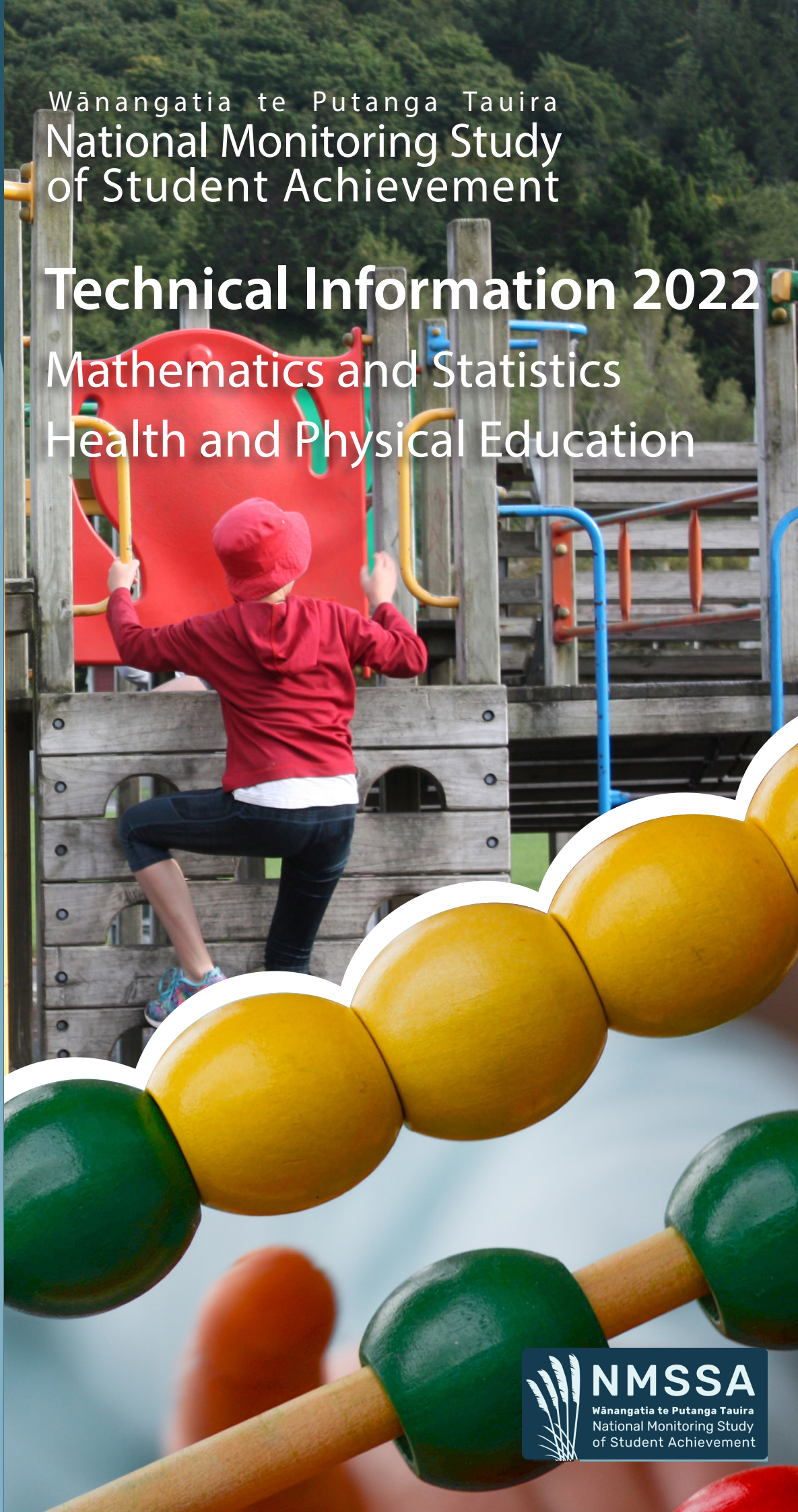
NMSSA • CYCLE 2

NMSSA Report 32 : TECHNICAL INFORMATION 2022 - Mathematics and Statistics; Health and Physical Education

Wānangatia te Putanga Tauira
National Monitoring Study of Student Achievement

Technical Information 2022

Mathematics and Statistics
Health and Physical Education



NMSSA
Wānangatia te Putanga Tauira
National Monitoring Study of Student Achievement

CYCLE 2
NMSSA Report 32

Wānangatia te Putanga Tauira

National Monitoring Study of Student Achievement

Technical Information 2022

Mathematics and Statistics
Health and Physical Education

Educational Assessment Research Unit
and
New Zealand Council for Educational Research

NMSSA Report 32

Technical Information 2022 Mathematics and Statistics | Health and Physical Education
(available online at https://nmssa.otago.ac.nz/nmssa-data/)

National Monitoring Study of Student Achievement Report 32:
Technical Information 2022 – Mathematics and Statistics | Health and Physical Education
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: 978-1-927286-83-8 (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

2022 Project Team Mathematics & Statistics; Health & PE	**EARU**	**NZCER**
Project	Mustafa Asil David Berg Doris Lancaster Catherine Morrison Jenny Ward Jane White Sharon Young	David Coblentz Charles Darr Meredith Dooley Jonathan Fisher Bronwyn Gibb Elliot Lawes Teresa Maguire Jess Mazengarb
Operations	Lee Baker Ruth Barton Linda Jenkins James Rae Alison Richardson	
External Advisors: Jeffrey Smith – University of Otago, Marama Pohatu – Te Rangatahi Ltd		

- *NMSSA Project image, this page: Marelda O'Rourke Gallagher*

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 panel in the learning areas of mathematics & statistics, and health & physical education (PE)
- 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 2022 main 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.

Introduction

This report is comprised of a set of technical appendices that supplement the suite of 2022 NMSSA Key Findings reports. The appendices in this report outline the methods and procedures used to design, develop, implement, and report the results of NMSSA 2022. This report is organized into eight appendices:

- Appendix 1: Sample Characteristics for 2022
- Appendix 2: Methodology for the 2022 NMSSA Programme
- Appendix 3: NMSSA Sample Weights 2022
- Appendix 4: Variance Estimation: NMSSA 2022
- Appendix 5: Linking Maths across Cycle 2 and Cycle 2b
- Appendix 6: Linking Critical Thinking in Health across Cycle 2 and Cycle 2b
- Appendix 7: Assessment Framework for Maths 2022
- Appendix 8: Assessment framework for Health and Physical Education 2022

Appendix 1:
Sample Characteristics for 2022

Contents:

Tables:

Samples for 2022

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

Because the implementation was scheduled for Term 3 2022, the Ministry of Education July 2021 school returns for Year 3 and Year 7 were used for estimating the enrolment of Year 4 and Year 8 students in 2022.

A stratified random sampling approach was taken to select 100 state and state-integrated schools at Year 3 and 100 schools at Year 7. A maximum of 27 students were randomly selected from each school to form national samples at Year 4 and Year 8.

1. Sampling of schools

Sampling is done using Ministry of Education school roll return and school directory information available via the Education Counts website. The algorithm below refers directly to the variables included in those data sets.

Sampling algorithm

From the complete list of New Zealand 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
    - Schools with decile 99 (schools which have not been allocated a decile)
    - Trial schools
    - Chatham Island schools
    - Authority in:
        - Private: Fully Registered
        - Private: Provisionally Registered
    - School Type in:
        - Special School
        - Teen Parent Unit
        - Correspondence School
        - Secondary (Year 9-15)
        - Secondary (Year 11-15)
    - Definition in:
        - Kura Kaupapa Māori
        - Designated Character School
- Stratify the sampling frame by region (using the Regional Council variable) and quintile[1].
- Within each region-by-quintile stratum, order the schools by Year 3 roll size[2].

