 2004 (i.e. the youngest children)
Group 2: Children born July 2003 to October 2003 (i.e. 2nd youngest group)
Group 3: Children born March 2003 to June 2003 (i.e. 2nd oldest group)
Group 4: Children born on or before February 2003 (i.e. the oldest group) – these are all children who deferred entry to Primary 1 (most of which will be born in Jan/Feb 2003).

And just in case you're wondering, a "mothballed" school is a school that is technically still 'open'. However, there are currently no children in it so it is 'temporarily closed' until either children/families in the area need the school to re-open or for the school to be formally 'closed'.

Does that help?

[redacted]

****
****

gov.scot

[redacted]
Statistician

Education Analytical Services
Learning Analysis Unit

T: [redacted]
M: [redacted] (if required)

Scottish Government
Victoria Quay
Commercial Street
Edinburgh EH6 6QQ

From: [redacted] [mailto: [redacted]]
Sent: 03 August 2017 08:46
To: [redacted]; [redacted]; [redacted]
Subject: exclusions / non - attendance

[redacted], [redacted] and [redacted]

We discussed the idea of categories for exclusion and other non-response that might work in the SNSA context. (Just to emphasise though, I don't know whether it's going to be possible to implement this in the upcoming norming studies given that there is not much time left to do so.)

As a starting point, I extracted the international definitions used for PISA. These might have been modified somewhat ('adapted') to better suit the national context. They usually appear with a preamble along the lines that the intention is for the assessment to be as inclusive of students as possible, but that for various reasons a very small number of students may need to be excluded. I think the key component of each category is that condition the student has means that the he or she CANNOT take the test.

Within-school exclusions - exclusion of students from potential assessment because of one of the following:

They are functionally disabled in such a way that they cannot take the PISA test. Functionally disabled students are those with a moderate to severe permanent physical disability.

They have a cognitive, behavioural or emotional disability confirmed by qualified staff, meaning they cannot take the PISA test. These are students who are cognitively, behaviourally or emotionally unable to follow even the general instructions of the assessment.

They have insufficient assessment language experience to take the PISA test. Students who have insufficient assessment language experience are those who meet all the following three criteria: − they are not native speakers of the assessment language, − they have limited proficiency in the assessment language, and − they have received less than one year of instruction in the assessment language.

There are no materials available in the language in which the student is taught.

They cannot be assessed for some other reason as agreed upon.

The last category will likely be particularly problematic. I doubt there will be an easy mechanism for 'reaching agreement' during the field operation.

I just offer these as a starting point, but you may well have better categories that you've used before, or additional considerations that need to be taken into account.

In addition to within school exclusion codes, we will also need some other non-response codes – e.g. for absent, refusal, no longer enrolled etc.

Cheers,
[redacted]

From: [redacted] [mailto: [redacted]]
Sent: 03 August 2017 08:06
To[redacted]; [redacted] ; [redacted]

Subject: RE: SNSA - Sampling Teleconference call

Dear **[redacted]; [redacted];** and **[redacted]**

Thanks very much for the phone call yesterday, it was very useful. Thanks also **[redacted]**
 for the updated schools contact list. Dare I ask, can this list be supplemented with stage level enrolments, and/or enrolments by gender (overall or by stage)?

As promised, a summary below of outcomes from the call (**[redacted]**
 = **[redacted];**, **[redacted];**, **[redacted];** MM = me).

If I've missed any important points or misrepresented things, or other refinements or edits required, please do so!

1) Stratification
    a.   Agreement to recommend the following stratification structure: LA(32); Location (6); Deprivation (5); Age within stage (4); Gender (2)

2) School level exclusions
    a.   A common understanding that the following schools are to be excluded: all special schools (local authority and grant aided), independent schools, closed and mothballed schools. To be confirmed.

3) Pupil substitution

a.   [redacted] : There will be cases where teachers will exclude pupils from the assessment (e.g. a pupil who, due to their specific complex needs, may never be expected to achieve CfE Early Level.) There could also be data lost through absence, parent refusal etc. Response rates may not be as high as the very high rates achieved in SSLN – for example with less advanced notice given to schools about the norming study. A concern that this loss of data might affect the quality of the scaling and calibration.
    b.   MM: The scaling methodology is quite robust, so that even with a higher than expected loss of sample through refusal/non-response – say 10%, 20% or even more – it should not substantively affect our capacity to develop good scales of pupil ability and item difficulty.
    c.   Agreement to recommend that so long as ACER thinks that scaling and calibration will not be unduly compromised with a loss of data of that order through non-response and exclusions, that pupil substitution could be avoided.

4) Distinguishing between within school exclusions and other non-response (e.g. absence, refusal)
   a. Agreement that if possible, it would be useful to be able to distinguish between different types of non-response. This might involve an approach where teachers are given a series of codes for different categories of exclusion / non-response in a manner similar to PISA for example.
5) Data source
   a) Agreement that the table of SEEMiS variables (copied below) and the school details data (http://www.gov.scot/Topics/Statistics/Browse/School-Education/Datasets/contactdetails) have the necessary variables for stratification and for sampling.
   b) SG can provide the most up-to-date school details data from which schools to be excluded and closed and mothballed schools can be removed prior to sampling.

