NMSSA • CYCLE 2

NMSSA Report 21 : TECHNICAL INFORMATION 2018 – Mathematics & Statistics • Social Studies

Wānangatia te Putanga Tauira

National Monitoring Study of Student Achievement

Technical Information 2018

Mathematics and Statistics • Social Studies

NMSSA
Wānangatia te Putanga Tauira
National Monitoring Study of Student Achievement

CYCLE 2
NMSSA Report 21

Wānangatia te Putanga Tauira

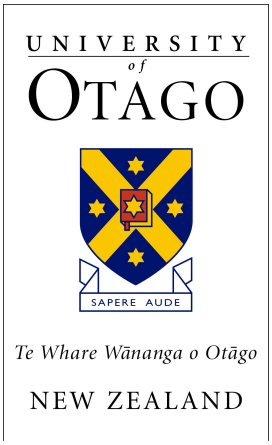
National Monitoring Study of Student Achievement

Technical Information 2018

Mathematics and Statistics • Social Studies

Educational Assessment Research Unit
and
New Zealand Council for Educational Research

NMSSA Report 21



Technical Information 2018 Mathematics and Statistics • Social Studies
(all available online at http://nmssa.otago.ac.nz/reports/index.htm)

National Monitoring Study of Student Achievement Report 21: Technical Information 2018 – Mathematics and Statistics, Social Studies
published by Educational Assessment Research Unit, University of Otago, and New Zealand Council for Educational Research
under contract to the Ministry of Education, New Zealand

ISSN: 2350-3238 (Online) ISBN: 978-1-927286-52-4 (Online only)

National Monitoring Study of Student Achievement
Educational Assessment Research Unit, University of Otago, PO Box 56, Dunedin 9054, New Zealand
Tel: 64 3 479 8561 • Email: nmssa@otago.ac.nz

Contents

2018 Project Team	EARU	NZCER
Management Team	Sharon Young Albert Liau Lynette Jones Jane White	Charles Darr
Design/Statistics/ Psychometrics/Reporting	Albert Liau Alison Gilmore Mustafa Asil	Charles Darr Hilary Ferral Jess Mazengarb
Curriculum/Assessment	Sharon Young Jane White Doris Lancaster	Jonathan Fisher Linda Bonne Teresa Maquire
Programme Support	Lynette Jones Linda Jenkins James Rae Fiona Rae Lee Baker	Jess Mazengarb

External Advisors: Jeffrey Smith – University of Otago, Marama Pohatu – Te Rangatahi Ltd

Cover photos: Ruby Jones • NMSSA Project image, this page: Marelda O'Rourke Gallaher

Acknowledgements

The NMSSA project team wishes to acknowledge the very important and valuable support and contributions of many people to this project, including:

- members of the technical advisory panel
- members of the curriculum advisory panels in mathematics and statistics, and social studies
- principals, teachers and students of the schools where the tasks were piloted and trials were conducted
- principals, teachers and Board of Trustees' members of the schools that participated in the 2018 main study including the linking study
- the students who participated in the assessments and their parents, whānau and caregivers
- the teachers who administered the assessments to the students
- the teachers and senior initial teacher education students who undertook the marking
- the Ministry of Education Research Team and Steering Committee.

Appendix 1:
Sample Characteristics for 2018

Contents:

Tables:

Samples for 2018

A two-stage sampling design was used to select nationally representative samples of students at Year 4 and at Year 8. The first stage involved sampling schools; the second stage involved sampling students within schools.

A stratified random sampling approach was taken to select 100 schools at Year 4 and 100 schools at Year 8. A maximum of 25 students were randomly selected from each school to form national samples at Year 4 and Year 8.

The Ministry of Education July 2017 school returns for Year 3 and Year 7 were used to inform the selection of Year 4 and Year 8 schools in 2018.

1. Sampling of schools

Sampling algorithm

From the complete list of NZ schools select two datasets – one for Year 3 students and one for Year 7 students.

For the Year 3 sample:

- Exclude:
    - schools which have fewer than eight Year 3 students
    - private schools
    - special schools
    - Correspondence School
    - Kura Kaupapa Māori
    - trial schools
    - Chatham Island schools.
- Stratify the sampling frame by region and quintile[1].
- Within each region-by-quintile stratum, order the schools by Year 3 roll size[2].
- Arrange the strata alternately in increasing and decreasing order of roll size[3].
- Select a random starting point.
- From the random starting point, cumulate the Year 3 roll.
- Because 100 schools are required in the sample, the sampling interval is calculated as:

$$\frac{\text{Total number of Year 3 students}}{100}$$

- Assign each school to a 'selection group' using this calculation:

$$Selection\ group\ =\ ceiling\left(\frac{cumulative\ roll}{sampling\ interval}\right)$$

- Select the first school in each selection group to form the final sample.

Follow the same process for the Year 7 sample.

If a school is selected in both the Year 3 and Year 7 samples, randomly assign it to one of the two samples. Locate the school in the unassigned sample and select a replacement school (next on list). Repeat the process for each school selected in both samples.

[1] Decile 1 and 2 comprises Quintile 1; Decile 3 and 4 comprises Quintile 2; Decile 5 and 6 comprises Quintile 3; Decile 7 and 8 comprises Quintile 4; and Decile 9 and 10 comprises Quintile 5.

[2] Roll size refers to the year level in question e.g. roll size for Year 3 students.

[3] This is done so that when replacements are made across stratum boundaries the replacement school is of a similar size to the one it is replacing.

Substitution procedure

The sampling frames constituted 1,497 schools for Year 3 and 961 schools for Year 7 after exclusions had been applied. No schools were listed in both samples.

Selected schools were invited to participate in 2018. Therefore 'Year 3 schools' became 'Year 4 schools' and similarly 'Year 7 schools' became 'Year 8 schools'. Those that declined to participate were substituted using the following procedure.

- From the school sampling frame, select the school one row below the school withdrawn.
- If this school is not available, re-select by going to one row above the school withdrawn.
- If this school is not available, select the school two rows below the school withdrawn. Continue in this sequence until a substitute is found.

In total, 43 schools at Year 4 and 58 schools at Year 8 declined to participate, before a sample of 100 schools at each of Years 4 and 8 was achieved. Of the 43 Year 4 schools, 29 were from the original sample and 14 were replacement schools who also withdrew. Of the 58 Year 8 schools, 39 were from the original sample and 19 were replacement schools who also withdrew.

2. Sampling of students

Four nested student samples were required for the assessment programme:

- A group-administered task (GAT) sample for mathematics that included up to 25 students per school who completed the assessment in mathematics and questionnaires in mathematics and social studies.
- A subset of up to 12 students per school formed the group-administered task (GAT) sample for social studies. These students completed the social studies computer-based assessments.
- A subset of up to eight students formed the sample that participated in the in-depth (InD) assessment in social studies.
- A subset of up to six students formed the sample that participated in the InD assessment in mathematics.

The procedure for selecting students for the samples was as follows.

- Participating schools were asked to provide a list of all students in their school at the relevant year level (Year 4 or Year 8) in 2018, identifying any students who should be excluded for logistical reasons, or because the experience would be inappropriate (e.g. high special needs (ORS), very limited English language (ESOL), Māori Immersion Level 1, would be absent during the visit, had left the school, other health or behavioural issues).
- For each school, students identified for exclusion from the sampling frame were removed from the list, and a computer-generated random number between 1 and 1 million was assigned to each of the remaining students. They were then ranked in order of their random number from lowest to highest.
- The first 25 students in the ordered list were identified as belonging to the GAT mathematics sample. The first 12 students also belonged to the GAT social studies sample, the first 8 students to the sample that participated in the in-depth assessment for social studies, and the first 6 students to the sample that participated in the in-depth assessment for mathematics.
- The names of selected students were returned to schools for approval. Principals or contact people were given a second opportunity to identify students for whom the NMSSA assessment would be inappropriate. Any students identified for withdrawal were replaced with students ranked 26 onwards from the ordered list. The resultant sample was confirmed and letters of consent were sent to the parents of selected students on our behalf via the schools.
- The children of parents who declined to have their child participate were withdrawn from the sample and were replaced in the same way as above (if there were sufficient eligible students).

However, no replacements were added within two weeks of the date of the school visit, as there was insufficient time to seek parental permission.

- On-site replacements of students by teacher assessors (TAs) were made if any of the students 1–8 (the InD sample) were absent or withdrawn on the first day, prior to the start of assessments. They were replaced by students ranked 9–25, on a best-match basis (e.g. using our gender/ethnicity replacement priorities).
- If students were absent or withdrawn after the start of the assessment programme, no replacements were made.

The following sections describe the achieved GAT and InD samples of students at Year 4 and Year 8 and contrast their demographic characteristics with those of their respective national populations (through comparison with the sample frame of eligible schools). This allows us to assess the national representativeness of the samples in relation to those characteristics.

Achieved samples at Year 4

Across the 100 schools participating at Year 4, principals identified 325 students for whom the experience would be unsuitable. These students were not considered for inclusion in the school sample.

The initial sample (the first 25 students in each school's ordered list) consisted of 2,390 randomly selected students. Principals or parents withdrew 193 students after the sample was drawn. Substitute (replacement) students numbered 155. A further 194 students were withdrawn without sufficient time for replacement, were absent or did not respond for other reasons during the assessment period. The achieved GAT mathematics sample included 2,158 students.

Table A1.1 The selection of Year 4 students for the GAT and InD samples from 100 schools

	Group administered tasks		In-depth tasks	
Learning area	**Maths**	**Social Studies**	**Maths**	**Social Studies**
Initial sample numbers	**2390**	**1198**	**600**	**800**
Students withdrawn by parents or principals after sampling	-193	-4	-	-1
Substitute students used (replacements for above)	155	-	-	-
Late withdrawals made prior to commencement of assessment programme	-35	-	-	-
Absences, non-responses during assessment period	-159	-	-	-
Achieved sample numbers	**2158**	**1194**	**600**	**799**

Table A1.2 contrasts the characteristics of the samples with the sample frame across a number of key demographic variables.

Table A1.2 The composition of the Year 4 samples in comparison with the sample frame by gender, ethnicity, school quintile, school type and education region

	Population *N=61,844* (%)	Samples			
		GAT		InD	
		Maths	Social Studies	Maths	Social Studies
Gender					
Boys	**51**	51	52	50	51
Girls	**49**	49	48	50	49
Ethnicity					
European	**52**	55	56	56	57
Māori	**21**	19	19	20	20
Pacific	**12**	12	11	11	11
Asian	**12**	11	10	9	10
Other	**3**	4	3	3	3
School quintile					
1	**17**	15	16	16	16
2	**17**	18	19	19	19
3	**17**	16	16	16	16
4	**22**	20	21	21	21
5	**28**	30	28	28	28
School type					
Contributing (Years 1-6)	**61**	64	62	62	62
Full Primary (Years 1-8)	**37**	33	35	35	35
Composite (Years 1-15)	**3**	3	3	3	3
MOE region					
Auckland	**35**	36	34	34	34
Bay of Plenty/Waiariki	**8**	8	8	8	8
Canterbury	**12**	11	11	11	11
Hawkes Bay/Tairāwhiti	**5**	5	5	5	5
Nelson/Marlborough/West Coast	**4**	4	4	4	4
Otago/Southland	**6**	7	7	7	7
Northland/Tai Tokerau	**4**	4	4	4	4
Taranaki/Whanganui/Manawatu	**7**	6	6	6	6
Waikato	**9**	8	9	9	9
Wellington	**12**	12	12	12	12

Note: Ministry of Education July 2018 school returns for Year 4

Achieved samples at Year 8

Principals in schools selected at Year 8 identified 482 students for whom the NMSSA assessment experience would be unsuitable. The initial sample at Year 8 consisted of 2,335 randomly selected students. Principals or parents withdrew 195 students after the sample was drawn. Substitute (replacement) students numbered 157. A further 214 students were withdrawn late, were absent or did not respond for other reasons during the assessment period. The achieved GAT mathematics sample was 2,083 students at Year 8.