[1] Quintile 1 comprises state/state-integrated deciles 1 and 2 schools; Quintile 2 comprises state/state-integrated deciles 3 and 4 schools; Quintile 3 comprises state/state-integrated deciles 5 and 6 schools; Quintile 4 comprises state/state-integrated deciles 7 and 8 schools; and Quintile 5 comprises state/state-integrated deciles 9 and 10 schools.

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

- 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, 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.

Substitution procedure

The sampling frames constituted 1413 schools for Year 3 and 878 schools for Year 7 after exclusions had been applied.

Selected schools were invited to participate in 2022 based on 2021 July roll returns. Therefore '2021 Year 3 schools' became '2022 Year 4 schools' and similarly '2021 Year 7 schools' became ' 2022 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.
- As in the initial selection process, schools were assigned to only one sample. A school already selected for one of the samples was then ineligible as a substitute in the other.

In total, 142 schools were invited for Year 4, with 42 declining (29 were from the original sample and 13 replacement schools) before a sample of 100 schools was confirmed for Year 4.

For Year 8 schools, a total of 151 were invited with 51 declining (35 were from the original sample and 16 were replacement schools) before a sample of 100 schools was confirmed for Year 8.

[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.

2. Sampling of students

The sampling plan for selection of students is detailed in this section.

Four nested student samples were intended for the study:

1. A sample that included up to 27 students per school to complete group-administered task (GAT) assessments in *mathematics.*
2. A subset of up to 20 students per school for group-administered task (GAT) assessments in *health and PE*.
3. A subset of up to eight students per school for in-depth (InD) assessment in *mathematics* and *health and PE*.
4. A subset of up to eight students per school for in-depth (InD) assessment in *movement*.

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 2022, 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, and other health or behavioural issues).
- For each school, a computer-generated random number between 1 and 1 million was assigned to each student and they were then ranked in order of their random number from lowest to highest.
- The first 27 students in the ordered list were identified as belonging to the GAT sample for mathematics.

The procedure for selecting students to the in-depth (InD) assessments was as follows:

mathematics and health and PE = students 1-8

movement = students 1-8

- 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 listed 28 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 (the InD sample) were absent or withdrawn on the first day, prior to the start of assessments. They were replaced by students ranked, on a best-match basis (e.g. using the 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 all students in eligible schools). This allows us to assess the national representativeness of the samples in relation to those characteristics.

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.

Achieved samples at Year 4

The initial sample consisted of 2,482 randomly selected students. Principals or parents withdrew 206 students. Substitute students numbered 156. Another 368 students were withdrawn without sufficient time for replacement, were absent or did not respond for other reasons during the assessment period. The achieved GAT sample for mathematics included 2,064 students. The achieved sample for each assessment is displayed in the bottom row of Table A1.1.

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

	Group administered tasks (GAT)		In depth tasks (InD)	
Learning Area	**Mathematics**	**Health and PE (HPE)**	**Maths and HPE**	**Movement**
Maximum students per school	*27*	*20*	*8*	*8*
Students withdrawn by schools before sampling	**241**			
Initial sample:	**2482**	**1921**	**800**	**800**
Students withdrawn by parents or principals after sampling	-206	-	-	-
Substitute students used (replacements for above)	156	-	-	-
Absences, non-responses and withdrawals during assessment period	-368	-105	-122	-78
Achieved sample:	**2064**	**1816**	**678**	**722**

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

	Sample frame *N = 53866* **%**	**Group administered tasks**		**In-depth tasks**	
		Mathematics *N = 2064* **%**	**HPE** *N= 1816* **%**	**Maths and HPE** *N = 678* **%**	**Movement** *N = 722* **%**
Gender					
Boys	**51**	49	48	49	51
Girls	**49**	51	52	51	49
Ethnicity*					
European	**49**	46	46	46	46
Māori	**21**	18	19	20	19
Pacific	**11**	12	11	10	10
Asian	**15**	17	17	17	16
Other	**4**	7	7	7	8

	Sample frame *N = 53866* **%**	Group administered tasks		In-depth tasks	
		Mathematics *N = 2064* **%**	**HPE** *N= 1816* **%**	**Maths and HPE** *N = 678* **%**	**Movement** *N = 722* **%**
Quintile					
1	**17**	13	13	14	15
2	**17**	19	19	19	21
3	**17**	16	16	15	13
4	**22**	24	25	24	25
5	**28**	28	28	27	25
School type					
Composite (Year 1-15)	**1**	1	1	1	1
Contributing	**61**	64	64	62	63
Full primary	**38**	35	35	37	36
Region					
Auckland	**34**	36	34	34	32
Bay of Plenty/Waiariki	**7**	9	9	10	10
Canterbury	**12**	13	12	13	11
Hawkes Bay/Tairāwhiti	**5**	5	5	5	4
Nelson/Marlborough/ West Coast	**3**	3	3	2	3
Otago/Southland	**6**	6	6	7	7
Northland/Tai Tokerau	**4**	3	3	3	3
Taranaki/Whanganui/ Manawatu	**7**	5	6	6	7
Waikato	**9**	8	8	9	9
Wellington	**12**	13	13	10	14

Note: Ministry of Education July 2022 school returns for Year 4 were used for the population percentages.

* Ethnicity is based on the Ministry of Education's prioritised ethnicity statistics.

Achieved samples at Year 8

The initial sample consisted of 2490 randomly selected students. Principals or parents withdrew or excluded 235 students. Substitute students numbered 190. A further 485 students were withdrawn without sufficient time for replacement, were absent or did not respond for other reasons during the assessment period. The achieved GAT sample for mathematics included 1,960 students. The achieved sample for each assessment is displayed in the bottom row of Table A1.3. Due to time constraints, we did not withdraw/replace three Year 8 schools that had consented but had less than our criteria of 8 students – hence less than 800 for the initial sample for InD and movement.

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

	Group administered tasks (GAT)		In depth tasks (InD)	
Learning Area	**Mathematics**	**HPE**	**Maths and HPE**	**Movement**
Maximum students per school	*27*	*20*	*8*	*8*
Students withdrawn by schools before sampling	**-196**			
Initial sample:	**2490**	**1880**	**790**	**790**
Students withdrawn by parents or principals after sampling	-235	-	-	-
Substitute students used (replacements for above)	190	-	-	-
Absences, non-responses and withdrawals during assessment period	-485	-107	-95	-78
Achieved sample:	**1960**	**1773**	**695**	**712**

Table A1.4 contrasts the characteristics of the Year 8 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

	Sample frame *N = 49734* **%**	**Group administered tasks**		**In-depth tasks**	
		Mathematics *N = 1960* **%**	**HPE** *N= 1773* **%**	**Maths and HPE** *N = 695* **%**	**Movement** *N = 712* **%**
Gender					
Boys	**51**	51	51	48	51
Girls	**49**	49	49	52	49
Ethnicity*					
European	**51**	49	50	51	50
Māori	**22**	18	18	19	19
Pacific	**12**	13	13	12	12
Asian	**12**	13	13	13	13
Other	**4**	6	6	5	6
Quintile					
1	**16**	13	13	13	14
2	**18**	20	20	21	21
3	**22**	19	20	19	19
4	**22**	25	24	24	24
5	**23**	24	23	23	22
School type					
Full primary	**37**	44	43	45	46
Intermediate	**57**	52	53	52	51
Secondary (Year 7-15)	**3**	2	2	1	2
Composite (Year 1-15)	**3**	1	1	2	2

	Sample frame *N = 49734* **%**	Group administered tasks		In-depth tasks	
		Mathematics *N = 1960* **%**	**HPE** *N= 1773* **%**	**Maths and HPE** *N = 695* **%**	**Movement** *N = 712* **%**
Region					
Auckland	**33**	34	34	31	33
Bay of Plenty/Waiariki	**8**	7	8	8	8
Canterbury	**12**	15	14	15	11
Hawkes Bay/Tairāwhiti	**5**	6	5	5	5
Nelson/Marlborough/ West Coast	**4**	3	3	3	3
Otago/Southland	**4**	4	5	5	5
Northland/Tai Tokerau	**3**	3	3	3	3
Taranaki/Whanganui/ Manawatu	**7**	7	7	8	8
Waikato	**10**	9	9	10	10
Wellington	**13**	13	13	11	13

Appendix 2:
Methodology for the 2022 NMSSA Programme

Contents:

Tables:

Figures:

Introduction

This appendix outlines the methodology for the 2022 studies in health and physical education (PE), and mathematics and statistics undertaken by the National Monitoring Study of Student Achievement (NMSSA).