Field	Required	Field Length/Type	Nullable	Comments
SEEMiS ID	Y	char(25)	N	
Forename	Y	char(30)	Y	
Surname	Y	char(30)	Y	
Preferred name	Y	char(30)	Y	
SCN	Y	char(9)	Y	SCN = Scottish Candidate Number
SIMD	Y	int	Y	• Scottish Index of Multiple Deprivation, expressed as vigintile.
Gender	Y	char(1)	Y	• Currently M or F.
DOB	Y	datetime	Y	
Additional support needs	Y	bit	N	
Looked after (at home)	Y	bit	N	
Looked after and accommodated	Y	bit	N	
English as a second language	Y	bit	N	
Free school meal	Y	bit	N	
Ethnicity	Y	char(2)	N	
Stage	Y	char(2)	Y	• Technically nullable, but in practice this won't be.
Class/Registration group	Y	char(6)	Y	
School ID	?	char(7)	N	• Data is taken from a school, so this will only ever be 1.
Teacher ID*	?	[char(12)]	Y	• Can be empty if no registration class exists. • Can be more than 1.

Best Regards,
[redacted]

Hi [redacted],

I've attached an up to date version of the school contacts list. This has the school closures and new school openings that we are currently know about. I've sorted the list to separate out the schools that I understand should be excluded, i.e. all special schools (local authority and grant aided), independent schools, closed and mothballed schools.

There will inevitably be some further changes between now and when the testing is done but these should hopefully be minor such as very small schools being mothballed.

Let me know if you need any more help with this.

Thanks,

[redacted]

School, Staff and Pupil Census Statistics
Education Analytical Services
Scottish Government
[redacted]

From: [redacted] [mailto: [redacted]]
Sent: 02 August 2017 08:08
To: [redacted]; [redacted]; [redacted]
Subject: RE: SNSA - Sampling Teleconference call

Hi **[redacted]**

I've drafted some ==talking points== for our discussion this morning, see attached
Talk soon.

Regards,
[redacted]

==Talking Points attached as separate document (Number 6)==

From: [redacted]
Sent: Friday, 28 July 2017 6:54 PM
To: [redacted] <[redacted] >; [redacted]
[redacted]

Subject: FW: SNSA - Sampling Teleconference call

Hi all

To see in preparation for next week's meeting.

[redacted]

****
****

gov.scot

[redacted]
Statistician
Education Analytical Services
Learning Analysis Unit

T: [redacted]

M: [redacted] (if required)

Scottish Government
Victoria Quay
Commercial Street
Edinburgh EH6 6QQ

From: [redacted]
Sent: 24 July 2017 13:34
To: [redacted]
Subject: FW: SNSA - Sampling Teleconference call

[redacted]

Does this make sense on the "extract" front?

Cheers
[redacted]
Programme Manager
Directorate for Learning | The Scottish Government
2B South, Victoria Quay, Edinburgh
[redacted]
[redacted]

From: [redacted]
Sent: 24 July 2017 11:53
To: [redacted][redacted]
Subject: RE: SNSA - Sampling Teleconference call

Not much to add really – I suppose it's possible that the term "extract" was being bandied around and it was assumed that this would be something that ACER would explicitly receive from SEEMiS, but there should be nothing that they won't have already and, given the API processes running between SEEMiS and ACER, the ACER version of the data would never be more than 24 hours out of date at any given time.

From: [redacted]
Sent: 20 July 2017 16:45
To: [redacted]; [redacted]
Subject: RE: SNSA - Sampling Teleconference call

Hi **[redacted]**

 [redacted] is on a long weekend. But I think we might be talking at cross purposes here. I wasn't at the meeting, so I may be going off track, as per **[redacted]** note I think the extract should be from SNSA as it is already up to date at any point and you are right, ACER already have those details. As P1's don't get populated until end August, we'd have to wait for this to complete. Other risk factor to bear in mind is we may not have 100% of LA's which will have implications for sample size (hopefully not an issue if all data sharing issues resolved soon).

[redacted] may have more info when he is back on Monday

[redacted]
Service Manager
[redacted]
Internal: **[redacted]**
[redacted]

Digital Learning & Teaching
Scottish Government
Victoria Quay, Area 1H-North
Edinburgh, EH6 6QQ

From: [redacted]

Sent: 20 July 2017 15:41
To: [redacted]; [redacted]
Subject: FW: SNSA - Sampling Teleconference call

[redacted]

Any views on the 'extract'?