Table A1.3 The selection of Year 8 students for the GAT and InD samples from 100 schools

	Group administered tasks		In-depth tasks	
Learning area	**Maths**	**Social Studies**	**Maths**	**Social Studies**
Initial sample numbers	**2335**	**1191**	**600**	**799**
Students withdrawn by parents or principal after sampling	-195	-4	-2	-6
Substitute students used (replacements for above)	157	-	-	-
Late withdrawals	-28	-	-	-
Absences/non-responses during assessment period	-186	-	-	-
Achieved sample	**2083**	**1187**	**598**	**793**

Table A1.4 contrasts the characteristics of the samples with the sample frame across a number of key demographic variables.

Table A1.4 The composition of the Year 8 samples in comparison with the sample frame by gender, ethnicity, school quintile, school type and education region

	Population *N=55,003 (%)*	Samples			
		GAT		InD	
		Maths	Social Studies	Maths	Social Studies
Gender					
Boys	51	52	52	52	52
Girls	49	48	48	48	48
Ethnicity					
European	54	53	53	54	54
Māori	21	19	21	21	21
Pacific	11	11	11	11	11
Asian	11	11	9	9	9
Other	3	5	6	6	5
School quintile					
1	14	12	14	14	14
2	16	16	16	16	16
3	22	22	22	22	22
4	24	26	25	25	25
5	24	24	23	23	23
School type					
Full primary (Year 1-8)	31	38	40	40	40
Intermediate	48	44	42	42	42
Secondary (Year 7-15)	16	14	13	13	13
Composite (Year 1-15)	5	5	5	5	5
MOE region					
Auckland	33	36	35	35	35
Bay of Plenty/Waiariki	8	8	7	7	7
Canterbury	12	12	12	12	12
Hawkes Bay/Tairāwhiti	5	5	5	5	5
Nelson/Marlborough/West Coast	4	4	4	4	4
Otago/Southland	7	7	7	7	7
Northland/Tai Tokerau	4	3	4	4	4
Taranaki/Whanganui/Manawatu	7	7	7	7	7
Waikato	9	8	8	8	8
Wellington	11	11	12	11	11

Note: Ministry of Education July 2018 school returns for Year 8

At both year levels the student samples closely matched the characteristics of the population (as represented by the sample frame) in relation to the identified demographic variables. We have confidence in their national representativeness.

Appendix 2:
Methodology for the 2018 NMSSA Programme

Contents:

Tables:

This appendix outlines the methodology for the 2018 social studies and mathematics and statistics (hereafter mathematics) studies undertaken by the National Monitoring Study of Student Achievement (NMSSA).

1. The 2018 social studies assessment programme

The 2018 social studies assessment programme built upon the assessment framework and associated assessment programme developed for the 2014 social studies study. In 2018, we used a variety of assessment approaches to assess the Nature of Social Studies (NSS). The first approach involved a group-administered task (GAT) delivered via laptop to 1200 students at Year 4 and 1200 students at Year 8. The second approach included a set of in-depth group and individual tasks undertaken by a subset of 800 students. The group tasks encouraged group discussion and participation. The third approach consisted of one-to-one interview tasks conducted with a subset of 600 students at each year level. In-depth assessment, including interviews, provided an opportunity to explore students' knowledge and understanding of social studies without the limitations inherent in requiring students to communicate responses in written form. The NSS consisted of 17 tasks, six of which were link tasks from the 2014 study. Table A2.1 summarises the key differences between the assessment programmes in 2014 and 2018. See Appendix 7 for the 2018 assessment framework.

Table A2.1 The key features of the 2014 and 2018 social studies assessment programmes

	2014	2018
Assessment approaches	The Nature of Social Studies (NSS) assessment was made up of in-depth tasks* using interviews and individual or group activities. The tasks covered the strands of Conceptual Understanding, Active Participation in Society, and Values and Perspectives. Responses from the NSS tasks were used to create an IRT measurement scale. All assessments were videoed.	The NSS scale was made up a diverse range of assessment approaches. The assessment combined a GAT administered on laptops and in-depth tasks (interviews, and group and individual activities). NSS was expanded to include an additional strand: Information gathering and analysis. All in-depth tasks were videoed.
Number of students	Eight students per school participated in the in-depth tasks, giving a total of 800 students at Year 4 and 800 students at Year 8.	Up to 12 students per school participated in the GAT. Eight students per school participated in the group and individual activities, and six students per school participated in interviews.

NB *A task is an assessment context. Each task has several questions.

Development and trialling of social studies tasks

The NMSSA team reviewed all 2014 social studies tasks for possible inclusion in the 2018 assessment programme. Some tasks were retained in their original format to be used as link tasks, necessary for making comparisons between 2014 and 2018. Tasks were based on the focus of the social studies learning area, which is defined as being, 'about how societies work and how people can participate as critical, active, informed, and responsible citizens' (NZC[4], p. 30).

New and modified tasks were piloted in local schools before being used in a NMSSA trial in March 2018 involving schools in Auckland and Otago/Southland. The student responses from the pilots and the trial were used to refine the tasks and support the development of appropriate scoring guides. An Item Response Theory (IRT) model[5] was applied to the trial data to help refine the tasks, inform the selection of tasks for the main study and explore the development of the reporting scale: Nature of Social Studies (NSS) that paralleled and extended the 2014 scale.

[4] Ministry of Education (2007). *The New Zealand Curriculum*. Wellington: Learning Media Limited.

[5] IRT is an approach to constructing and scoring assessments and surveys that measure mental competencies and attitudes. IRT seeks to establish a mathematical model to describe the relationship between people (in terms of their levels of ability or the strengths of their attitude) and the probability of observing a correct answer or a particular level of response to individual questions. IRT approaches provide flexible techniques for linking assessments made up of different questions to a common reporting scale. The common scale allows the performance of students to be compared regardless of which form of the assessment they were administered.

Administration of the assessment tasks

Eleven teacher assessors were trained in the administration of tasks during a five-day training programme prior to the main study. Teacher assessors were carefully monitored and received feedback to ensure consistency of administration. Student responses were captured on video and paper and stored electronically for marking.

2. 2018 mathematics and statistics assessment programme

The 2018 mathematics assessment programme built upon the mathematics programme used in 2013. The programme retained many of the tasks used in 2013 and included a range of new tasks. Table A2.2 compares the assessment programmes for 2013 and 2018. The biggest change was the construction of one rather than two scales.

Table A2.2 The key features of the 2013 and 2018 mathematics assessment programmes

	2013	2018
Assessment approaches	Two separate assessments: • a 45-minute, group-administered task (GAT). This was a paper-and-pencil assessment involving selected response and short answer questions called the Knowledge and Application of Mathematics Ideas (KAMSI) assessment • a selection of individual one-to-one interview tasks and individual and group performance activities called the Mathematical and Statistical Proficiencies (MSP) assessment. A separate scale was constructed for each assessment: the KAMSI scale and the MSP scale.	One assessment made up of two parts: • a paper-and-pencil, 45-minute group-administered task (GAT) component involving selected response and short answer questions • a selection of 'in-depth' tasks involving student interviews, independent 'station', and group activities. Responses from both components were used to construct one scale: the Mathematics and Statistics (MS) scale.
Number of students	Up to 25 students per school participated in the paper-and-pencil assessment. Eight of these students per school participated in the in-depth tasks.	Up to 25 students per school participated in the paper-and-pencil assessment. Up to six of these students per school participated in the in-depth tasks.

Development of the group-administered part of the MS assessment

The group-administered (GAT) part of the MS assessment was based on the questions developed for the group-administered assessment used in the 2013 study. Assessment development staff within the NMSSA project reviewed the existing items in order to identify areas where new items could be added to support the assessment framework and broaden the pool of questions. They then wrote a collection of new questions to cover these areas. All new questions were carefully reviewed before being piloted in a range of schools in the Wellington area. The results from the piloting were used to select and fine-tune questions for a larger national trial.

The national item trial was held in March of 2018. The trial involved about 400 students at each of Year 4 and Year 8, and enabled the development team to refine the new items as needed and then select a final bank of questions for use in the main study.

A total of 14 GAT forms were constructed for the 2018 study, based on the final pool of questions (eight forms at Year 8 and six at Year 4). Each form was linked to the other forms using common questions.

Development of the in-depth tasks for mathematics

A selection of in-depth tasks was also developed as part of the MS assessment. These were designed to be more open-ended than the GAT and to stimulate extended responses from students.

Development began with a review of the in-depth tasks used in 2013. Some of these tasks were adapted for use in 2018. A selection of new tasks were also developed. Most of the tasks were designed to be administered as part of a one-to-one interview with a teacher assessor, while some were designed to be completed independently as part of a group of 'stations' activities. There was also one task designed to be

completed by a pair of students in each school, that explored their ability to problem-solve co-operatively. Many of the in-depth tasks required students to use equipment or consider a rich stimulus.

An initial group of in-depth tasks were piloted in local schools in Wellington and Auckland in late 2017 and early 2018. Some of these were then used in a larger item trial held in March 2018 that involved a selection of schools in Auckland and Otago/Southland. Data from the pilots and trials were used to refine the tasks and their associated scoring rubrics. As a result of the development process, eleven in-depth tasks were selected for use in the main 2018 study.

Use of the MS assessment in the 2018 NMSSA study

Teacher assessors were instructed on how to administer the MS assessment during a five-day training session prior to the main study.

The group-administered part of the MS assessment was administered to up to 25 students in each school. Up to six students in each school completed the in-depth tasks.

Linking Year 4 and Year 8 results in mathematics

To enable achievement to be linked across Year 4 and Year 8, four additional GAT forms were constructed using a mix of questions from both year levels. These were administered to a sample of about 800 Year 6 students from schools across the country. The Year 6 schools used were additional schools not already involved in the NMSSA study.

3. Marking social studies and mathematics

Teacher markers, some of whom had been teacher assessors, and third-year University of Otago College of Education students were employed to mark the tasks. All markers were trained, and quality assurance procedures were used to ensure consistency of marking. The marking schedules were refined as necessary to ensure they reflected the range of responses found in the main study. Students' scores were entered directly by the markers into the electronic database.

Inter-rater reliability was calculated using Fleiss Kappa for 203 items in social studies and mathematics. Inter-rater reliability was 'perfect' (greater than 0.81) for 59 percent of the questions, 'substantial' (between 0.61 and 0.80) for 17 percent of the questions, and 'moderate' (between 0.41and 0.60) for 24 percent of the questions[6] (Cicchetti, 1994)[7].