1. The health and physical education (health and PE) assessment programme

The 2022 the health and PE assessment programme built on the NMSSA assessment framework for health and PE (see Appendix 8) and included two major components: the critical thinking in Health and PE (CT) assessment; and Learning Through Movement (LTM)assessment.

The CT assessment focussed on three important aspects of learning in health and PE: critical thinking, critical action, and creative thinking in relation to self, others, and society. The assessment involved two parts. In the first part, students responded to tasks involving a range of stimulus material presented on a laptop or through static images. Students recorded their answers in writing using an answer booklet. About 1800 students per year level participated in these assessments. The second part of the assessment involved a series of tasks incorporated into one-to-one interviews with up to 8 students per school with a teacher assessor. Some of the tasks contained material that was used at both Year 4 and Year 8.

The LTM assessment was undertaken by about 625 students at each year level. Students worked individually, in pairs, or in groups of four and were assessed as they participated in physical games. Follow-up interviews were videotaped for later marking and analysis.

Table A2.1 summarises the key differences between the assessment programmes for the health and physical education learning area in 2013/2017 and 2022.

Table A2.1 The key features of the 2017 and 2022 health and PE assessment programmes

	Cycle 2 (2017)	**Cycle 2 extension** (2022)
Programme components	In 2017, the health and PE programme involved three components. **1. Student, teacher and principal questionnaires.** *Student questionnaires* • Attitude and confidence including attitude to school • Opportunities to learn *Teacher questionnaires* • Satisfaction in teaching • Attitude and confidence • Provision of learning opportunities • Professional support *Principal questionnaires* • School policies and practices for HPE teaching and learning • Provision for learning in HPE • Use of external providers **2. Achievement on the Critical Thinking (CT) in HPE assessment** The CT scale was expanded to include more health and movement contexts. The assessment combined new group administered tasks (GAT) using stimuli presented on laptops, and InD tasks (interviews and movement tasks)	In 2022, the health and PE programme involved the same three components. **1. Student, teacher and principal questionnaires** *Student questionnaires* • Attitude and confidence including attitude to school. • Opportunities to learn *Teacher questionnaires* • Satisfaction in teaching • Attitude and confidence • Provision of learning opportunities • Professional support *Principal questionnaires* • School policies and practices for HPE teaching and learning • Provision for learning in HPE • Use of external providers All questionnaires included a brief section on the impact of COVID-19 on teaching and learning. The principals' questionnaire included an additional section on provision of aquatic education. **2. Achievement on the Critical Thinking (CT) in HPE scale** Additional tasks were developed to expand the existing item bank Additional items were developed within existing tasks to interrogate students understanding of hauora across a range of contexts

	Cycle 2 (2017)	Cycle 2 extension (2022)
	In 2013 – all CT data was collected through interview. 3. **Achievement on the Learning Through Movement (LTM) assessment.** The number of tasks assessing movement skills was increased and responses used to form a new measurement scale called Learning Through Movement (LTM)	The wellbeing task was retained from the 2013 and 2017 iterations with a CT element added to the marking rubric. **3. Achievement on the Learning Through Movement (LTM) assessment** A scale was not created for LTM in 2022
Numbers of participants	Up to 12 students per school participated in the GAT. Eight students per school participated in the movement tasks, and 8 students per school participated in the CT and LTM interviews. Twelve students per school responded to the HPE questionnaire.	Up to 20 students per school participated in the GAT. Eight students per school participated in the movement tasks, and 8 students per school participated in the CT and LTM interviews. The attitude and confidence questions were answered by 20 students per school. Ten students responded to the opportunities to learn in health questions, and 10 responded to the opportunities to learn in PE items.

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

Development and trialling of health and PE tasks for the 2022 study

The NMSSA team reviewed all previously used health and PE tasks for possible inclusion in the 2022 assessment programme. Eleven of the 15 tasks from 2017 were retained with some items within these kept in their original format to be used as link tasks, necessary for making comparisons between 2017 and 2022. Others were modified to take into account advice from the HPE curriculum advisory team. A further five tasks were developed.

New and modified tasks were piloted in local schools before being used in a NMSSA trial in March 2022 involving schools in Otago. The student responses from the pilots and the trial were used to refine the tasks and support the development of appropriate marking rubrics. An Item Response Theory (IRT) model[4] 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.

2. The mathematics and statistics (mathematics) assessment programme

The 2022 mathematics assessment programme was based around a similar programme to the one used in 2018 (see Table A2.2). As in 2018, the 2022 programme combined a group-administered assessment, with in-depth activities. Data was also collected using questionnaires for students, teachers, and principals). One difference was that in 2022, the in-depth programme included two student focus groups. One was based on how ākonga Māori and Pacific learners saw their own cultures within their mathematics learning. The second focus group looked at student perspectives on learning opportunities in mathematics.

Table A2.2 The key features of the 2018 and 2022 mathematics and statistics assessment programmes

	Cycle 2 (2018)	Cycle 2 extension (2022)
Programme components	**1. Student, teacher and principal questionnaires** *Student questionnaires* • Attitude and confidence • Opportunities to learn *Teacher questionnaires* • Qualifications in mathematics • Attitude and confidence	**1. Student, teacher and principal questionnaires** *Student questionnaires* • Attitude and confidence • Opportunities to learn *Teacher questionnaires* • Qualifications in mathematics • Attitude and confidence

[4]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.

	Cycle 2 (2018)	Cycle 2 extension (2022)
	• Provision of learning opportunities • Professional support *Principal questionnaires* • School policies and practices for mathematics teaching and learning • Provision for learning in HPE • Use of external providers **2. Achievement in mathematics** A 45-minute group-administered paper-and- pencil assessment (the Mathematics and Statistics assessment) incorporating a mixture of selected response and short constructed response questions. A series of 'in-depth' tasks, most of which were administered in one-to-one student interviews.	• Provision of learning opportunities • Professional support *Principal questionnaires* • School policies and practices for mathematics teaching and learning • Provision for learning in HPE • Use of external providers **2. Achievement in mathematics** A 45-minute group-administered assessment (the Mathematics and Statistics assessment) incorporating a mixture of selected response and short constructed response questions. A series of in-depth task focussed on fractions and algebra. **3. Student perceptions of learning in mathematics** Small focus groups focussed on students' perceptions of culture in mathematics and their learning opportunities in mathematics.
Number of participants	Up to 25 students in each school completed the group administered MS assessment. Up to 8 students in each school completed the in-depth tasks. Up to 12 students in each school completed the student questionnaire. Up to 3 teachers in each school completed the teacher questionnaire.	Up to 25 students in each school completed the group administered MS assessment. Up to 8 students in each school completed the in-depth tasks. Up to 4 students took part in each focus group. Up to 25 students in each school completed the student questionnaire. Up to 3 teachers in each school completed the teacher questionnaire.