Cheers

[redacted]
Programme Manager
Directorate for Learning | The Scottish Government
2B South, Victoria Quay, Edinburgh
[redacted]
[redacted]

From: [redacted]
Sent: 20 July 2017 12:41
To: [redacted]
Subject: RE: SNSA - Sampling Teleconference call

Thanks for this [redacted]

From one of the previous SNSA meetings I attended, my understanding was that ACER were to receive an 'extract' of pupils from SEEMIS for use in the SNSA (e.g. a list of pupils in P1, P4, P7 and S3) – as this was need to obtain specific pupil characteristic information (as outlined below) – as this would help with any 'reporting' of results by pupil characteristics.

The question I have is when they are expecting to receive this file. If it's August, then this would likely be the best 'population' file they would have to draw the 'normalising' sample from (particularly for the sample needed for November).

This file could also be used to draw the necessary sample needed for the March 'normalising' sample. However, children do unfortunately move around Scotland, so there may be some merit to get a more up-to-date list from SEEMIS when drawing that sample (to minimise the risk that a chosen child is no longer in a particular school (or indeed in Scotland)).

We will discuss this with [redacted] and see where we get to. I get a sense he is looking more for advice on what 'factors' need to be considered when drawing the sample (e.g. deprivation, gender, ages within stages, LAs, etc.) rather us actually drawing the sample for him.

Will keep you posted.

[redacted]

****
****

gov.scot

[redacted]
Statistician
Education Analytical Services
Learning Analysis Unit

T: [redacted]
M: [redacted] (if required)

Scottish Government
Victoria Quay
Commercial Street
Edinburgh EH6 6QQ

From: [redacted]
Sent: 20 July 2017 12:13
To: [redacted]
Subject: FW: SNSA - Sampling Teleconference call

[redacted]

Attached information may be useful for [redacted], I'll let you decide. It's something we think ACER already have anyway.

Cheers

[redacted]
Programme Manager
Directorate for Learning | The Scottish Government
2B South, Victoria Quay, Edinburgh
[redacted]
[redacted]

From: [redacted]
Sent: 20 July 2017 11:45
To: [redacted]
Cc: [redacted]
Subject: RE: SNSA - Sampling Teleconference call

Below is the list of fields that we get for pupils, along with the format. Not to say that SEEMiS don't potentially hold some of this information in more granular forms (eg some that are yes/no might actually have degrees within SEEMiS but that's not something we would have access to, and probably don't have any right to have access to since it would be beyond what's been agreed).

Attached is a spreadsheet of dummy data from SEEMiS based on that specification.

The pupil info could also be cross-referenced (via a look-up of SEEMiS ID -> SEED number) against the school details that are available at:

http://www.gov.scot/Topics/Statistics/Browse/School-Education/Datasets/contactdetails

This has school roll, the urban/rural classification in two forms (6 stage and 8 stage), proportion of pupils from minority ethnic groups and proportion of pupils drawn from the most deprived quintile.

Field	Required	Field Length/Type	Nullable	Comments
SEEMiS ID	Y	char(25)	N	
Forename	Y	char(30)	Y	
Surname	Y	char(30)	Y	
Preferred name	Y	char(30)	Y	
SCN	Y	char(9)	Y	SCN = Scottish Candidate Number
SIMD	Y	int	Y	• Scottish Index of Multiple Deprivation, expressed as vigintile.
Gender	Y	char(1)	Y	• Currently M or F.
DOB	Y	datetime	Y	
Additional support needs	Y	bit	N	
Looked after (at home)	Y	bit	N	
Looked after and accommodated	Y	bit	N	
English as a second language	Y	bit	N	
Free school meal	Y	bit	N	
Ethnicity	Y	char(2)	N	
Stage	Y	char(2)	Y	• Technically nullable, but in practice this won't be.
Class/Registration group	Y	char(6)	Y	
School ID	?	char(7)	N	• Data is taken from a school, so this will only ever be 1.
Teacher ID[†]	?	[char(12)]	Y	• Can be empty if no registration class exists. • Can be more than 1.

From: [redacted]
Sent: 20 July 2017 11:13
To: [redacted]
Cc: [redacted]
Subject: FW: SNSA - Sampling Teleconference call

[redacted]

Do we have information that may be useful to address [redacted]'s request?

"I understand that I will not have access to the actual database until I am in the UK (i.e. 4-8 September). If possible, it will be extremely helpful for my preparations if I can access a codebook for that database, and some dummy data. I'm not sure yet exactly the mechanisms for getting those

things prior to my visit, or for access to the data itself when I am in Edinburgh. Let me discuss with [redacted] first, and I may need to come back to you after that. "

Regards

[redacted]
Programme Manager
Directorate for Learning | The Scottish Government
2B South, Victoria Quay, Edinburgh
[redacted]
[redacted]

From: [redacted]
Sent: 18 July 2017 09:33
To: [redacted]
Cc: [redacted]; [redacted]; Leng D (David); [redacted]; [redacted]; [redacted]
Subject: RE: SNSA - Sampling Teleconference call

Hi [redacted]

The Data Specification and Guidance Notes we have for our Pupil Census is attached below. However, I suspect this will not match exactly with how the 'raw' data is held on SEEMIS as some LAs may use their own local code lists (i.e. some information may be 'mapped' onto their own local code lists on to our required standard Codes).

http://www.gov.scot/Topics/Statistics/ScotXed/SchoolEducation/SchoolPupilCensus/SurveyDocument ation

Also, we also add on additional variables to this data extract (e.g. we will add on the deprivation information based on the child's home postcode we receive).