4. Creating the achievement scales for social studies and mathematics

The Rasch IRT model was applied to student responses from the items in the NSS and MS assessments. This approach included analysing the items used in the assessments for any differential item functioning with respect to year level, gender and ethnicity.

The IRT approach allowed a set of plausible values to be generated for each student involved in the study. Plausible values account for the imprecision associated with scores in an assessment, which can produce biased estimates of how much achievement varies across a population. Each set of plausible values represents the range of achievement levels a student might reasonably be expected to attain given their responses to the assessment items. Plausible values provide more accurate estimates of population and subgroup statistics, especially when the number of items answered by each student is relatively small.

6 Six items were re-marked to improve their inter-marker reliability.

7 Cicchetti, D. V. (1994). Guidelines, criteria, and rules of thumb for evaluating normed and standardized assessment instruments in psychology. *Psychological Assessment*, *6*(4), 284.

Standardising the scales

For ease of understanding, each scale was standardised so that:

- the mean of Year 4 and Year 8 students combined was equal to 100 scale score units
- the average standard deviation for the two year-levels was equal to 20 scale score units.

Achievement on the scales ranged from about 20 to 180 units.

The scales locate both student achievement and relative task difficulty on the same measurement continuums using scale scores.

Scale descriptions

The scales for NSS and MS were described to indicate the range of knowledge and skills assessed.

To create the scale descriptions, the scoring categories for each item (e.g. 0, 1 or 2) in the NSS and MS assessments were located on the respective scales. This meant identifying where the students who scored in each category were most likely to have achieved overall on the scale. Once this had been done for all items, the NMSSA team identified the competencies exhibited as the scale locations associated with the different scoring categories increased, and students' responses became more sophisticated. The result was a multi-part description for each scale, providing a broad indication of what students typically know and can do when achieving at different places on the scale.

The descriptions were provided to give readers of NMSSA reports a strong sense of how social studies and mathematics were assessed. The scale descriptors were not written to necessarily 'line up' with curriculum levels or achievement objectives. They were a direct reflection of what was assessed and how relatively hard or easy students found the content of the assessment.

5. Linking results from cycle 1 to cycle 2

In order to compare results from cycle 1 with those from 2018, separate scale-linking exercises were carried out for social studies and mathematics. The exercises involved comparing the scale locations of the common questions used in the assessments at the different points of time. As part of the exercises, the cycle 1 scales were reconstructed using the same plausible values approach that was used in 2018 (plausible values were not used when social studies and mathematics were assessed in cycle 1). The linking exercises indicated that simple transformations could be used to link the scales. These transformations were applied, allowing results from both cycles to be compared. Further information about the linking processes can be found in Appendix 5 (social studies) and Appendix 6 (mathematics).

6. Reporting achievement against curriculum levels

For mathematics, a curriculum alignment exercise in 2013 was used to determine achievement expectations (cut scores) on the 2013 mathematics scale associated with achievement at different curriculum levels. Linking the 2013 scale to the 2018 MS scale allowed these cut scores to be located on the MS scale. A similar curriculum alignment for social studies was carried out in 2014. This, along with scale linking for social studies, allowed achievement on the 2018 NSS scale to be reported against curriculum levels.

7. Development of questionnaires for examining contextual information

In order to gain a better understanding of student achievement in New Zealand, NMSSA collects contextual information through questionnaires to students, teachers and principals. A conceptual framework for describing the contextual information to be collected by NMSSA during cycle 2 sought to:

- build (and improve) on the contextual information collected in the first cycle
- learn from the literature about important factors that influence achievement and consider them for inclusion in NMSSA
- address the thematic contextual questions set out in the respective assessment plans.[8]

One new development in cycle 2 was the creation of measurement scales to report on different aspects of the contextual information.

For the student questionnaire, items were developed to construct the following scales:

- Attitude to Social Studies
- Attitude to Mathematics
- Confidence in Social Studies
- Confidence in Mathematics.

For the teacher questionnaire, items were developed to construct the following scales:

- Attitude to Teaching Social Studies
- Attitude to Teaching Mathematics
- Confidence in Teaching Social Studies
- Confidence in Teaching Mathematics.

The scales were constructed using the Rasch model. This approach included analysing the items used in the questionnaires for any differential item functioning with respect to year level, gender and ethnicity. Unlike the achievement measures, plausible values were not generated for the contextual scales. Each contextual scale was standardised in the same way as the achievement scales.

To aid interpretation of the contextual scales, they were divided into separate score ranges to provide different reporting categories. For instance, the Attitude to Social Studies scale was broken down into three score ranges. The 'very positive' part of the scale was associated with students mainly using the 'totally agree' category to respond to each of the questionnaire statements related to attitude, the 'positive' section of the scale was associated with students mainly using either 'agree a lot' or 'agree a little', and the 'not positive' part of the scale was associated with students mainly using 'do not agree at all'.

8. Administration of the questionnaires

Students who participated in the mathematics GAT and the social studies GAT were all expected to respond to the associated student questionnaire items. Up to three teachers from each school were invited to complete the teacher questionnaire. These were classroom teachers, social studies specialist teachers and mathematics specialist teachers. The principal or a designated school leader (if principal unavailable) from each school completed the principal questionnaire.

[8] Gilmore, A. (2016). Towards a NMSSA conceptual framework. NMSSA Working Paper.

Appendix 3:
NMSSA Sample Weights 2018

Contents:

Tables:

The methodology for calculating sample weights on an annual basis is detailed in *NMSSA Approach to Sample Weighting,* available online at https://nmssa.otago.ac.nz/reports/Sample_Weighting_NMSSA.pdf.

Each year we set out a brief summary of the effect of applying sample weights in the analysis of the current year's data, and make a recommendation as to whether weights should be used.

Tables of estimated[9] means and standard errors calculated with and without sample weights follow. In 2018, NMSSA measured achievement in social studies and mathematics. Information about the respective samples can be found in Appendix 1.

Tables 1 and 2 report the NMSSA estimated means and standard errors (in scale score units) for the Year 4 and Year 8 social studies samples, respectively; Tables 3 and 4 report these for the mathematics samples.

Summary

All weighted estimates were well within one standard error of the estimated unweighted mean.

The recommendation was to proceed with the 2018 analyses without using sample weights.

[9] All estimates of means and standard errors are calculated using the full sample size rather than the *effective sample size* defined by the design effect calculations. See Appendix 4.

Social Studies

Table A3.1 NMSSA social studies achievement Year 4: Comparison of estimates using unweighted and weighted data

	Year 4					
	Using unweighted data		Using weighted data			
	Mean	SE	Mean	SE	Difference	N
All	80.3	0.6	80.2	0.6	0.1	1195
Girls	82.7	0.8	82.6	0.8	0.1	574
Boys	78.0	0.8	78.0	0.8	0.0	621
NZE	84.7	0.7	84.8	0.7	-0.1	772
NZE girls	87.0	0.9	87.2	0.9	-0.2	372
NZE boys	82.5	0.9	82.6	0.9	-0.1	400
Māori	72.5	1.3	72.3	1.3	0.2	268
Māori girls	73.8	1.8	73.6	1.8	0.2	137
Māori boys	71.0	1.7	70.9	1.7	0.1	131
Pacific	69.3	1.6	69.2	1.6	0.1	154
Pacific girls	72.8	2.2	72.9	2.2	-0.1	74
Pacific boys	66.0	2.3	65.7	2.3	0.3	80
Asian	83.3	1.6	83.3	1.6	0.0	143
Asian girls	86.5	2.3	86.5	2.3	0.0	71
Asian boys	80.1	2.1	80.1	2.1	0.0	72
Quintile 1	67.0	1.4	66.9	1.4	0.1	195
Quintile 2	74.8	1.2	74.7	1.2	0.1	224
Quintile 3	78.2	1.5	78.2	1.5	0.0	192
Quintile 4	85.5	1.2	85.5	1.2	0.0	249
Quintile 5	88.9	1.0	88.9	1.0	0.0	335

Table A3.2 NMSSA social studies achievement Year 8: Comparison of estimates using unweighted and weighted data

	Year 8					
	Using unweighted data		Using weighted data			
	Mean	SE	Mean	SE	Difference	N
All	119.7	0.6	119.9	0.6	-0.2	1182
Girls	123.7	0.8	123.7	0.8	0.0	569
Boys	116.1	0.8	116.1	0.8	0.0	613
NZE	123.1	0.7	123.2	0.7	-0.1	725
NZE girls	126.4	0.9	126.4	0.9	0.0	361
NZE boys	119.9	0.9	119.9	0.9	0.0	364
Māori	111.5	1.1	111.6	1.1	-0.1	288
Māori girls	118.2	1.6	118.2	1.6	0.0	116
Māori boys	106.9	1.4	106.9	1.4	0.0	172
Pacific	109.8	1.7	109.8	1.7	0.0	148
Pacific girls	111.2	2.8	111.2	2.8	0.0	64
Pacific boys	108.7	2.1	108.7	2.1	0.0	84
Asian	124.1	1.7	124.1	1.7	0.0	129
Asian girls	126.7	2.3	126.7	2.3	0.0	74
Asian boys	120.6	2.5	120.6	2.5	0.0	55
Quintile 1	106.8	1.5	106.9	1.5	-0.1	165
Quintile 2	113.8	1.3	113.9	1.3	-0.1	189
Quintile 3	117.9	1.2	118.1	1.2	-0.2	258
Quintile 4	124.5	1.0	124.6	1.0	-0.1	295
Quintile 5	128.2	1.1	128.3	1.1	-0.1	275

Mathematics

Table A3.3 NMSSA mathematics achievement Year 4: Comparison of estimates using unweighted and weighted samples data

	Year 4					
	Using unweighted data		Using weighted data			
	Mean	SE	Mean	SE	Difference	N
All	83.9	0.4	83.5	0.4	0.4	2105
Girls	82.5	0.6	82.1	0.6	0.4	1035
Boys	85.2	0.6	84.8	0.6	0.4	1070
NZE	87.5	0.5	87.4	0.5	0.1	1328
NZE girls	85.9	0.7	85.8	0.7	0.1	655
NZE boys	89.0	0.7	88.9	0.7	0.1	673
Māori	75.2	0.9	74.8	0.9	0.4	439
Māori girls	73.0	1.3	72.6	1.3	0.4	217
Māori boys	77.3	1.2	77.0	1.2	0.3	222
Pacific	71.1	1.1	70.8	1.1	0.3	287
Pacific girls	71.0	1.6	70.4	1.6	0.6	142
Pacific boys	71.3	1.5	71.1	1.5	0.2	145
Asian	91.4	1.1	91.2	1.1	0.2	263
Asian girls	91.8	1.5	91.6	1.5	0.2	120
Asian boys	91.0	1.6	90.9	1.6	0.1	143
Quintile 1	68.7	1.0	68.6	1.0	0.1	316
Quintile 2	78.1	0.9	78.1	0.9	0.0	386
Quintile 3	83.6	0.9	83.6	0.9	0.0	347
Quintile 4	87.6	0.9	87.7	0.9	-0.1	419
Quintile 5	92.6	0.7	92.6	0.7	0.0	637