Development and trialling of TRM questions

The questions used in the group-administered mathematics assessment programme built on the bank of questions developed for use in 2018. All existing questions were reviewed to ensure the items were inclusive for Māori and Pacific learners. Where needed, changes were made to items. The review included careful consideration of how contexts and illustrations were used. Some additional questions were also developed to extend the bank.

The in-depth tasks were developed to generate insights into students' thinking and capabilities around fractions and algebra. The tasks were open ended and used a range of response modes, including one-to-one interviews with teacher assessors that were videoed. Rubrics were developed to code the student responses to the in-depth items. The rubrics considered students' explanations and reasoning.

The in-depth tasks were trialled in March of 2022 and the responses used to fine-tune the rubrics, questions, and administration instructions.

3. Administration of the assessment tasks

The 2022 study was carried out in Term 3 of 2022. Twenty-four teacher assessors were trained in the administration of tasks during a five-day training programme prior to the main study. During the study, the 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 (responses on paper were scanned).

4. Marking

Marking occurred immediately after the administration stage had concluded. Teacher markers, four of whom had been teacher assessors in the 2022 implementation phase, and final-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. This included double marking of tasks and the consideration of inter-marker agreement rates.

In preparation for marking and based on student samples from the main study, the marking schedules were refined, as necessary, to ensure they reflected the range of responses found in the field. Students' scores were entered directly by the markers into the electronic database.

5. Creating the achievement scales

The Rasch IRT model was applied to student responses from the study to construct scales associated with achievement. This approach included analysing the items used in the assessments for any differential item functioning (DIF) with respect to year level, gender and ethnicity. Items that showed DIF were examined by the task developers, and if their inclusion could not be defended, responses to these items were not included in the scale. In the case of DIF related to year level, the affected items were sometimes split into separate Year 4 and Year 8 items. Very few items showed DIF.

The IRT approach allowed sets of plausible values to be generated for each student involved in the study related to achievement on each of the scales. 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 a random sample of the possible scores 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.

Three scales were developed in 2022 across the two learning areas. These were:

- Critical Thinking in health and PE (CT)
- Mathematics and Statistics (MS)

The scales developed for critical thinking in health and PE, and for mathematics and statistics represented a continuation of the scales developed for the respective studies in Cycle 1- 2.

Standardising the scales

When NMSSA scales are constructed, they are standardised so that:

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

Scales used over more than one cycle can 'lose' these means and standard deviations as achievement patterns change over time. Achievement on the scales generally ranges from about 20 to 180 units.

Scale descriptions

Each of the scales in the two learning areas were described to indicate the range of knowledge and skills assessed. To create the scale descriptions, the scoring categories used to score responses to each item (e.g. 0, 1, 2 or 3) were located on the respective scales. This identified 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 the NMSSA key findings reports a strong sense of what kinds of capabilities were associated with increasing levels of success on the assessments. 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 assessments.

6. Reporting achievement against curriculum levels

The curriculum alignment exercises carried out in Cycle 1 for (health and PE, and mathematics) allowed the results in 2022 to also be reported against curriculum levels. In health and PE, substantial changes were made to the LTM assessment used in 2017 on advice of the HPE curriculum advisory panel. This meant that results from 2013/2017 and 2022 could not be compared using the same scale. We did not undertake a new curriculum alignment exercise for the 2022 movement assessment because of the upcoming refresh of the NZC.

Learning through movement

The complexity of the LTM assessment tasks meant that fewer students were assessed at each year level than was the case for the CT assessment (about 600 students in each year level for LTM compared with 1800 for CT). Analysis indicated that it was not appropriate to develop an LTM measurement scale for reporting purposes using IRT. However, because some tasks in the LTM assessment included items used in 2017, it was possible to make some achievement comparisons on a task-by-task basis.

An indication of how achievement in learning through movement had changed across cycles was provided in a comparison of scores on items used in both cycles. These were presented graphically using 'barbell plots' (see Figure A2.1). Further information about the process used to link the Cycle 1 and Cycle 2 scales can be found in Appendices 5, 6, and 7.

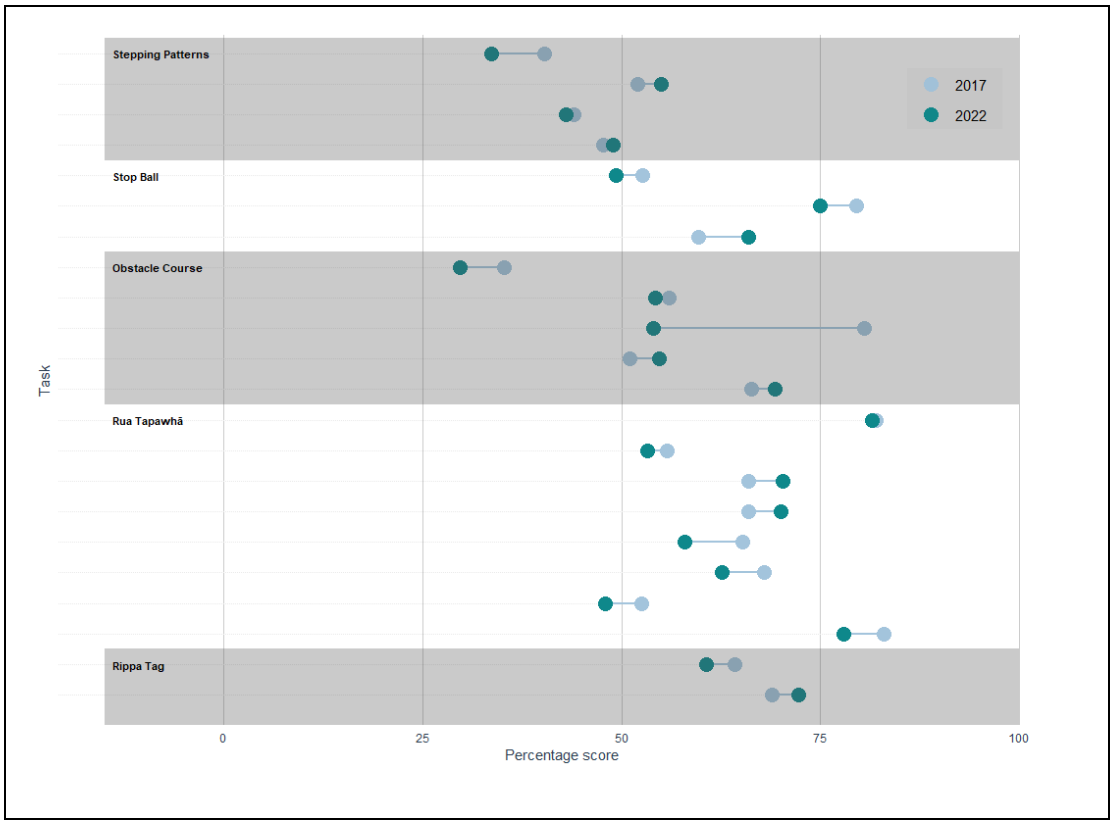


Figure A2.1 An example of the barbell plot used to show how scores on common items had changed from 2017 to 2022 for Learning Through Movement (LTM) assessment.

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.