Unfortunately, we will not receive this information for September 2017 until mid/late October 2017, so we won't ourselves have this to help pick a sample for the 2017/18 school year.

However, a potential option to consider is that we could use our September 2016 (2016/17 school year) file to pick a sample of children for P4, P7 and S3 for the 2017/18 school year using P3, P6 and S2 pupils from last year. The issues with doing this are some children will have moved school (or left Scotland) and some children have now moved to Scotland (which we currently don't know about). We still would not have anything to sample this year's P1 pupils (as we do not currently collect individual child-level for children in early learning and childcare settings).

Anyway, I hope this helps with your thinking in relation to this exercise.

[redacted]

****
****

[redacted]
Statistician
Education Analytical Services
Learning Analysis Unit

T: [redacted]
M: [redacted]if required)

Scottish Government
Victoria Quay
Commercial Street
Edinburgh EH6 6QQ

From: [redacted] [[mailto:](#) [redacted]]
Sent: 18 July 2017 09:04
To: [redacted]
Cc: [redacted]; [redacted]; Leng D (David); [redacted]; [redacted];
MacDougall A (Audrey); [redacted]

Subject: RE: SNSA - Sampling Teleconference call

Thanks very much **[redacted]** for organising that date so quickly.

Your question about 'what sampling frame' is a good one! Yes I think it will be an extract of the SEEMIS database.

I understand that I will not have access to the actual database until I am in the UK (i.e. 4-8 September). If possible, it will be extremely helpful for my preparations if I can access a codebook for that database, and some dummy data. I'm not sure yet exactly the mechanisms for getting those things prior to my visit, or for access to the data itself when I am in Edinburgh. Let me discuss with **[redacted]** first, and I may need to come back to you after that.

In the meantime I will develop a specific set of preparation tasks prior to our meeting over the next week or so, and share these with you.

Regards,
[redacted]

From: [redacted][redacted]

Sent: Monday, 17 July 2017 9:21 PM
To: [redacted] <[redacted] >
Cc: [redacted];[redacted]; [David.Leng@gov.scot](#); [redacted]; [redacted]
[Audrey.MacDougall@gov.scot](#); [redacted]

Subject: RE: SNSA - Sampling Teleconference call

Hi [redacted]

Thanks for this. You will see that I've set up a meeting for 2 August to discuss this further.

The report for the PISA 2015 states that in
Scotland: http://www.gov.scot/Publications/2016/12/7252/2

11. The school sample was randomly selected by NFER following submission of sampling forms to the consortium. The sample was stratified on the basis of previous exam performance (split into five categories), whether schools were publicly funded or independent, urban/rural location and school size, and whether schools were single-sex or mixed.

12. In total, 109 secondary schools participated in the survey. One hundred and two of these were from the main sample (87 per cent response rate), and seven from the back-up samples (resulting in 93 per cent participation rate after replacements were added in). This exceeded the OECD's minimum standard of 85 per cent participation.

13. Within each school 40 students were randomly sampled by NFER using software supplied by the Consortium. In total 4,283 students were drawn in the sample. Schools were able to withdraw a certain number of students where it was deemed that participation would be difficult due to additional support needs or language issues. Similarly students that had left the school in the interim were not considered part of the target sample. In total 3,610 students were deemed eligible participants. Of these a total of 3,123 students took part, with the balance being those who did not wish to take part (both students and their parents were given the opportunity to opt out of the survey), those who were absent on the day of the test or were withdrawn by the school because of their additional support needs.

Just to clarify my Urban/Rural reference below, then this can either be 2, 3, 6 or 8-fold (the most common used being 6 or 8-fold):
http://www.gov.scot/Topics/Statistics/About/Methodology/UrbanRuralClassification

In relation to the stratification structure (and strata), then I think we need to make sure we don't have too many (particularly if we're only looking to select 1,600/2,400 pupils from each stage). For example, even with:

Gender (M/F), Deprivation (Quintiles (or Deciles)), Local Authority (32 LAs) and Ages within Stages (e.g. 2 age groupings – younger/older children at time of assessment)

We would have quite a few strata (2 * 5 * 32 * 2 = 640 strata).

I don't think we need Urban/Rural (as well as LA) - together with the fact that we don't see much of an effect on achievement of CfE levels by this breakdown.

You raise an interesting point about whether we would want to use these 'normalising' samples for further research/statistical secondary analyses (and not one I had thought of before). Again, something we could discuss on the call, but this links to other continuing discussions as to what information the Scottish Government will request/hold that is produced from the SNSAs.