Table A3.4 NMSSA mathematics achievement Year 8: Comparison of estimates using unweighted and weighted samples data

Year 8						
	Using unweighted data		Using weighted data			
	Mean	SE	Mean	SE	Difference	N
All	117.1	0.5	116.6	0.5	0.5	1985
Girls	116.1	0.6	115.7	0.6	0.4	949
Boys	118.0	0.7	117.5	0.7	0.5	1036
NZE	119.5	0.6	119.4	0.6	0.1	1222
NZE girls	118.2	0.8	118.2	0.8	0.0	575
NZE boys	120.7	0.8	120.5	0.8	0.2	647
Māori	108.3	0.9	107.9	0.9	0.4	448
Māori girls	107.9	1.3	107.4	1.3	0.5	192
Māori boys	108.6	1.2	108.3	1.2	0.3	256
Pacific	105.6	1.2	105.0	1.2	0.6	245
Pacific girls	104.8	1.7	104.0	1.7	0.8	107
Pacific boys	106.3	1.6	105.8	1.6	0.5	138
Asian	127.0	1.4	126.6	1.4	0.4	256
Asian girls	125.9	1.8	125.5	1.8	0.4	146
Asian boys	128.5	2.1	128.2	2.1	0.3	110
Quintile 1	102.3	1.1	102.3	1.1	0.0	239
Quintile 2	110.4	1.1	110.4	1.1	0.0	309
Quintile 3	115.7	1.0	115.7	1.0	0.0	446
Quintile 4	121.8	0.9	121.8	0.9	0.0	519
Quintile 5	125.0	0.9	124.9	0.9	0.1	472

Appendix 4:
Variance Estimation: NMSSA 2018

Contents:

Tables:

1. Introduction

This brief summary supports the general NMSSA variance estimation paper[10], with specific findings relating to data in NMSSA 2018.

Design effects were calculated using the data collected for the NMSSA 2018 mathematics assessment. The NMSSA mathematics assessment was completed by the entire NMSSA sample, and therefore provides the most complete information regarding the clustering of students in schools, and consequently the effect on variance estimation.

Design effects for the whole sample, and key sub-groups were investigated.

In general, through experience with calculating design effects each year, it has been noted that reducing the sample size by a factor of 0.7 for calculation of population statistics, accounts for most of the design effect related to the clustered nature of the NMSSA sample.

Design effects in 2018 mostly varied between about 1.0 and 2.0. While the design effects in some cases are fairly large, the effect on the width of confidence intervals is small in practice. In each case, the increase in width of the 95% confidence intervals is less than 1 NMSSA scale score point.

It was recommended that for ease of calculation, and to absorb most of the variance bias caused by the NMSSA complex sample design, the standard multiplier of 0.7 should be used to form an effective sample size in the calculation of statistics dependent on sample size.

Tables follow showing the effect of the NMSSA complex sample design for the 2018 mathematics assessment.

[10] A standard routine for assessing design effects in NMSSA was developed using NMSSA data over the years 2014 and 2015. See *Variance Estimation in NMSSA*, at https://nmssa.otago.ac.nz/reports/Variance_Estimation_NMSSA.pdf.

2. Tables of design effects

Table A4.1 Mathematics Year 4: Comparison of results for different variance estimation methods

Year 4	Mean[11] (SRS[12])	SE (SRS)	SE (TSL[13])	CI (SRS) (lower)	CI (SRS) (upper)	CI (TSL) (lower)	CI (TSL) (upper)	Design effect	CI width increase	CI width increase %	N	Effective N
All Year 4	0.51	0.03	0.04	0.45	0.56	0.44	0.58	1.86	0.0195	36	2105	1132
NZE[14]	0.77	0.03	0.04	0.71	0.84	0.69	0.86	1.65	0.0189	29	1158	702
Māori	-0.07	0.06	0.06	-0.18	0.04	-0.19	0.05	1.26	0.0138	12	426	338
Pacific	-0.46	0.09	0.08	-0.63	-0.29	-0.61	-0.31	0.82	-0.0159	-9	184	226
Asian	1.08	0.07	0.08	0.94	1.22	0.92	1.24	1.35	0.0228	16	232	173
Female	0.42	0.04	0.05	0.35	0.50	0.32	0.52	1.81	0.0263	35	1020	564
Male	0.60	0.04	0.05	0.52	0.67	0.49	0.70	1.88	0.0285	37	1047	557
Female NZE	0.67	0.05	0.06	0.58	0.76	0.56	0.78	1.48	0.0202	22	560	379
Female Māori	-0.19	0.08	0.09	-0.34	-0.03	-0.36	-0.01	1.20	0.0152	9	217	182
Female Pacific	-0.51	0.13	0.11	-0.76	-0.26	-0.73	-0.29	0.78	-0.0319	-13	89	121
Female Asian	1.06	0.10	0.12	0.87	1.26	0.83	1.30	1.47	0.0419	21	108	74
Male NZE	0.87	0.05	0.06	0.78	0.96	0.75	0.99	1.70	0.0284	30	598	353
Male Māori	0.05	0.08	0.09	-0.10	0.20	-0.12	0.22	1.31	0.0216	14	209	160
Male Pacific	-0.41	0.11	0.10	-0.64	-0.19	-0.62	-0.21	0.85	-0.0184	-8	95	114
Male Asian	1.09	0.10	0.11	0.90	1.29	0.87	1.32	1.26	0.0242	12	124	100
Low decile	-0.27	0.05	0.06	-0.38	-0.17	-0.40	-0.15	1.30	0.0151	14	483	373
Mid decile	0.47	0.04	0.05	0.39	0.55	0.37	0.57	1.42	0.0157	19	772	545
High decile	0.98	0.04	0.05	0.91	1.06	0.89	1.08	1.60	0.0198	26	850	532

[11] All results in table are quoted in logit units except where indicated.
[12] Simple random sample
[13] Taylor series linearisation method
[14] New Zealand European

Table A4.2 Mathematics Year 8: Comparison of results for different variance estimation methods

Year 8	Mean[15] (SRS[16])	SE (SRS)	SE (TSL[17])	CI (SRS) (lower)	CI (SRS) (upper)	CI (TSL) (lower)	CI (TSL) (upper)	Design effect	CI width increase	CI width increase %	N	Effective N
All Year 8	2.62	0.03	0.04	2.56	2.68	2.54	2.70	1.82	0.0207	35	1985	1089
NZE	2.78	0.04	0.05	2.71	2.86	2.69	2.87	1.56	0.0186	25	1094	703
Māori	2.06	0.06	0.06	1.95	2.17	1.93	2.19	1.37	0.0188	17	448	328
Pacific	1.63	0.09	0.10	1.46	1.81	1.43	1.83	1.30	0.0250	14	141	109
Asian	3.33	0.09	0.11	3.15	3.50	3.11	3.54	1.40	0.0333	18	231	166
Female	2.57	0.04	0.06	2.48	2.65	2.46	2.68	1.81	0.0288	35	912	504
Male	2.68	0.04	0.06	2.60	2.76	2.57	2.79	1.86	0.0308	36	1034	557
Female NZE	2.70	0.05	0.06	2.59	2.80	2.57	2.82	1.55	0.0254	24	504	326
Female Māori	2.04	0.08	0.10	1.87	2.20	1.85	2.23	1.35	0.0270	16	192	142
Female Pacific	1.57	0.13	0.13	1.32	1.82	1.33	1.82	0.99	-0.0031	-1	63	67
Female Asian	3.20	0.12	0.13	2.97	3.43	2.94	3.46	1.34	0.0359	15	141	106
Male NZE	2.86	0.05	0.07	2.75	2.96	2.73	2.99	1.53	0.0251	24	590	387
Male Māori	2.08	0.07	0.09	1.93	2.23	1.91	2.25	1.38	0.0262	18	256	186
Male Pacific	1.68	0.12	0.15	1.44	1.93	1.39	1.97	1.44	0.0490	20	78	55
Male Asian	3.52	0.15	0.18	3.23	3.81	3.18	3.87	1.44	0.0583	20	90	63
Low decile	1.90	0.06	0.07	1.78	2.01	1.76	2.03	1.36	0.0193	17	425	312
Mid decile	2.55	0.05	0.06	2.46	2.64	2.44	2.66	1.48	0.0202	22	739	499
High decile	3.06	0.04	0.06	2.98	3.15	2.96	3.17	1.61	0.0233	27	821	512

[15] All results in table are quoted in logit units except where indicated.
[16] Simple random sample
[17] Taylor series linearisation method

Appendix 5:
Linking Social Studies across Cycle 1 and Cycle 2

Contents:

Figures:

Tables:

1. Introduction

In order to make comparisons across cycles, the National Monitoring Study of Student Achievement (NMSSA) carries out analyses in each learning area to link the assessment results. This document summarises the steps conducted to link the Nature of Social Studies (NSS) assessments in 2014 and 2018.

In 2014, the NSS scale was constructed using in-depth (interview and group/team) items administered to eight students per school. However, in 2018, both group-administered tasks (presented on computer taken by up to 12 students per school) and in-depth items were used to construct the NSS scale. Thus, the 2018 scale was considered to be a richer measure of the same construct. Both scales were psychometrically sound and robust measures. Because of these reasons NMSSA decided to link the 2014 scale to the existing 2018 scale (rather than the other way round).

2. Technical differences 2014 to 2018

Some technical details regarding estimation have changed between 2014 and 2018. Primarily, plausible values have been introduced (since 2015) for calculating population estimates. Generating sets of plausible values for the student sample requires a slightly different estimation technique from that used in 2014 for calculating item parameters. These technical changes necessitated a re-analysis of the assessment data from 2014 so that it could be properly compared with the 2018 data.

The re-analysis of 2014 data has been done solely for the purposes of the NMSSA trend analysis. It means that estimates recorded in the 2014 NMSSA social studies report cannot be directly compared with those in the 2018 report. Meaningful comparisons across time are restricted to those reported in the trend analysis sections of the 2018 reports.

3. Reconstruction of the 2014 NSS scale

The 2014 social studies data was re-analysed with a process that replicated the 2018 analysis as precisely as possible. In 2014 NMSSA used joint maximum likelihood estimation (JMLE) procedures to estimate both item and person parameters. The reconstruction of the data involved using marginal maximum likelihood (MML) to estimate item parameters. Both estimation methods apply the Rasch model. The main difference between the two estimation procedures is that MML assumes an underlying normal distribution for the student population, whereas JMLE does not.

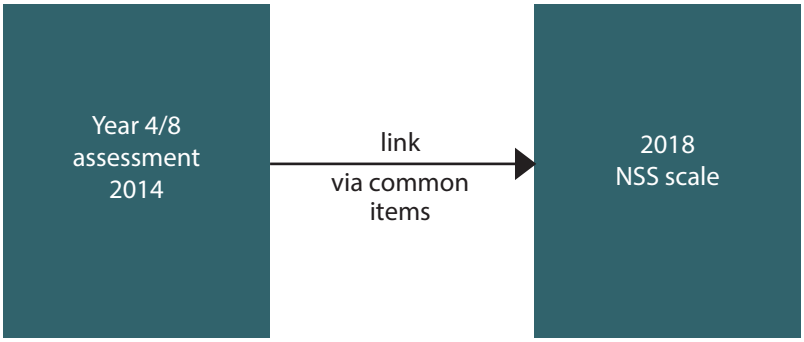
MML item parameters were generated for the 2014 data, and link item calibrations at both time-points were examined.