Student questionnaire

The student questionnaire gathered information about the languages students speak at home. It also gathered information about students' attitudes to school. Within each of the two learning areas (health and PE and mathematics), questions were focused around four themes: students' attitudes to the learning area, students' confidence in the learning area, the learning opportunities students had experienced related to the learning area, and their perceptions of the impact of Covid-19 on their learning.

Four IRT scales were constructed from the student questionnaire data:

- Attitude to health,
- Attitude to PE
- Attitude to mathematics
- Confidence in mathematics

Teacher questionnaire

The teacher questionnaire gathered demographic information about teachers. This included their gender, ethnicity, and teaching experience. Questions for teachers in each of the two learning areas focused on five themes. These were teachers' attitudes to the learning area, their confidence in the learning area, the learning opportunities they had provided for students, the professional support they received for teaching (for example, the professional development they had received), and their responsibility within the learning area, in particular, whether they were a specialist teacher. Additional sections focused on their satisfaction with their teaching role, and their perceptions of the impact of Covid-19 on teaching and learning.

Principal questionnaire

The principal questionnaire included questions focused on demographic information (gender), and school characteristics (the proportion of students with English as a second language). Questions within each of the learning areas focused on three themes. These included school structures that support learning (for example the use of specialist teachers or external providers to deliver programmes and the recency of PLD), teaching and learning (for example, schoolwide processes to support planning, assessment, and reporting) and resourcing. Principals were also asked to rate the impact of Covid-19 on school resourcing, and on student learning.

An additional section was included in the principals' health and PE questionnaire, probing schools' provision of aquatic education.

Measurement scales for the questionnaires

The scales associated with the questionnaires were constructed using the Rasch model. 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, the scales were divided into separate score ranges to provide different reporting categories. For instance, the Confidence in Mathematics scale was broken down into three score ranges: very confident, confident, and not confident. The 'very confident' part of the scale was associated with students mainly using the 'totally agree' category to respond to each of the questionnaire statements related to confidence, the 'confident' section of the scale was associated with students mainly using either 'agree a lot' or 'agree a little', and the 'not confident' part of the scale was associated with students mainly using 'do not agree at all'.

8. Administration of the questionnaires

The student questionnaire was administered on laptop computers supplied by NMSSA. There were three questionnaires types: one for mathematics and one for each component of health and PE. Students responded to the appropriate questionnaire after completing group assessment tasks from the learning area.

Up to three teachers from each school were invited to complete the teacher questionnaire. This included any specialists teaching health and PE or mathematics to the students selected for the study, and the classroom teachers in each school with the most students selected. The principal in each school was invited to complete the principal questionnaire or delegate it to a designated school leader. Teachers and principals had the option of completing the questionnaire online or in a hard-copy.

Appendix 3:
NMSSA Sample Weights 2022

Contents:

Tables:

Introduction

The methodology for calculating sample weights on an annual basis is detailed in *NMSSA Approach to Sample Weighting,* available online at:

https://nmssa-production.s3.amazonaws.com/documents/Sample_Weighting_NMSSA.pdf

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

Tables of estimated[5] means and standard errors calculated with and without sample weights follow. These calculations use the data from the 2022 NMSSA Mathematics and Statistics assessment. Information about the sample can be found in Appendix 1.

Tables 1 and 2 report the estimated means and standard errors (in scale score units) for the Year 4 and Year 8 samples on the 2022 Mathematics and Statistics scale.

1. Summary

Weighted estimates were within one standard error of the estimated unweighted mean.

The recommendation was to proceed with the 2022 analyses without using sample weights.

[5] 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.

Table A3.1 NMSSA Mathematics and Statistics 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	84.0	0.5	83.6	0.5	0.4	2064
Girls	82.4	0.7	82.0	0.7	0.4	1060
Boys	85.7	0.7	85.2	0.7	0.5	1004
NZE	86.2	0.7	86.2	0.7	0.0	955
NZE girls	84.4	0.9	84.3	0.9	0.1	484
NZE boys	88.1	1.0	88.0	1.0	0.1	471
Māori	75.3	0.9	74.9	0.9	0.4	448
Māori girls	74.4	1.2	74.0	1.2	0.4	248
Māori boys	76.3	1.4	75.9	1.4	0.4	200
Pacific	72.9	1.2	72.6	1.2	0.3	269
Pacific girls	72.7	1.6	72.7	1.6	0.0	133
Pacific boys	73.2	1.7	72.6	1.7	0.6	136
Asian	94.0	1.1	94.0	1.1	0.0	413
Asian girls	91.7	1.5	91.6	1.5	0.1	204
Asian boys	96.2	1.5	96.0	1.5	0.2	209
Quintile 1	71.3	1.2	71.3	1.2	0.0	268
Quintile 2	78.2	1.1	78.2	1.1	0.0	389
Quintile 3	83.8	1.1	83.8	1.1	0.0	324
Quintile 4	89.1	1.0	89.1	1.0	0.0	503
Quintile 5	89.5	0.9	89.5	0.9	0.0	580

Table A3.2 NMSSA Mathematics and Statistics 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	115.8	0.6	115.5	0.6	0.3	1960
Girls	113.3	0.8	112.9	0.8	0.4	955
Boys	118.1	0.8	117.9	0.8	0.2	1005
NZE	119.0	0.7	119.0	0.7	0.0	1025
NZE girls	115.4	1.0	115.4	1.0	0.0	481
NZE boys	122.2	1.0	122.2	1.0	0.0	544
Māori	105.0	1.0	105.0	1.0	0.0	423
Māori girls	103.2	1.4	103.1	1.4	0.1	210
Māori boys	106.9	1.4	106.9	1.4	0.0	213
Pacific	101.2	1.2	101.2	1.2	0.0	283
Pacific girls	100.4	1.5	100.4	1.5	0.0	140
Pacific boys	102.0	1.8	101.9	1.8	0.1	143
Asian	129.5	1.5	129.3	1.5	0.2	308
Asian girls	128.6	2.1	128.4	2.1	0.2	149
Asian boys	130.3	2.1	130.1	2.1	0.2	159
Quintile 1	103.6	1.4	103.7	1.4	-0.1	245
Quintile 2	108.0	1.1	108.0	1.1	0.0	384
Quintile 3	113.4	1.3	113.5	1.3	-0.1	372
Quintile 4	120.6	1.1	120.3	1.1	0.3	495
Quintile 5	125.3	1.1	125.1	1.1	0.2	464

Appendix 4:
Variance estimation: NMSSA 2022

Contents:

Tables:

1. Summary

This summary supports the general NMSSA variance estimation paper[6] with specific findings relating to data in NMSSA 2022.

Design effects were calculated using the data collected for the NMSSA 2022 Mathematics and Statistics assessment. This assessment was completed by the majority of the 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 effective 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 2022 varied between 0.9 and 1.9. While the design effects in some cases are fairly large, the effect on the width of confidence intervals is small in practice (less than 1 scale score unit in all but one case).

On the basis of these calculations, it was considered appropriate to maintain the standard multiplier of **0.7** to form an effective sample size in the calculation of statistics dependent on sample size.

Tables showing the effect of the NMSSA complex sample design on the 2022 Mathematics and Statistics assessment follow.