A question I have to you is "**What sampling frame will the sample be drawn from?**". Will this be from a population extract you will be receiving from SEEMIS? If so, when do you expect to have this extract? If not, then what?

Looking forward to our catch up on 2 August, but feel free to email **[redacted]**/**[redacted]**/**[redacted]** /myself if you have any further 'data' queries.

[redacted]

[redacted]
Statistician
Education Analytical Services
Learning Analysis Unit

T: [redacted]
M: [redacted] (if required)

Scottish Government
Victoria Quay
Commercial Street
Edinburgh EH6 6QQ

From: [redacted] [mailto: [redacted]]
Sent: 17 July 2017 05:07
To:] [redacted]
Cc:] [redacted];] [redacted]; Leng D (David);] [redacted];] [redacted]; MacDougall A (Audrey);] [redacted]
Subject: RE: SNSA - Sampling Teleconference call

Hi] [redacted] (and hello again] [redacted] and team),

Yes, your mail provides a pretty good summary of the task ahead. I will be in Scotland in the week 4-8 September for selecting the samples. Prior to that time (or by early that week at the latest) we will need to have the stratification structure decided upon and population coverage and exclusions (nationally, school level, pupil level) identified and quantified.

As far as meeting times go, these will probably work best as your morning / my evening. 6pm in Melbourne is 9am in Edinburgh, so 6pm or later if necessary should be fine for me. Unfortunately, this week is a bit difficult, but next Tuesday or Wednesday 25- 26 July, or the same days in the following week would be best for me, but I'm happy to accommodate other days if that's necessary.

Your list of stratification variables looks to cover the major possibilities. Other variables to consider are attainment, school type, school size. It's a good idea to include variables that are likely to be strongly related to the outcome measures (i.e. achievement), so that puts attainment into consideration. It may be that some variables (e.g. attainment?) are available at some levels and not others. Also in Scotland's participation for both TIMSS and PISA, a 6 level location classification was used, rather than urban/rural, so something else to consider.

Below is my brief understanding of the stratification structure for recent surveys in Scotland included in my presentation. I've also attached that presentation.

- SSLN
    – By gender within school
    – Some monitoring of deprivation status profile
- PISA
    – School level
        • Funding (2) / attainment (6) / gender (3) / area (6)
    – Student level
        • Grade / gender
- TIMSS 2007
    – School type (3) / area (6) / deprivation status (4)

Re the use of pupil substitutes, my recommendation is to avoid this. Technically, substitutes are used to impute the outcomes of pupils who were sampled but did not participate in the assessment.   They are best used when other methods for addressing that non-response – for example weighting adjustments or other forms of imputation are considered inferior. I think we will

be able to make a better imputation of the nonrespondent's outcomes by drawing upon the actual outcomes of that nonrespondent when s/he does the online assessment at the time the teacher chooses, as well as modelling of outcomes and progress over time drawing on the analysis of the two norming samples' data. Implementing a pupil substitution approach will also throw up some technical and operational complications which we can talk through in our discussions. Nevertheless I recognise that this is something that your team have some experience with through SSLN and the weight of the argument might fall in favour implementing a pupil substitution procedure, in which case I'm sure we can work that in.

I really believe that these norming studies – as well as serving the immediate tasks of initiating the scaling and calibration, and forming student distributions – will be important sources of data for other purposes. I remember on my last visit [redacted] you were discussing your analysis work in relation to teacher judgements and this norming study data will be very valuable for that work. And perhaps also for other researchers. (One question will be whether these data will be made available for secondary analyses in the same way that TIMSS and PISA data are made public?) One quality indicator for the database will be the extent to which decisions such as coverage, exclusions, stratification and non-response adjustment are recorded and quantified. I don't think that will be a huge task – similar to filling in the sampling forms for PISA for example - but it will be worthwhile.

Looking forward to our further discussions as we prepare to implement this very exciting project!

Best Regards,
[redacted]

[redacted]
[redacted]
Australian Council for Educational Research

19 Prospect Hill Road Camberwell
Camberwell VIC 3124 Australia
t: [redacted]
e: [redacted]

www.acer.edu.au

Melbourne | Sydney | Brisbane | Perth | Adelaide | Dubai | Delhi | Jakarta | London

From:] [redacted]] [redacted]
Sent: Friday, 14 July 2017 8:09 PM
To: [redacted]< [redacted]>
Cc:] [redacted]] [redacted] David.Leng@gov.scot;] [redacted]] [redacted] Audrey.MacDougall@gov.scot;] [redacted]
Subject: SNSA - Sampling Teleconference call

Hi [redacted]

As promised yesterday, I said that I would try to set up a meeting with you to discuss how the 'sampling' of children for the 'normalising' needed for the Scottish National Standardised Assessments (SNSA). **Can you let me know what days/times would be best for you (especially given the time difference)?**

To summarise my understanding from yesterday's meeting. We need to select a sample of pupils for the 'normalising' samples (i.e. for P1, P4, P7 and S3 children to undertake the SNSA is a fixed 'window' – i.e. 3 weeks in November 2017 and March 2018).