Linking process

We applied two approaches (see Figure A5.1) to link the 2014 and 2018 NSS scales:

- Design A: linking the (Year 4/8) 2014 scale to the (Year 4/8) 2018 scale
- Design B: separately linking the Year 4 and Year 8 data from 2014 to the 2018 scale.

Linking design A



Linking design B

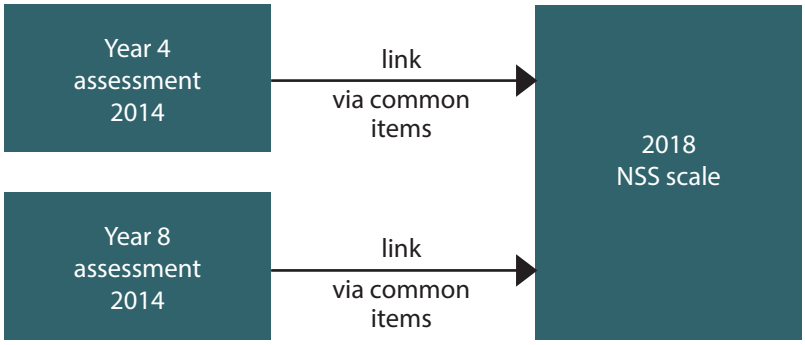


Figure A5.1 Linking schemes for NMSSA Nature of Social Studies

To create a strong link between scales, the two sets of item calibrations at different time points ideally require:

- as many items as possible
- a good spread of items across the scale
- strong correlation between the two sets
- similar standard deviation in the two sets.

Both linking methods yielded very similar trend outcomes.

It was decided to use the linking design A for reporting because, when compared with linking Year 4 and Year 8 separately, whole scale linking resulted in a set of linking items that better represented the construct (fewer items had to be deleted), higher correlations between item calibrations, and a smaller difference between standard deviations.

EAP[18] person estimates were generated for the 2014 data using transformed MML item parameter estimates, and the usual procedure for generating plausible values was carried out.

Linking outcomes

Of the 19 items chosen for linking, two items did not correlate well enough to be included in the link calculation. These items were eliminated from ensuing calculations. The remaining 17 items had a correlation of 0.98, and showed a good spread across the NMSSA NSS scale. The two sets of item parameters also recorded a similar standard deviation at both time points: 1.31 logits and 1.33 logits at 2014 and 2018, respectively.

[18] An expected a posteriori (EAP) estimate refers to the expected value of the posterior probability distribution of latent trait scores in a given case.

The standard deviations were sufficiently similar to warrant a simple shift on the NSS scale to bring the 2014 calibrations in line with the 2018 calibrations.

The transformation which takes the 2014 MML item parameters to the 2018 scale is:

$$\delta_i^{2017} = \delta_i^{2012} - 0.15\ logits,$$ where δ_i is the estimated parameter of *item i*

4. Trend analysis

The Year 4 achievement distributions shows an improvement between 2014 and 2018, while the overall Year 8 mean shows no change. Figure A5.2 depicts theoretical normal distributions based on the observed mean and standard deviation (in logits).

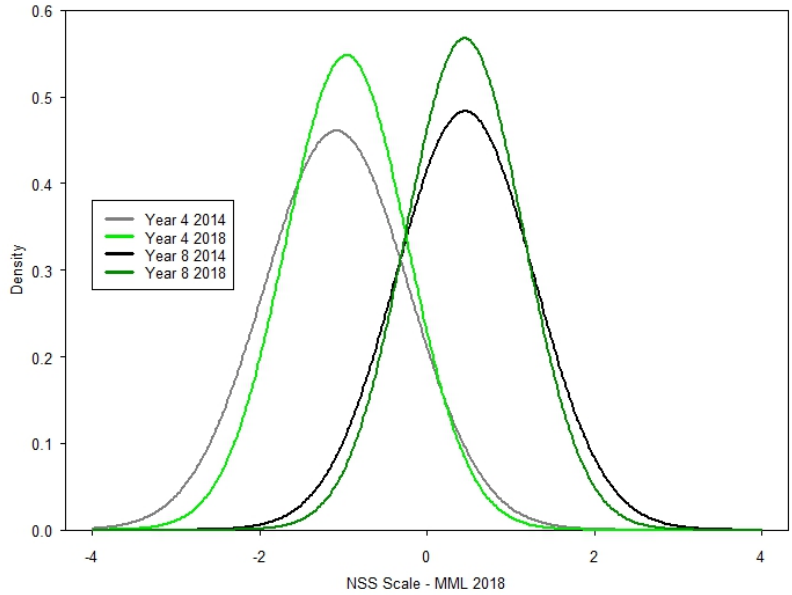


Figure A5.2 Comparative estimated distributions in 2014 and 2018

More detailed findings from trend analysis are included in the 2018 Social Studies report[19].

Linking error

When linking two scales such as these, a linking error should always be considered in the analysis. The size of the linking error is dependent on the differences between pairs of item parameters. In this case, since the correlation between the item parameters is very strong, the linking error is small (0.01 logits). The linking error was calculated as

$$linking\ error = \sqrt{\sum_{i=1}^{L}(\delta_i - \delta_i')^2 * \frac{L}{L-1}}$$, where L is the number of link items

Standard error on differences between means

The trend analysis involves examining differences between means at the two time-points for complete year levels and for key subgroups. The general formula for calculating confidence intervals around an observed difference is

$$1.96 * \sqrt{se_{pooled}^2 + linking\ error^2}$$

[19] *NMSSA Report 20: Social Studies 2018 – Key Findings*.

5. Alignment of the 2018 NSS scale to the NZ Curriculum

NMSSA has a particular interest in the achievement level of Year 4 students against level 2 of the *New Zealand Curriculum*, and the achievement level of Year 8 students against level 4 of the curriculum.

The 2014 curriculum alignment generated boundaries on the 2014 NSS scale to indicate curriculum level cut-scores. The cut-scores were developed by a group of teachers and social studies curriculum specialists in a curriculum alignment exercise described in the 2014 NMSSA social studies report. These cut-scores were then used to estimate how the Year 4 and Year 8 student populations were achieving against year-level appropriate curriculum expectations.

The 2014 curriculum cut-scores were located on a scale, which had been constructed using JMLE estimation. There is no direct transformation from the 2014 JMLE scale to the 2018 MML scale. The first step was to re-estimate the 2014 scale using plausible values, a procedure that was adopted after 2014. The 2014 MML scale was reconstructed and used for equating with the 2018 scale. Percentile equating was used to locate the curriculum cut-scores from the 2014 NSS scale on the 2018 NSS scale. Percentile equating assumes that the proportion of Year 4 students estimated to be achieving at or above level 2, and the proportion of Year 8 students estimated to be achieving at level 4 or above, should not vary with the estimation method. In other words, when the 2014 results are placed on the 2018 scale, the proportion of students meeting expectations should be the same as was reported in 2014 against the original 2014 NSS scale. The cut-scores for 2018 using percentile equating were as follows:

Table A5.1 Final curriculum cut-scores on the 2018 NMSSA Nature of Social Studies scale

	Level 2 (logit)	**Level 2** (NSS units)	**Level 3** (logit)	**Level 3** (NSS units)	**Level 4** (logit)	**Level 4** (NSS units)
Year 4 & 8	-1.37	69	-0.33	98	0.70	127

Appendix 6:
Linking Mathematics and Statistics across Cycle 1 and Cycle 2

Contents:

Figures:

Tables:

1. Introduction

In order to make comparisons across cycles, the National Monitoring Study of Student Achievement (NMSSA) carries out analyses in each learning area to link the assessment results. This document summarises the steps conducted to link the Knowledge and Application of Mathematical and Statistical Ideas (KAMSI) assessment from 2013 with the Mathematics and Statistics (MS) assessment from 2018.

In 2013, the KAMSI scale was constructed using items from group-administered paper-and-pencil assessment. Another scale, called the Mathematical and Statistical Proficiencies (MSP) scale was constructed from in-depth interview tasks and performance activities, administered on a one-to-one basis. In 2018, both group-administered and in-depth items were used to construct the single MS scale. The scale was initially constructed using only the group-administered items. The in-depth items were then added while the calibrations of the group-administered items were anchored.

While both 2013 and 2018 scales were psychometrically sound and robust measures, the 2018 scale was considered to be a richer measure of the same construct, and the decision was made to link the 2013 results to the 2018 scale, using common group-administered items.

2. Technical differences 2013 to 2018

Some technical details regarding estimation have changed between 2013 and 2018. Primarily, plausible values were introduced in 2015 for calculating population estimates. Generating sets of plausible values for the student sample requires a slightly different estimation technique from that used in 2013 for calculating item parameters. These technical changes necessitated a re-analysis of the assessment data from 2013 so that it can be properly compared with the 2018 data.

The re-analysis of 2013 data has been done solely for the purposes of the NMSSA trend analysis. It means that estimates recorded in the 2013 NMSSA mathematics and statistics report cannot be directly compared with those in the 2018 report. Meaningful comparisons across time are restricted to those reported in the trend analysis sections of the 2018 reports.

3. Reconstruction of the 2013 KAMSI scale

At each of Years 4 and 8, the data from the 2013 KAMSI assessment were re-analysed with a process that replicated the 2018 analysis as precisely as possible. This re-analysis involved using marginal maximum likelihood (MML) to estimate item parameters, where joint maximum likelihood estimation (JMLE) was used in 2013. Both estimation methods apply the Rasch model. The main difference between the two estimation procedures is that MML assumes an underlying normal distribution for the student population, whereas JMLE does not.

Linking Process

The 2013 KAMSI scale was constructed with a vertical link through about 800 Year 6 students who completed questions from both Year 4 and Year 8 assessments. This is a standard NMSSA method used to locate Year 4 and Year 8 students on a single scale. This method was used again in 2018. To minimise the aggregation of the vertical linking error across the two cycles, it was decided that MML parameters should be generated separately for the 2013 Year 4 data and the 2013 Year 8 data, with each of the year levels then being independently linked to the 2018 scale (see Figure A6.1).

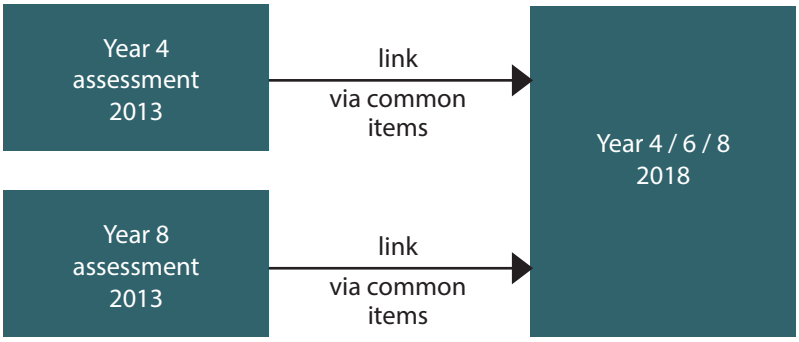


Figure A6.1 Linking scheme for NMSSA mathematics and statistics

For strong linking, there should be a high correlation between the calibrations of the link items for 2013 and 2018, standard deviations should be approximately the same for both sets of calibrations, and item difficulties (calibrations) should cover as wide a range of the scale as possible. In order to achieve a strong link, it is sometimes necessary to remove some items from the link item pool. In this case a small number of items was removed at each of the year levels. Using the remaining link items, a shift was applied to bring the 2013 calibrations in line with the 2018 ones. This shift was calculated separately for Year 4 and Year 8.