[6] See *Variance Estimation in NMSSA*, at https://nmssa-production.s3.amazonaws.com/documents/Variance_Estimation_NMSSA.pdf

Tables of design effects

Table A4.1 Mathematics and Statistics Year 4 - Comparison of results for different variance estimation methods

Year 4	Mean[7] (SRS[8])	SE (SRS)	SE (TSL[9])	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	0.52	0.03	0.04	0.46	0.57	0.45	0.58	1.81	0.0182	35	2064	1141
NZE[10]	0.68	0.04	0.05	0.60	0.75	0.58	0.77	1.68	0.0229	30	845	502
Māori	-0.04	0.05	0.06	-0.14	0.05	-0.15	0.07	1.35	0.0158	16	448	334
Pacific	-0.26	0.08	0.10	-0.41	-0.12	-0.45	-0.07	1.67	0.0436	29	184	111
Asian	1.22	0.06	0.07	1.10	1.34	1.08	1.36	1.38	0.0211	18	377	274
Female	0.41	0.04	0.05	0.34	0.48	0.32	0.51	1.84	0.0255	36	1038	565
Male	0.62	0.04	0.05	0.55	0.70	0.52	0.73	1.81	0.0272	35	989	546
Female NZE	0.56	0.05	0.06	0.46	0.66	0.44	0.68	1.49	0.0222	22	427	289
Female Māori	-0.10	0.07	0.08	-0.23	0.03	-0.25	0.06	1.38	0.0229	17	248	181
Female Pacific	-0.30	0.10	0.13	-0.50	-0.10	-0.55	-0.04	1.64	0.0564	28	87	54
Female Asian	1.07	0.09	0.11	0.91	1.24	0.87	1.28	1.53	0.0402	24	189	125
Male NZE	0.80	0.06	0.08	0.68	0.91	0.64	0.95	1.84	0.0410	36	418	227
Male Māori	0.03	0.07	0.08	-0.12	0.17	-0.13	0.18	1.25	0.0168	12	200	162
Male Pacific	-0.23	0.11	0.14	-0.45	-0.02	-0.51	0.05	1.73	0.0683	31	97	57
Male Asian	1.36	0.08	0.09	1.20	1.53	1.18	1.55	1.27	0.0207	12	188	150
Low decile	-0.20	0.05	0.06	-0.30	-0.10	-0.32	-0.08	1.57	0.0253	25	437	280
Mid decile	0.52	0.04	0.05	0.45	0.60	0.43	0.62	1.55	0.0194	24	817	528
High decile	0.89	0.04	0.05	0.81	0.97	0.79	0.99	1.61	0.0218	27	810	503

[7] All results in table are quoted in logit units except where indicated

[8] Simple random sample

[9] Taylor series linearisation

[10] New Zealand European

Table A4.2 Mathematics and Statistics Year 8 - Comparison of results for different variance estimation methods

Year 8	Mean[11] (SRS[12])	SE (SRS)	SE (TSL)	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	2.54	0.03	0.04	2.48	2.60	2.46	2.62	1.85	0.0221	36	1960	1059
NZE	2.80	0.04	0.05	2.72	2.88	2.69	2.90	1.58	0.0214	26	918	582
Māori	1.86	0.05	0.06	1.75	1.96	1.74	1.97	1.16	0.0078	7	423	368
Pacific	1.52	0.08	0.09	1.37	1.67	1.34	1.69	1.31	0.0224	14	166	129
Asian	3.57	0.08	0.09	3.41	3.74	3.39	3.76	1.28	0.0220	13	258	202
Female	2.38	0.04	0.06	2.30	2.47	2.27	2.50	1.86	0.0306	36	949	510
Male	2.70	0.04	0.06	2.61	2.78	2.58	2.81	1.78	0.0295	33	996	561
Female NZE	2.58	0.06	0.07	2.46	2.69	2.44	2.72	1.55	0.0283	25	423	273
Female Māori	1.74	0.07	0.08	1.60	1.88	1.58	1.90	1.30	0.0203	14	210	162
Female Pacific	1.55	0.10	0.11	1.34	1.75	1.33	1.76	1.13	0.0120	6	85	79
Female Asian	3.48	0.12	0.14	3.25	3.71	3.21	3.75	1.39	0.0419	18	125	91
Male NZE	2.99	0.06	0.07	2.87	3.10	2.85	3.13	1.45	0.0237	20	495	343
Male Māori	1.97	0.08	0.07	1.82	2.12	1.82	2.12	0.95	-0.0040	-3	213	226
Male Pacific	1.49	0.12	0.14	1.26	1.72	1.21	1.77	1.46	0.0481	21	81	57
Male Asian	3.66	0.12	0.13	3.43	3.89	3.41	3.91	1.19	0.0215	9	133	112
Low decile	1.73	0.06	0.07	1.62	1.84	1.60	1.86	1.38	0.0197	18	409	297
Mid decile	2.52	0.04	0.05	2.44	2.61	2.42	2.62	1.41	0.0164	19	930	658
High decile	3.10	0.05	0.07	3.00	3.20	2.96	3.23	1.78	0.0339	33	621	350

[11] All results in table are quoted in logit units except where indicated

[12] Simple random sample

Appendix 5:
Linking Mathematics and Statistics across Cycle 2 and Cycle 2b

Contents:

Figures:

Introduction

This appendix describes the process used to link results from the Mathematics and Statistics achievement scale across Cycle 2 and Cycle 3, for the purposes of comparing student achievement.

In 2018, a Mathematics and Statistics scale was constructed using items from a group-administered paper-and-pencil assessment. In 2022, Mathematics and Statistics was assessed using a combination of items from the 2018 assessment, and new items developed for 2022. Most of the items were multiple choice with a smaller number being short constructed response.

1. Linking approach

In 2018, vertical linking of the Year 4 and Year 8 samples on the Mathematics and Statistics scale was achieved using results from about 800 Year 6 students who completed questions from both Year 4 and Year 8 assessments. In 2022, the Year 4 and Year 8 samples' results were calibrated separately and independently aligned with the 2018 scale via a shift based on a set of designated linking items. There were no items in common between Year 4 and Year 8, and the scale alignment process relied on the existing vertical linking to determine the relative positions of the Year 4 and Year 8 samples.

Many of the assessment items that contributed to the scale in 2018 were retained for 2022. However, some of those had changes that could potentially affect item functioning to varying degrees. Only dichotomous items with no changes between 2018 and 2022 were considered for linking. Three of these items were deleted from the 2022 calibration due to poor fit. This left 40 unchanged dichotomous items at Year 4, and 50 at Year 8.

Once the scales were aligned, they were considered directly comparable.

2. Linking results

In order to finalise a set of appropriate linking items for each of Year 4 and Year 8, the difficulties of all potential linking items were zero-centred for 2018 and 2022, respectively, and paired differences obtained. Ninety-five percent confidence intervals were constructed for the differences. Any items with a significant difference were excluded from the linking pool.

Year 4

Among the Year 4 items, 9 out of the potential 40 were significantly different between 2018 and 2022. This left 31 Year 4 items for linking. Aligning the means of those 31 linking items across the 2018 and 2022 calibrations required a shift of 0.65 (2dp) logits. The correlation between 2018 and 2022 for all potential linking items is 0.98 (2dp), and for the final linking set the correlation is 0.99 (2dp).

Figure A5.1 shows the paired differences and their associated confidence intervals. Figure A5.2 shows the centred difficulties of the potential linking items from the 2018 calibration against the centred difficulties of those same items from the 2022 calibration. Items that were excluded due to significant difference are shown in orange.