Each 'normalisation' sample (November & March) would require:

- 1,600 P1 pupils
- 2,400 P4 pupils
- 2,400 P7 pupils
- 2,400 S3 pupils
- **8,800 pupils in total.**

I presume we will need to sample different pupils for each 'normalising' sample (i.e. no pupil would appear in both each of these samples)? Therefore, we need to pick **17,600 pupils** in total (out of the approx. **210,000 P1/4/7/S3 pupils** we would have available to select from).

We need to consider what 'stratification' criteria we need to consider to ensure the 'normalisation' is as representative of Scotland as possible. For example:

- Gender
- Ages within stage
- Deprivation
- Local Authority
- Urban/Rural

[redacted] – what 'strata' did SSLN use?

Also, to consider having a 'substitution' list too (e.g. if a sampled pupil is deemed not appropriate to be included in the normalisation sample, then there is a similar pupil selected to substitute for them). For example, a pupil may be selected who, due to their specific complex needs, may never be expected to achieve CfE Early Level. Therefore, there is probably a sound argument that these pupils should not 'deselected' in the normalising sample and replaced by another 'similar' pupil.

Also, as this sample is needed for the 'normalisation' of English medium SNSAs, then it probably also makes sense not include P1 and P4 pupils based in Gaelic Medium Primary Classes for the 'literacy' sample, and P1 pupils based in Gaelic Medium Primary classes for the 'numeracy' sample (as there is no 'expectation' that these pupils should be achieving the same CfE level as pupils in equivalent English Medium classes).

It was also agreed that we should not include pupils based independent schools, or pupils in standalone special schools/units, in the 'normalising' samples (as the SNSA is to normalise for pupils for whom the assessment is for – i.e. primarily pupils in publicly funded primary and secondary schools).

There is also on going consideration for an 'equating' sample (Jan/Feb 2018) to be drawn, which would involve around 400 pupils from each P2, P3, P5, P6, S1, and S2 stage (i.e. 2,400 pupils) – the purpose of this sample being to assist in assessing the appropriateness of the difficulty of items in each stage-specific SNSA. The thinking here, though, is that local authorities/schools would 'opt in' to be a part of this work.

I look forward to hearing from you in order to set this up. I have a webex / teleconference account and will be able to give you the local dial in details

[redacted]

gov.scot

[redacted]
Statistician
Education Analytical Services
Learning Analysis Unit

T: [redacted]
M: [redacted] (if required)

Scottish Government
Victoria Quay
Commercial Street
Edinburgh EH6 6QQ

==Document 6 "Talking points" from email 2nd August – email chain SG official and ACER==

Talking points: Teleconference: Sampling for SNSA Norming Studies

1) Population coverage
    a. LA participation – DPA issues
    b. Public sector only
    c. Any other reductions on population coverage?
    d. Documenting and quantifying

2) School level exclusions
    a. Special schools (?) (see School Contact Details file discussed below)
    b. Other?
    c. List of school IDs to link to SEEMIS?
    d. Documenting and quantifying

3) Student level exclusions
    a. Can some or all of these be identified on the SEEMIS database prior to sampling?
    b. Will some only be identifiable after contact with the school has been made?
        i. Operation for coding non-respondents as exclusions/absentees etc.
    c. Standard definitions used in Scotland that are understood at the school level?
    d. Distinction between exclusion and non-response
        i. Any student identified for exclusion should NOT be substituted
        ii. A population estimate of exclusion rates will be an outcome of the survey

4) Non-response (e.g. parent/student refusal, illness)

   i. Operational implications (e.g parent permission / assessments over a period of time, to cover absence on a particular day)
    1. If assessments can be conducted over a period of time (e.g. a week or more), absence due to illness etc. should be minimal
  ii. Use of pupil substitution?
    1. If so, we need to retain the IDs and background data of both the SAMPLED and the substitute

5) Data
 a. ACER UK confirms that they have access to the data fields listed in the table in [redacted] mail:

Field	Required	Field Length/Type	Nullable	Comments
SEEMiS ID	Y	char(25)	N	
Forename	Y	char(30)	Y	
Surname	Y	char(30)	Y	
Preferred name	Y	char(30)	Y	
SCN	Y	char(9)	Y	SCN = Scottish Candidate Number
SIMD	Y	int	Y	• Scottish Index of Multiple Deprivation, expressed as vigintile.
Gender	Y	char(1)	Y	• Currently M or F.
DOB	Y	datetime	Y	
Additional support needs	Y	bit	N	
Looked after (at home)	Y	bit	N	
Looked after and accommodated	Y	bit	N	
English as a second language	Y	bit	N	
Free school meal	Y	bit	N	
Ethnicity	Y	char(2)	N	
Stage	Y	char(2)	Y	• Technically nullable, but in practice this won't be.
Class/Registration group	Y	char(6)	Y	
School ID	?	char(7)	N	• Data is taken from a school, so this will only ever be 1.
Teacher ID†	?	[char(12)]	Y	• Can be empty if no registration class exists. • Can be more than 1.