For each year level, EAP[20] person estimates were generated for the 2013 data using transformed MML item parameter estimates, and the usual procedure for generating plausible values was carried out.

Year 4 linking outcomes

Of the 30 items that were available for linking, 28 were included in the link calculation. Those 28 items had a correlation of 0.98 and showed a reasonable spread across the scale. The two sets of item parameters had the same standard deviation (1.10 logits).

To bring the 2013 calibrations in line with the 2018 calibrations, the following transformation was applied at Year 4:

$$\delta_i^{2018} = \delta_i^{2013} - 0.18\ logits$$, where δ_i is the estimated parameter of *item i.*

Year 8 linking outcomes

There were 41 items available for linking, and 36 were included in the link calculation. The included items had a correlation of 0.98 and were spread across the scale. The standard deviations of the item parameters were similar across the two time points: 0.90 logits for 2013 and 0.91 logits for 2018. These standard deviations were similar enough to warrant a simple shift.

To bring the 2013 calibrations in line with the 2018 calibrations, the following transformation was applied at Year 8:

$$\delta_i^{2018} = \delta_i^{2013} + 1.89\ logits$$, where δ_i is the estimated parameter of *item i.*

[20] An expected a posteriori (EAP) estimate refers to the expected value of the posterior probability distribution of latent trait scores in a given case.

4. Trend analysis

The Year 4 achievement distributions showed very little difference between 2013 and 2018, while the overall Year 8 mean showed a slight improvement. Figure A6.2 depicts the theoretical normal distributions based on the observed means and standard deviations (in logits).

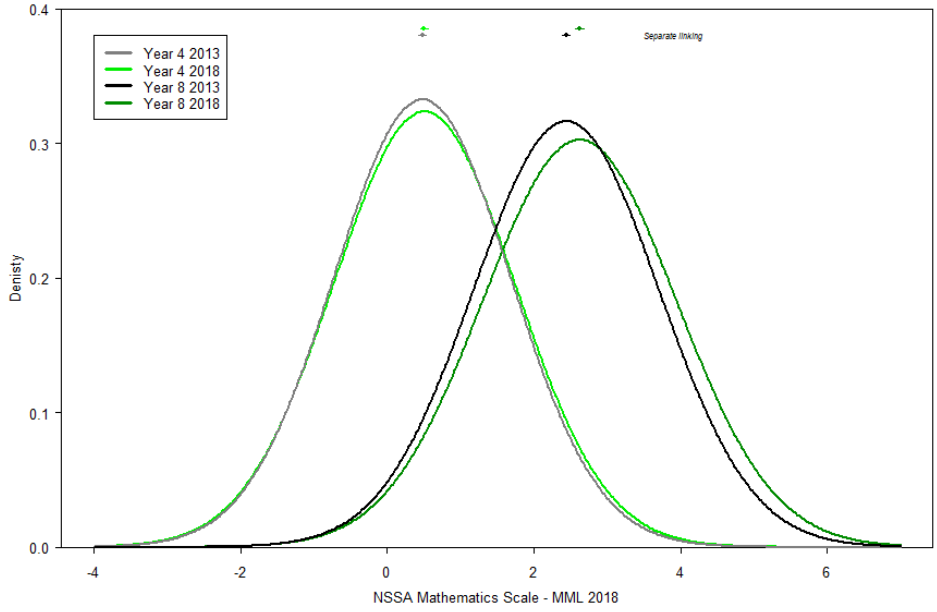


Figure A6.2 Estimated distributions of Mathematics and Statistics scale scores for 2013 and 2018

More detailed findings from trend analysis are included in the 2018 Mathematics and Statistics report[21].

Linking error

When linking two scales, a linking error should be considered in the analysis. The size of the linking error is dependent on the differences between pairs of item parameters. In this case, since the correlation between the item parameters was very strong, the linking error was small: 0.04 logits at Year 4 and 0.03 logits at Year 8.

The linking error was calculated as: $\sqrt{\sum_{i=1}^{L}(\delta_i - \delta_i')^2 * \frac{L}{L-1}}$, where L is the number of link items

Standard error for differences between means

Trend analysis involved examining differences between means at the two time-points for complete year levels and for key sub-groups. The formula used for calculating the confidence interval around an observed difference was:

$$1.96 * \sqrt{se_{pooled}^2 + linking\ error^2}.$$

[21] *NMSSA Report 21: Mathematics and Statistics 2018 – Key Findings.*

5. Alignment of the 2018 MS scale to the NZ Curriculum

NMSSA has a particular interest in the achievement of Year 4 students against level 2 of the *New Zealand Curriculum*, and the achievement of Year 8 students against level 4 of the curriculum.

The 2013 curriculum alignment generated boundaries on the KAMSI scale to indicate curriculum level cut-scores. The cut-scores were developed by a group of teachers and mathematics curriculum specialists in a book-marking exercise described in the 2013 NMSSA mathematics and statistics report. These cut-scores were then used to estimate how the Year 4 and Year 8 student populations were achieving against year-level appropriate curriculum expectations.

The 2013 curriculum cut-scores were located on a scale which had been constructed using JMLE estimation. As there is no direct transformation from the 2013 JMLE scale to the 2018 MML scale, the first step was to re-estimate the 2013 scale using plausible values, a procedure that was adopted in 2015. The 2013 MML scale was reconstructed and used for equating with the 2018 scale. Percentile equating was used to locate the curriculum cut-scores from the 2013 KAMSI scale on the 2018 MS scale. Percentile equating assumes that the proportion of Year 4 students estimated to be achieving at or above level 2, and the proportion of Year 8 students estimated to be achieving at level 4 or above, should not vary with the estimation method. In other words, when the 2013 results are placed on the 2018 scale, the proportion of students meeting expectations should be the same as was reported in 2013 against the original 2013 KAMSI scale.

Because Year 4 and Year 8 results from 2013 have been linked separately to the 2018 scale, the mappings of the 2013 cut-scores on to the 2018 scale differ slightly when using percentile equating (see Table A6.1).

Table A6.1 Curriculum cut-scores on the 2018 NMSSA Mathematics and Statistics scale

	Level 2 (logit)	**Level 2** (MS units)	**Level 3** (logit)	**Level 3** (MS units)	**Level 4** (logit)	**Level 4** (MS units)
Year 4	-0.59	67	0.92	90	2.44	114
Year 8	-0.78	64	0.89	90	2.74	119

In order that the percentage of Year 4 students achieving at Level 2 or above, and the percentage of Year 8 students achieving at level 4 and above, remain consistent with what was reported in 2013, the level 2 cut-score was placed at -0.59 logits, and the level 4 cut-score at 2.74 logits.

In locating the level 3 cut-score, the mid-point between the Year 4 level 3 cut-score (0.92) and the Year 8 level 3 cut-score (0.89) was used. The level 3 cut-score was therefore located at 0.90 logits (logit cut-scores reported here have been rounded to 2 decimal places).

Appendix 7:
NMSSA Assessment Framework for Social Studies 2018

Contents:

Tables:

1. Introduction

This appendix describes the assessment approach that the National Monitoring Study of Student Achievement (NMSSA) took to assess social studies in 2018. It describes how social studies is set out in the *New Zealand Curriculum*[22] (NZC) and outlines the conceptual framework that guided the development of the Nature of Social Studies (NSS) assessment.

2. Social Studies in The New Zealand Curriculum

The NZC states:

> The social sciences learning area is about how societies work and how people can participate as critical, active, informed and responsible citizens. Contexts are drawn from the past, present and future, and from places within and beyond New Zealand.
>
> Through the social sciences, students develop the knowledge and skills to enable them to: better understand, participate in, and contribute to the local, national, and global communities in which they live and work; engage critically with societal issues; and evaluate the sustainability of alternative social, economic, political and environmental practices.
>
> Students explore the unique bicultural nature of New Zealand society that derives from the Treaty of Waitangi. They learn about people, places, cultures, histories, and the economic world, within and beyond New Zealand. They develop understandings about how societies are organised and function, and how the ways in which people and communities respond are shaped by different perspectives, values and viewpoints. As they explore how others see themselves, students clarify their own identities in relation to their particular heritages and contexts. (p. 30)

The assessment of the Nature of Social Studies (NSS) was derived from the achievement objectives in social studies and focused on four constructs: conceptual understanding, active participation in society, values and perspectives, and using information. The constructs covered one or more of four conceptual strands of the social studies learning area.

Definition of constructs

Conceptual understanding (CU)

These are big ideas that students develop about society within social studies. The concepts relate to the four interrelated conceptual strands of social studies in the NZC (see next section). Students were assessed on the extent to which they were able to:

- be informed and critical in their understanding of social studies concepts
- apply, transfer and extend their conceptual understandings across a range of contexts
- demonstrate connections between multiple contexts.

Active participation in society (APS)

APS is to be constructively involved in participating in, or observing, critically informed actions in relation to local or global issues. Students were assessed on the extent to which they were able to:

- recognise the impact of their role in society
- identify opportunities for themselves or others to participate in society
- identify issues or problems
- identify how they or others can take action and/or make decisions based on knowledge and understandings
- recognise the personal or social significance of the contributions of self or others to society.

[22] Ministry of Education. (2007). *The New Zealand Curriculum.* Wellington: Learning Media.

Values/perspectives (VP)

Values are deeply held beliefs about what is important or desirable. They are expressed through the ways in which people think and act. Students were assessed on the extent to which they were able to:

- express values using evidence-based justifications
- recognise diverse values and perspectives in society
- critically analyse values and actions based on these values.

Using information (UI)

Using information requires the gathering and analysis of useful information to inform conclusions and support decision making. Students were assessed on the extent to which they were able to:

- frame questions for an inquiry
- identify appropriate sources of information
- analyse and respond to mathematical information
- present information using appropriate conventions
- use information to make or recognise valid generalisations or references.

Definition of conceptual strands

Identity, Culture and Organisation (ICO)

Students learn about society and communities and how they function. They also learn about the diverse cultures and identities of people within those communities and about the effects of these on the participation of groups and individuals.

Place and Environment (PE)

Students learn about how people perceive, represent, interpret, and interact with places and environments. They come to understand the relationships that exist between people and the environment.

Continuity and Change (CC)

Students learn about past events, experiences and actions, and the changing ways in which these have been interpreted over time. This helps them to understand the past and the present and to imagine possible futures.

The Economic World (EW)

Students learn about the ways in which people participate in economic activities and about the consumption, production and distribution of goods and services. They develop an understanding of their role in the economy and of how economic decisions affect individuals and communities.

What does progress in social studies look like?

The broad expectations of social studies in the NZC for levels 2 and 4 are expressed as the following Achievement Objectives (AOs) in Table A7.1.