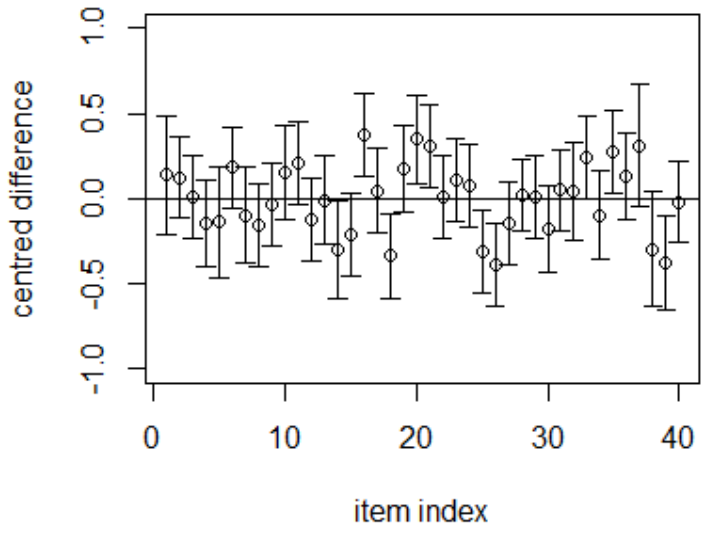


Figure A5.1 Paired differences in Year 4 item difficulties for Mathematics and Statistics, 2018 and 2022

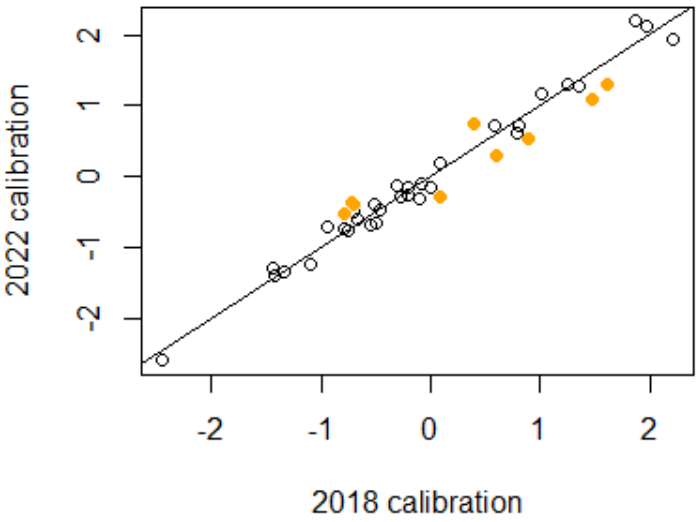


Figure A5.2 Centred Year 4 item difficulties for Mathematics and Statistics, 2018 and 2022

Year 8

Among the Year 8 items, 9 out of the potential 50 were significantly different between 2018 and 2022. This left 41 Year 8 items for linking. Aligning the means of those 41 linking items across the 2018 and 2022 calibrations required a shift of 2.68 (2dp) logits. The correlation between 2018 and 2022 for all potential linking items is 0.98 (2dp), and for the final linking set the correlation is 0.99 (2dp).

Figure A5.3 shows the paired differences and their associated confidence intervals. Figure A5.4 shows the centred difficulties of the potential linking items from the 2018 calibration against the centred difficulties of those same items from the 2022 calibration. Items that were excluded due to significant difference are shown in orange.

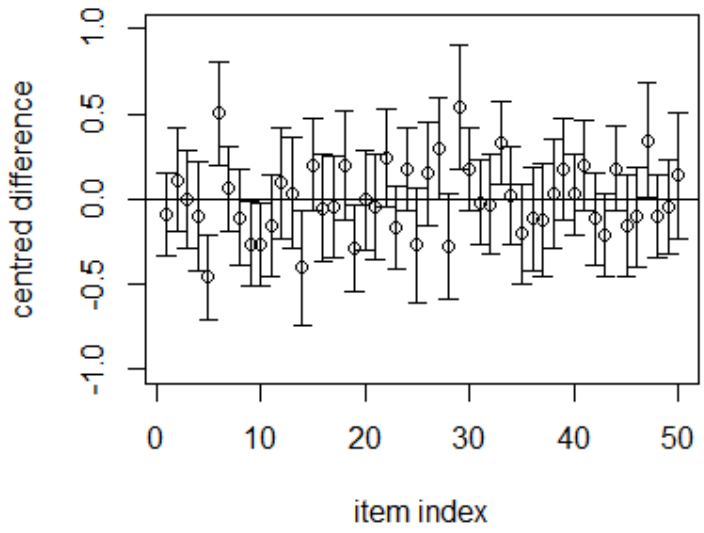


Figure A5.3 Paired differences in Year 8 item difficulties for Mathematics and Statistics, 2018 and 2022

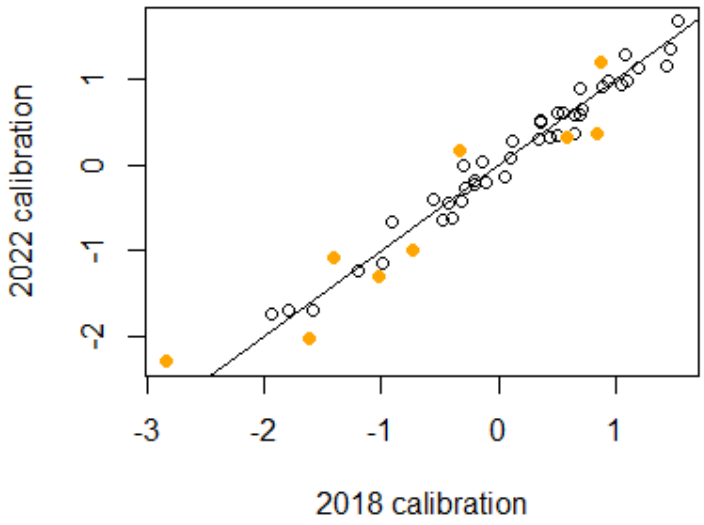


Figure A5.4 Centred Year 8 item difficulties for Mathematics and Statistics, 2018 and 2022

3. Aligning the scales

Across both cycles (2 and 3), marginal maximum likelihood estimation was used to create a Rasch measurement model, and plausible values were generated based on Expected A Posteriori distributions for the calculation of population statistics. The item difficulties obtained through the 2022 Year 4 calibration were shifted by + 0.65 logits to align with 2018, and the item difficulties obtained through the 2022 Year 8 calibration were shifted by + 2.68 logits to align with the 2018 scale. Following these shifts, the 2018 and 2022 scales were considered directly comparable. Because of this direct comparability, there was no need to relocate curriculum cut scores. Instead the cut scores established in 2018 were added to the 2022 scale.

4. Uncertainty associated with linking

Although the 2018 and 2022 Mathematics and Statistics scales are considered directly comparable, the linking process does have some effect on uncertainty around estimates that rely on linking. In order to approximate a measure of the uncertainty associated with linking, pairwise differences between the 2018 and 2022 item difficulties were used to calculate a linking error component:

$$\sqrt{\sum_{i=1}^{L}(\delta_i - \delta_i')^2 * \frac{1}{L(L-1)}}$$

where L is the number of link items, δi represents the average of the thresholds for item i in 2022 and $\delta i'$ represents the average of the thresholds for item i in 2018.

At each of Year 4 and Year 8, linking error was approximated at 0.02 logits.

Linking error has been incorporated, as shown, into the formula used for calculating a confidence interval around an observed difference in means between 2018 and 2022:

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

Appendix 6:
Linking Critical Thinking in Health across Cycle 2 and Cycle 2b

Contents:

Figures:

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 Critical Thinking in Health and Physical Education assessments in 2017 and 2022.

In both 2017 and 2022, the Critical Thinking (CT) scale was constructed using group-administered tasks and in-depth (interview and group/team) items. The 2022 programme built upon the programme used in 2013 and 2017 with items newly developed for 2022, and focussed on three important aspects of learning in Health and Physical Education (HPE): critical thinking, critical action, and creative thinking in relation to self, others, and society.

This document is divided into two sections. The first section describes the linking process. The second section presents the linking error.