School Contact details (Sep 2016), from http://www.gov.scot/Topics/Statistics/Browse/School-Education/Datasets/contactdetails

 b. Sufficient basis for determining school level exclusions of special schools?
 c. A school level deprivation average?
  i. I presume SIMD vigintiles in above table are student level, yes?
 d. Stage enrolments is missing – can they be linked in from somewhere?

6) Recommended stratification
 a. Using above data, possible stratification (for each stage):
  i. LA
  ii. Location (2? 6? 8?)
  iii. Deprivation (school level) (4?)
  iv. School
  v. Age within stage (4)
  vi. Gender (2)
 b. Other possibilities? (e.g. free meals/ proportion minority ethnic/ proportion most deprived quintile?)
 c. Process for formal agreement of stratification

7) Access to dummy data

      a. ACER UK making a request for a 'training set' for my use, with fictional data.

NATIONAL IMPROVEMENT FRAMEWORK – NEW NATIONAL STANDARDISED ASSESSMENTS

[sections 1 to 3 redacted – out of scope of request]

4. Scottish sample for standardisation

The aim is for pupil results to be standardised against a representative Scottish sample of pupils. Those assessments currently used in schools are standardised across a wider sample, usually UK based. It takes time for a large enough sample of data to be collected and analysed in order to be used in the assessment process. Given the timescales for this project, it is likely there will not be sufficient time for a fully Scottish sample to be developed.

We are proposing to allow contractors to use existing standardisation methods at the beginning of the contract if necessary but that they should commit to developing and implementing a Scottish standardisation process as soon as possible during the contract period.

The sample will be specific to the assessments used and will not be available to be used by a different supplier at the end of the contract period.

[section 5 – out of scope of request]

6. National level datasets

The assessments will provide different levels of data to different parts of the education system. Our development work has identified necessary detail at the pupil, class, school and local authority level and these elements are provided for within currently available products.

If we receive pupil level data at the national level we will have the ability to undertake full pupil characteristic analysis, bearing in mind that as pupils will take the assessments at a time decided by their teachers rather than all during a restricted period, the results will not be properly comparative and able to be aggregated. If we do not specify the ability to receive detailed data, there will likely be additional costs and problems with requesting data during the contract period, if additional analysis becomes necessary.

However there are implications from the SG holding data that can identify individual pupils. It may be possible for us to collect the data without unique pupil identification numbers but with pupil characteristic information.

We do not yet know what, if anything, we want to do with national level data. It is our intention to ask in the requirements for contractors to have the ability to provide a detailed dataset with pupil level data and pupil characteristics. The national level data we actually receive and use will be determined during the development phase of the contract.

[redacted sections 7 – 11 – out of scope of request]

Learning Directorate
March 2016

Any analysis or advice (internal or external) provided to the Scottish Government regarding the reliability/statistical robustness of the Scottish Survey of Literacy and Numeracy (SSLN)

Document 8 SSLN Desk Instructions – Time Series methodology Note 2014

SSLN Time Series Note 2014

The SSLN survey design resulted in some items being released after the 2012 survey and replaced for 2014 survey, meaning there were some assessment booklets that had new content in 2014. There were also some additional booklets added in for the 2014 survey. I have performed a series of data checks on the 2012 and 2014 SSLN data to assess whether the results can be compared. The main steps in the process have been outlined below.

All analysis used the final pupil level datasets as the starting point, where pupils had to complete both the reading assessments, a written booklet and an online assessment, to be included. For the purposes of the time series work, the 2012 reading data was revised to incorporate a small amount of Gaelic data and to change the deprivation category of independent pupils from the least deprived category to the unknown deprivation category, which ensured the methodology was consistent with the numeracy survey analysis.

I started by producing unweighted reading performance by stage tables for **all** booklets completed in 2012 and 2014. Due to booklet allocation design, in 2014 pupils participating in the reading assessment would have completed one of the following combinations:

- two booklets consistent between 2012 and 2014 (trend booklets)
- one booklet consistent between 2012 and 2014 and one booklet containing replaced items
- two booklets added to the 2014 survey (new booklets)

To investigate whether the results were comparable we wanted to ensure those pupils who completed the new booklets added in 2014 had similar performance to the pupils completing the trend booklets, and the overall results.

I identified any data items that had been released after the 2012 survey and replaced for the 2014 survey and the new booklets added for 2014, and created a dataset that excluded any **booklets** in these categories. It was agreed with [redacted] and [redacted] that full booklets should be removed even if only a section of the booklet had been released, as the context of the booklet would have changed as a result. This left a dataset consisting of only the trend booklets that had remained consistent between 2012 and 2014.