Table A7.1 Curriculum level 2 and level 4 achievement objectives in social studies

Social Studies achievement objectives	
Level 2	**Level 4**
Students will gain the knowledge, skills and experience to understand: • that people have social, cultural and economic roles, rights and responsibilities • how people make choices to meet their needs and wants • how cultural practices reflect and express people's customs, traditions and values • how time and change affect people's lives • how places influence people and people influence places • how people make significant contributions to New Zealand's society • how the status of Māori as tangata whenua is significant for communities in New Zealand.	Students will gain knowledge, skills and experience to understand: • how the ways in which leadership of groups is acquired and exercised have consequences for communities and societies • how people pass on and sustain culture and heritage for different reasons and that this has consequences for people • how exploration and innovation create opportunities and challenges for people, places and environments • that events have causes and effects • how producers and consumers exercise their rights and meet their responsibilities • how formal and informal groups make decisions that impact on communities • how people participate individually and collectively in response to community challenges.

3. Curriculum coverage in the NSS assessment

Many questions in the NSS assessment covered more than one construct and more than one conceptual strand. Seventeen tasks were included in the assessment programme, six of which were link tasks from 2014. A range of assessment approaches were used to assess the tasks: computer and paper-and-pencil presented tasks, interviews, group, team (of 4); and the settings in which the questions were located: New Zealand, global or 'other'.

Table A7.2 summarises the relative weighting given to each of the elements in the assessment framework for the NSS. Represented most strongly in the NSS assessment were: questions about identity, culture and organisation (14), conceptual understanding (17), and located in a New Zealand context (14).

Table A7.2 The number of tasks by strand, aspect, setting and assessment approach

Element	**Number of tasks with a focus on each strand, construct, setting and assessment approach** (Total number of tasks = 17)			
Strand	Identity, Culture & Organisation 14	Place & Environment 7	Continuity & Change 5	Economic World 3
Construct	Conceptual understanding 17	Active participation in society 8	Values / perspectives 12	Gathering and analysing information 11
Setting	New Zealand 14	Global 5	Other 3	
Assessment approach	Interview 7	Group (GAT) 8	Team 3	

Table A7.3 shows the coverage of the NSS framework by task, strand, concept, construct, setting and assessment approach. The Kai Moana task, for example, covered two constructs – ICO and PE; two conceptual strands – CU and VP: Status of Māori as tangata whenua and Use of resources and Sustainability. The task was set in a New Zealand context and the assessment approach used an interview. The marking schedule for Kai Moana is presented in Table A7.4. It illustrates how constructs and strands overlap within a task.

Table A7.3 Curriculum coverage in the Nature of Social Studies assessment

	Strands and Concepts		Construct				Setting[1]	Approach[2]
Task Title	**Strands**	**Concepts**	**CU**	**APS**	**VP**	**UI**		
Good Sorts	ICO EW	Benefitting community through social action Market and opportunity	✓	✓			NZ	G
Food Waste	ICO PE CC	Social action Sustainability Cause and effect; Future focus	✓	✓		✓	NZ/ Global	G
Making a Difference	ICO	Common good; Social action Roles and responsibilities Volunteering; Leadership	✓	✓	✓	✓	NZ	G
A Letter Sent Home	ICO CC	Impact of change on people's lives Roles in society	✓			✓	NZ/ Global	G
Special to Our People	ICO CC	Cultural diversity; Status of Māori Identity; Values; Treaty Society's attitudes changes over time	✓	✓	✓	✓	NZ	I
Scaring the Monkeys	ICO	Roles and responsibilities; Community; Consequences of actions	✓	✓	✓	✓	Other	G
Mapping Aotearoa/ New Zealand	ICO PE	Place names reflect identity, culture and heritage; Collaboration; Roles and responsibilities Place names; Location How people record important features of places and environments	✓	✓	✓	✓	NZ	T
Digging Up the Past	CC	Valuing the past	✓		✓	✓	NZ	I
Mission Possible	PE	Location; Natural features; Environment; How people record features of important places	✓		✓	✓	Other	G
Fiapule	ICO	Families; Culture; Celebrations; Relationships	✓		✓	✓	NZ	G
Kai ora Honey	ICO PE EW	Whanau/family; Cultural values; Ancestral land; Resources; Values-based business decision making	✓		✓	✓	Global	G
Fudge for the school fair (LINK)	EW	Factors affecting pricing; Profit; Factors influencing people's purchasing decisions	✓		✓	✓	NZ	I
Kai Moana (LINK)	ICO PE	Status of Māori as tangata whenua Use of resources; Sustainability	✓		✓		NZ	I
Graffiti (LINK)	ICO PE	How formal and informal groups make decisions How people view and use places differently	✓	✓	✓		NZ	I
Moving Here (LINK)	ICO PE CC	Cultural diversity and interaction Effect of people on the environment Cultural interaction can change culture over time	✓		✓		Global NZ	I
Cultural Symbols (LINK)	ICO	What constitutes culture; How symbols communicate identity	✓				NZ Asia	T
When Disaster Strikes (LINK)	ICO	Social responsibility; How people respond individually & collectively	✓	✓			Global NZ	T

Setting[1]: The settings identified for the 2018 social studies tasks were those identified in the 1997 Social Studies in the New Zealand Curriculum document, p. 20: New Zealand, The Pacific, Europe, Asia, Other Settings, Global Settings.

Approach[2]: I = Interview, G = Group, T = Team

Table A7.4 Marking rubric for the Kai Moana task

Title:	**KAI MOANA**	**Level:**	4 & 8
Task Info:	**3 cards, 3 'big ideas' words**	**Approach:**	Interview

Col 1	Q1. Why do we have laws about the amount of seafood and the size of seafood people can take?	CONSTRUCTS: Conceptual Understanding

SCORE:	0	1	2
Criteria:	No response/don't know/unsure Response not relevant e.g. meal size	Surface e.g. extinction of species; people will take too many; won't be any more left; they'll die out	Deeper e.g. sustainability considering others **and** the implications; considering future - leave small ones to grow; cultural value of giving back to the sea
Conceptual Understanding *Use of resources and sustainability*	*Demonstrates no understanding of the concepts*	*Demonstrates understanding of concepts (surface)*	*Demonstrates understanding of abstract concepts (deep)*

Col 2	Here is an opinion some people have about 'Customary Rights' for Māori. Q2. Why might people agree with that? Q3. What do you mean by _____? ***OR*** *Can you tell me more about* _____*?(* (student response from Q2) Q4. Why might people disagree with that? Q5. What do you mean by _____? ***OR*** *Can you tell me more about* _____*?(* (student response from Q4)	CONSTRUCTS: Values & Perspectives Conceptual Understanding

SCORE:	0	1	2
Criteria:	No response/don't know/ unsure Response not relevant Only gives 1 point of view at a **surface** level	Surface – **both** viewpoints recognised e.g. been given rights; they need more for a tangi; it's the law; unfair they get more	Deeper connects with culture/beliefs (one side) e.g. Māori as tangata whenua; we all have equal rights, so it should be the same for all; sacredness of seafood to Māori; discrimination
Recognising diverse values and perspectives	*Unable to explain others' values positions*	*Explains others' values on a simple/surface level*	*Explains others' values on a complex/deep level*
Conceptual Understanding *Status of Māori as tangata whenua*	*Demonstrates no understanding of the concepts*	*Demonstrates understanding of concrete concepts (surface)*	*Demonstrates understanding of abstract concepts (deep)*

Col 3	Q6. Sustainability, Rules, Responsibility Use these big ideas together, to tell me what they have to do with gathering seafood or kai moana.		CONSTRUCT: Conceptual Understanding
SCORE:	**0**	**1**	**2**
Criteria:	No response/don't know/ unsure Response not relevant Only 1 concept discussed	Talks about 2 or 3 individual concepts in relation to kai moana	Explanation with connections between 2 or 3 concepts in relation to kai moana
Connection between concepts	*Shows no awareness of how concepts are connected*	*Explanation of concepts within a specified context*	*Complex/deeper connection of concepts*

Col 4	Q7. Where else might these big ideas be important? (sustainability, responsibility and rules) Q8. How are these big ideas important in those places?		CONSTRUCT: Conceptual Understanding
SCORE:	**0**	**1**	**2**
Criteria:	No response/don't know/ unsure Response not relevant	Explains 1–3 individual words in another context/s e.g. follow rules on road or at school	Explains linked words clearly (2 or 3 words) to a context e.g. rules/responsibilities – follow road rules and this shows you are a responsible driver
Transferring and connecting concepts to different contexts	*Unable to transfer the concepts to a different context*	*Limited transfer of concepts to a different context*	*Clear transfer and linking of concepts to different context*

4. Key competencies, literacy and numeracy in social studies

All of the NZC key competencies are enacted within the social studies curriculum. NMSSA monitors the development of the key competencies by exploring how well students demonstrate these, through discussion or in writing, to meet curriculum purposes.

While the assessment tasks in social studies were focused on achievement, the literacy, numeracy and key competency demands were identified and assessed from these authentic social studies purposes. Table A7.5 sets out the definition of key competencies and how they were developed in social studies.

Table A7.5 Definition of key competencies and how they were developed in social studies

Key Competency (NZC)	Developed in social studies when students:
Thinking	• Pose questions, collect and analyse information, considering varying values and responses made by people and groups, and evaluate findings in a critical and informed manner • Are able to examine/challenge assumptions and perceptions
Using language, symbols and texts **Literacy and numeracy across the curriculum**	• Use and make meaning of the wide range of literacies inherent in the social sciences, including knowledge of texts such as newspapers, graphs, statistics, maps, visual and oral media and so on • Understand how ideas are represented, visualised and constructed e.g. map, table, timeline, graph
Managing self	• Manage themselves throughout a social inquiry approach by acting in ways that are enterprising, resourceful, reliable and resilient
Relating to others	• Interact effectively with others, listen and respond to other points of view, values and perspectives, and recognise alternative responses to social topics, themes and issues in society • Are able to work collaboratively with others
Participating and contributing	• Work and learn cooperatively in groups within the school and their communities and know about the rights, roles and responsibilities of themselves and others • Understand how they are able to be active in society now and in future • Acknowledge how critical, responsible and informed they are in their participation in society

References

Aitken, G. & Sinnema, C. (2008). *Effective pedagogy in social sciences/Tikanga a Iwi: Best evidence synthesis iteration* (BES). Wellington, NZ: Ministry of Education.

Education Review Office (2011). *Enterprise in the New Zealand Curriculum*. Retrieved from https://www.educationcounts.govt.nz/publications/series/iccs/what-do-new-zealand-students-understand-about-civic-knowledge-and-citizenship/how-well-prepared-are-our-year-9-students-to-be-future-citizens

Harcourt, M., Milligan, A., & Woon, B. (Eds.) (2016). *Teaching social studies for critical, active citizenship in Aotearoa New Zealand*. Wellington, NZ: NZCER Press.

Lang, K. (2010). *What do New Zealand students understand about civic knowledge and citizenship?* Retrieved from https://www.educationcounts.govt.nz/publications/series/iccs/what-do-new-zealand-students-understand-about-civic-knowledge-and-citizenship/how-well-prepared-are-our-year-9-students-to-be-future-citizens.

Wood, B. & Milligan, A (2016). Citizenship education in New Zealand policy and practice. *Policy Quarterly*, *12*(3), 65–73.

Wood, B., Taylor, R., Atkins, R., & Johnston, M. (2017). *Creating active citizens? Interpreting, implementing and assessing 'personal social action' in NCEA social studies*, NZCER, Wellington, NZ. Retrieved from http://www.tlri.org.nz/tlri-research/research-completed/school-sector/creating-active-citizens-interpreting-implementing on Sep 12, 2017.

Ministry of Education (2009). *Approaches to building conceptual understandings*. Wellington, NZ: Learning Media Limited.

Ministry of Education (2012). *Taking part in economic communities*. Wellington, NZ: Learning Media Limited.

Ministry of Education (2015) *Education for Enterprise* (E4E) Retrieved from http://nzcurriculum.tki.org.nz/Curriculum-resources/Education-for-Enterprise/Resources

Appendix 8:
NMSSA Assessment Framework for Mathematics and Statistics 2018

Contents:

Tables:

1. Introduction

This appendix outlines the conceptual framework used to support the development of the 2018 mathematics and statistics assessment.

2. Mathematics and statistics in The New Zealand Curriculum

Mathematics and statistics in *The New Zealand Curriculum* (NZC) (Ministry of Education, 2007) is about:

> … the exploration and use of patterns and relationships in quantities, space and time. Statistics is the exploration and use of patterns and relationships in data. These two disciplines are related but different ways of thinking and of solving problems. (p. 26)

The purpose for learning mathematics and statistics is to: 'equip students with effective means for investigating, interpreting, explaining and making sense of the world in which they live' (p. 26). Furthermore:

> By studying mathematics and statistics, students develop the ability to think creatively, critically, strategically and logically. They learn to structure and to organise, to carry out procedures flexibly and accurately, to process and communicate information, and to enjoy intellectual challenge. (p. 26)

Achievement objectives in the mathematics and statistics learning area are organised into three strands for Levels 1–6: number and algebra; geometry and measurement; and statistics. Relative weightings for the three strands at each of these levels are graphically represented in NZC as a venn diagram. According to the NZC, 'It is important that students can see and make sense of the many connections within and across these strands' (p. 26).

3. Continuity between the 2013 and 2018 mathematics and statistics frameworks

The NMSSA project team began development for the 2018 study with an established item collection used in 2013. Items in the 2013 group-administered tasks (GAT) were revised, taking into consideration their performance as measurement tasks in 2013, plus the balance of coverage across the strands and mathematical competencies of problem-solving, reasoning and communicating. A small collection of new items was developed to supplement those retained from 2013. The in-depth items, administered to individual students by a teacher assessor, were revised to strengthen a dual focus on spatial reasoning and problem-solving, as recommended by the NMSSA Curriculum Advisory Panel (CAP) for mathematics and statistics. A minor focus on fractions and percentages was also included, in line with the CAP recommendations. A key consideration for the review of the in-depth tasks was the opportunities they afforded students to demonstrate mathematical competencies.

Table A8.1 compares the achievement measure used in the 2018 programme with those used in 2013.

Table A8.1 Comparison of the achievement measures used in 2013 and 2018 for mathematics and statistics

2013	2018
Two achievement measures were developed in 2013, leading to two scales: *Knowledge and Application of Mathematical and Statistical Ideas (KAMSI);* and *Mathematical and Statistical Proficiencies (MSP)*. The correlation between the two measures was relatively high (.79 at Year 4 and .87 at Year 8) indicating that they measured similar skills and competencies.	The two achievement measures from 2013 were combined into one measure and one scale: *Mathematics and Statistics (MS)*. Existing assessment tasks and almost all new tasks, both paper-and-pencil and in-depth, contributed to the scale. Three new in-depth tasks were not developed to contribute to the scale (and will be reported on separately). These tasks focused on spatial reasoning and mathematical competencies: problem-solving (both individual and collaborative); reasoning; and communicating.
The KAMSI assessment was a group-administered paper-and-pencil assessment that covered the three strands of the mathematics and statistics learning area. The MSP tasks included performance and interview tasks to assess three areas of proficiency: understanding; reasoning strategies and mathematical procedures; and communication.	The MS assessment comprised both pencil-and-paper and in-depth tasks, which together covered the three strands of the mathematics and statistics learning area. The in-depth tasks also included explicit opportunities to assess students' mathematical competencies and collaboration skills.

4. The relationship of the framework to NZC

The 2018 NMSSA study assessed students' knowledge and application of mathematical and statistical ideas across the three content strands described by the mathematics and statistics achievement objectives in NZC. Across the strands, the assessment programme incorporated an emphasis on problem-solving, reasoning and communicating. As much as possible, assessment items were set in meaningful contexts, and involved students thinking mathematically and statistically, solving problems and modelling situations – consistent with NZC.

In order to be able to make an overall claim about students' achievement of the curriculum expectations at Year 4 and Year 8, aspects of the achievement objectives for mathematics and statistics have been broken into three key validity sub-claims at each level (see Tables A8.2 and A8.3). These have been further broken down into what students, who achieve highly at each year level, will be able to do and what they will know.

Table A8.2 Key validity sub-claims for the Mathematics and Statistics assessment: Year 4

	Sub-claims	Students will be able to:	Students will know:
NUMBER	Students solve number problems, using appropriate mental or written methods in flexible ways.	• use a range of additive strategies with whole numbers and fractions, including counting on and back, combining and partitioning • use simple multiplicative strategies with whole numbers and fractions, including equal sharing, skip counting, repeated addition, combining and partitioning.	• forward and backward counting sequences with whole numbers to at least 1,000 • how many ones, tens and hundreds are in whole numbers to at least 1,000 • fractions in everyday use • groupings to 10 • multiples of 10 and 100 that add to 100 and 1,000 • how to write simple equations • the order of whole numbers and unit fractions • numbers can be represented with structured equipment, e.g. on a number line, an abacus, or with place value blocks.
GEOMETRY	**Shape** Students can recognise and use the properties of shapes.	• identify the plane shapes found in objects • recognise drawings and models of simple objects.	• the names of simple two-dimensional shapes and some of their properties.
	Position and orientation Students can describe and interpret directions about position and movement.	• use simple maps to show position and direction • describe different views and pathways from locations on a map using grid references, turns and points of the compass • give clear instructions to re-orient an object.	• the language for turns (clockwise and anticlockwise, right and left), and the main compass points.
	Transformation Students can recognise and use the symmetries of shapes.	• predict and describe the transformations (reflection, rotation, translation) that have mapped one object onto another, and the symmetry of shapes.	• the language of transformation.
PROBLEM-SOLVING, REASONING & COMMUNICATING	Students can clearly communicate the strategies they plan to use to solve problems and evaluate their effectiveness.	• re-phrase a word problem and identify key information. describe an intended strategy • use and adapt their strategy as needed to solve a problem and give a clear explanation of the strategy they used • recognise when they are unsuccessful and suggest alternative strategies that could lead to a solution • identify patterns and relationships in a simple problem that will help them solve more difficult versions of the problem.	• recording (e.g. tally marks, simple equations) is useful for tracking thinking and supporting the communication of strategies.

Table A8.3 Key validity sub-claims for the Mathematics and Statistics assessment: Year 8

	Sub-claims	Students will be able to:	Students will know:
NUMBER	Students can calculate, using appropriate mental or written methods in flexible ways.	• use a range of multiplicative strategies flexibly when operating on whole numbers, fractions, decimals and percentages • use a range of addition and subtraction strategies flexibly on whole numbers, decimals, equivalent fractions and integers • find fractions, decimals and percentages of amounts expressed as whole numbers, simple fractions and decimals • apply linear proportions, including ordering fractions.	• equivalent decimal and percentage forms for everyday fractions • the relative size and place value structure of positive and negative integers and decimals to three places • fractions and percentages in everyday use • commonly used fraction, decimal and percentage conversions • the order of simple fractions and decimals • simple equivalent fractions • the notation for square roots • numbers can be represented with structured equipment, e.g. on a number line, an abacus, or with place value blocks.
GEOMETRY	**Shape** Students can recognise and use the properties of shapes.	• identify classes of two- and three-dimensional shapes by their geometric properties • relate three-dimensional models to two-dimensional representations, and vice versa.	• the names of simple two- and three-dimensional shapes, and the geometric terms for the properties of shapes.
	Position and orientation Students can describe and interpret directions about position and movement.	• interpret locations and directions using compass directions, distances, and grid references • describe different views and pathways from locations on a map using grid references, turns and points of the compass • give clear instructions to re-orient an object. interpret verbal instructions to visualise the re-orientation of an object.	• the language of direction and position.
	Transformation Students can recognise and use the symmetries of shapes.	• predict and describe the transformations (reflection, rotation, translation) that have mapped one object onto another • use the invariant properties of figures and objects under transformations.	• the language of transformation.
PROBLEM-SOLVING, REASONING & COMMUNICATING	Students can clearly communicate the strategies they plan to use to solve problems and evaluate their effectiveness.	• re-phrase a word problem and identify key information. describe a detailed and logical strategy • use and adapt their strategy as needed to solve a problem and give a clear detailed explanation of the strategy they used • recognise when they are unsuccessful and suggest alternative strategies that will almost certainly lead to a solution • generalise patterns and relationships in simple problems that will enable them to solve more difficult versions of the problem and predict unknown amounts in a number sequence.	• making a table to record results can facilitate identifying patterns • algebraic notation is an efficient way to predict unknown numbers in a given sequence.

Table A8.4 shows the spread of items developed for the 2018 study across the strands and competencies. The relative weighting of the three strands approximately reflects the NZC (number and algebra: 60 percent; measurement and geometry: 28 percent; and statistics: 12 percent).

Table A8.4 Coverage of items across strands and competencies in the mathematics and statistics learning area of NZC

Domain	Aspect	Year 4		Year 8	
		GAT items	In-depth items	GAT items	In-depth items
Number	Number knowledge	9	10	11	13
	Number strategies	24	3	26	3
Algebra	Patterns and relationships	6	1	5	1
	Equations and expressions	4	-	8	1
Measurement	Measurement	11	-	11	-
Geometry	Shape	6		5	
	Position and orientation	3	3	4	3
	Transformation	2	-	5	-
Statistics	Statistical investigation	5	-	10	-
	Statistical literacy	2	-	-	-
	Probability	3	-	3	-
Mathematical competencies*	Problem-solving, reasoning and communicating	-	4	-	4

* Four in-depth items explicitly focused on students' mathematical competencies. Three of these items did not contribute to the MS scale and are reported on descriptively as part of the Insights Report for teachers. Elements of mathematical competencies were also incorporated in some of the GAT tasks but were not an assessment focus, so are not indicated in the table.

NMSSA • CYCLE 2

NMSSA Report 21 : TECHNICAL INFORMATION 2018 – Mathematics & Statistics • Social Studies

MINISTRY OF EDUCATION
TE TĀHUHU O TE MĀTAURANGA

UNIVERSITY of OTAGO
Te Whare Wānanga o Otāgo
NEW ZEALAND

NZCER
Rangahau Mātauranga o Aotearoa

ISSN: 2350-3238 (Online only)

ISBN: 978-1-927286-52-4 (Online only)

New Zealand Government