This dataset was then refined to include only pupils who had completed two trend booklets and I recalculated the attainment results based on these remaining pupils. I used these results to re-run the unweighted performance table so I had unweighted results for **all** booklets and unweighted results for **trend booklets** for both 2012 and 2014. Pupils who completed two trend booklets consisted of just over 80% of the overall pupils in each stage in 2012 and about 70% of the overall total in 2014.

This process was then repeated to create a dataset and results based on pupils who had completed two new booklets added in 2014, This accounted for about 15% of pupils at each stage in 2014, but there were no comparable 2012 booklets.

These outputs enabled us to perform a series of unweighted within year and between year comparisons:

- Compare 2012 and 2014 results for all booklets, and also for trend booklets
- Compare results for all booklets against trend booklets for both 2012 and 2014
- Compare 2014 results for new booklets against trend booklets
- Compare 2014 results for new booklets against all booklets

The results showed that excluding booklets not consistent between the two surveys did not affect the in-year or between year comparisons as the results were similar. Hence these did not affect the overall picture of pupil performance, with the key messages staying the same regardless of whether replaced and new booklets were excluded or not.

Therefore it was concluded that the item release and replacement policy had produced comparable survey designs. On that basis, it was agreed that time series comparisons could be made and the preferred approach was to base the comparisons on all booklets, as reporting on the subset of consistent booklets would produce a second set of headline figures that may be confusing. Therefore the approach has been taken to report on all booklets for the 2014 report and supplementary tables.

To test for significant differences between 2012 and 2014, a series of T-tests were performed in Excel.

[redacted]
14/04/2015

From: [redacted]
Education Analytical Services
20th May 2016

Deputy First Minister and Cabinet Secretary for Education and Skills

Statistics Publication Notice: Scottish Survey of Literacy and Numeracy 2015 (Numeracy)

[redacted 10 paragraphs – out of scope of request]

[redacted]
Education Analytical Services
Ext: [redacted]

Copy List:	For Action	For Comment	For Information		
			Portfolio Interest	Constit Interest	General Awareness
First Minister					X
Leslie Evans	Permanent Secretary				
Paul Johnston	Director General Learning and Justice				
Liz Lloyd	Senior Special Advisor				
Colin McAllister	Special Advisor				
[redacted]	First Minister's Policy				
[redacted]	Communications Smarter				
[redacted]	Communications Smarter				
Fiona Robertson	Director of Learning				
Alan Johnston	Deputy Director Curriculum, Qualifications and Gaelic				
[redacted]	Head of Strategy and Performance Unit				
[redacted]	Performance and Strategy Unit				
[redacted]	Performance and Strategy Unit				
[redacted]	Head of Curriculum Unit				
[redacted]	Curriculum Unit				
[redacted]	Curriculum Unit				
[redacted]	Curriculum Unit				
[redacted]	Curriculum, Qualifications and Gaelic Division				
Donna Bell	Deputy Director Strategy and Performance				

[redacted]	Head of Raising Attainment Unit
[redacted]	Raising Attainment Unit
[redacted]	Raising Attainment Unit
Clare Hicks	Deputy Director People and Infrastructure
[redacted]	Head of People and Leadership Unit
[redacted]	People and Leadership Unit
Bill Maxwell	Chief Executive Education Scotland
[redacted]	Education Scotland
[redacted]	Education Scotland
[redacted]	Education Scotland
Graeme Logan	Education Scotland
[redacted]	Education Scotland
[redacted]	Education Scotland
[redacted]	Education Scotland
[redacted]	Chief Statistician
Audrey MacDougall	Deputy Director, Education Analytical Services
	Head of Learning Analysis, EAS
[redacted]	Learning Analysis
[redacted]	Learning Analysis
[redacted]	Learning Analysis
[redacted]	

Annex A – [redacted – out of scope]

Annex B – [redacted – out of scope of request]

Annex C - SSLN methodology

1. [redacted – one paragraph – out of scope of request]

2. The SSLN sample size is two P4, two P7 and 12 S2 pupils from every mainstream publicly funded and independent school in Scotland. Response rates are high, with 97% of publicly funded schools (33% of independent schools) taking part. This produces a target sample size of 4,000 pupils per stage (achieved sample size around 3,500 pupils per stage). Pupil results are weighted to produce robust results representative of the whole P4, P7 and S2 pupil population.

3. [redacted – 10 paragraphs – out of scope of request]

Annex D – [redacted – out of scope of request]

Any analysis or advice (internal or external) provided to the Scottish Government which compares the SNSA and the SSLN

The Scottish Government does not have any analysis or advice, internal or external, which compares the Scottish National Standardised Assessments and the Scottish Survey of Literacy and Numeracy, other than the bullet points under "key features of this design" (page 5) in document 3 above.

Any communications between the Office of the Researcher and other individuals/departments of the Scottish government regarding either the SNSA or the SSLN

The Scottish Government does not have any information on any communications between the Office of the Chief Researcher and any other individuals or departments of the Scottish Government regarding either the Scottish National Standardised Assessments or the Scottish Survey of Literacy and Numeracy, other than what has already been included in the documents above.