1. Linking approach

As mentioned above, NMSSA has used the CT scale to assess achievement in HPE in 2013, 2017, and 2022. Because the 2013 data was already placed on the 2017 scale, pragmatically we decided to locate 2022 data on the 2017 scale to facilitate comparison across the cycles.

The 2017 and 2022 CT scales were linked through common items. There were 37 items common to both 2017 and 2022 assessments, thirty-three of which were considered appropriate to use for linking because their scale locations (relative to the average scale location of all items offered in both 2017 and 2022) didn't change much from 2017 to 2022. To align the scales, an initial calibration of the 2022 items was carried out. The item thresholds from this calibration were then shifted so that the average 2022 initial threshold value of the 33 identified linking items was equal to the average 2017 initial threshold value of the same items. After this shift, the scales were considered directly comparable.

2. Linking error

Figure A6.1 shows the item difficulties (average item thresholds) from the 2017 calibration plotted against the item difficulties from the 2012 calibration. The correlation between the 2017 and 2022 item difficulties is 0.97 with standard deviations 1.04 and 1.05 respectively. While the correlation is high, there is some variance that should be incorporated in precision calculations as linking error, when making comparisons between the 2017 and 2022 administrations.

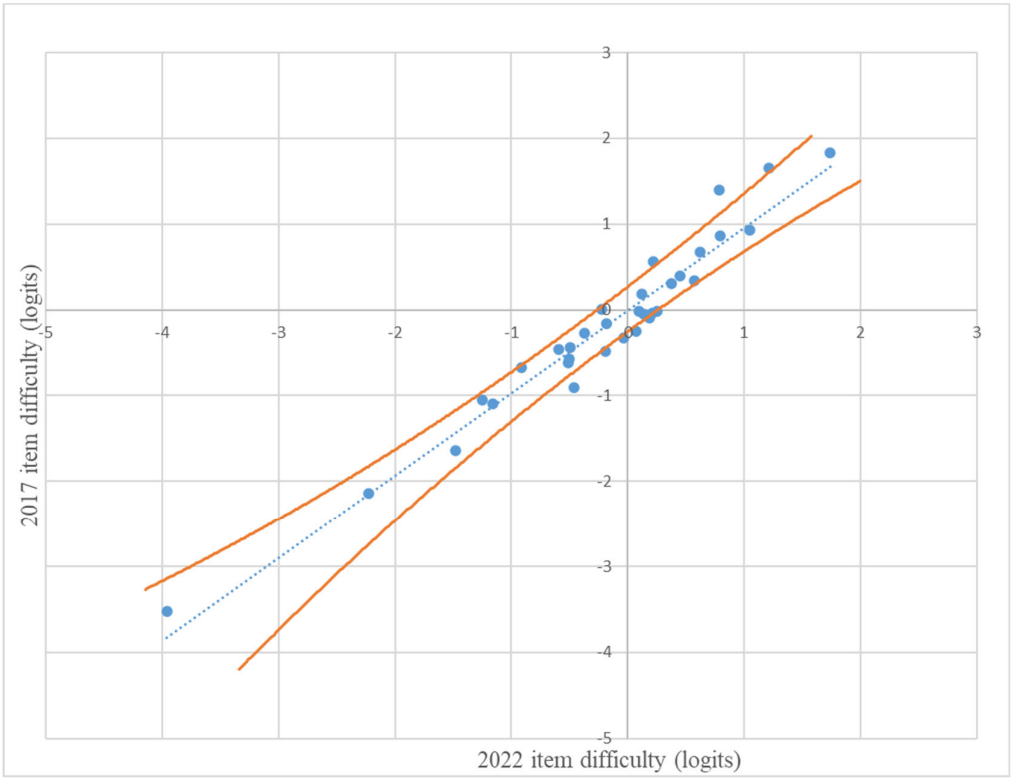


Figure A6.1 Item difficulty estimates for linking items from 2017 and 2022 calibrations

To estimate linking error, pairwise differences between the item estimates from the 2017 and 2022 calibrations, for those items common to both cycles, were used with the following formula applied:

$$\sqrt{\sum_{i=1}^{L}(\delta_i - \delta_i')^2 * \frac{1}{L(L-1)}}$$

where L is the number of link items, δi represents the average of the thresholds for item i in 2022 and $\delta i'$ represents the average of the thresholds for item i in 2017.

Linking error was incorporated in calculation of the confidence intervals around differences in means between the cycles (for the purposes of trend analysis). The formula used for calculating the confidence interval around an observed difference was:

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

Linking error was estimated at 0.04 (2dp).

Appendix 7:
Assessment Framework for Maths 2022

Contents:

Tables:

Introduction

This appendix presents the assessment framework used to support the development of the 2022 NMSSA mathematics and statistics (MS) assessment programme.

1. Mathematics and Statistics in The New Zealand Curriculum

The 2022 NMSSA MS assessment programme was designed to assess achievement at Year 4 and Year 8 in the MS learning area of *The New Zealand Curriculum* (NZC) (Ministry of Education, 2007). According to the NZC, MS is about:

> 1. …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 NZC notes that 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:

> 2. 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 of the NZC: number and algebra; geometry and measurement; and statistics. Relative weightings for the three strands at each of these levels are graphically represented in the NZC using 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).

2. Shape of the assessment

The 2022 assessment involved two components:

- a group-administered component completed by all Year 4 and Year 8 students involved in the 2022 NMSSA study, and
- an in-depth component with a selection of rich assessment tasks intended to give insights into the nature of students' understanding, completed by a subset of the students involved in the group-administered component.

The group-administered component

We began the development for the group-administered component with an established item bank consisting of selected response items and short constructed-response items. The items had been used in 2018. These items were revised and refreshed for use in 2022. In refreshing the bank, special attention was paid to making the items more reflective of, and engaging to, the diverse range of students in New Zealand, especially ākonga Māori, Pacific learners, and learners in lower decile schools. In addition, a small collection of new items was developed to supplement those retained from 2018. All items were administered in the context of a 40-minute paper-and-pencil assessment. Multiple assessment forms, connected through common items, were used to ensure all items in the banks were administered.

The 2022 group-administered component was designed to assess students' knowledge and application of mathematical and statistical ideas across the three content strands described by the MS achievement objectives in NZC. Where appropriate, assessment items were set in meaningful contexts, and involved students thinking mathematically and statistically, solving problems, and modelling situations.

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 MS were broken into three key validity sub-claims at each level (see Tables A7.1 and A7.2). These were further broken down into what students, who achieve highly at each year level, will be able to do and what they will know.

Table A7.1 Key validity sub-claims for the Mathematics and Statistics assessment 2022: 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 1000. • How many ones, tens, and hundreds are in whole numbers to at least 1000. • Fractions in everyday use. • Groupings to 10. • Multiples of 10 and 100 that add to 100 and 1000. • 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 A7.2 Key validity sub-claims for the Mathematics and Statistics assessment 2022: 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 A7.3 shows the spread of items used in the 2022 study across the strands and competencies. The relative weighting of the three strands approximately reflects the NZC (number and algebra: 55 percent; measurement and geometry: 32 percent; and statistics: 13 percent).

Table A7.3 Coverage in the Mathematics and Statistics assessment across strands of the mathematics and statistics learning area

Domain	Aspect	Year 4	Year 8
Number	Number knowledge	8	8
	Number strategies	26	33
Algebra	Patterns and relationships	7	6
	Equations and expressions	6	5
Measurement	Measurement	13	16
Geometry	Shape	7	5
	Position and orientation	2	5
	Transformation	4	5
Statistics	Statistical investigation	6	6
	Statistical literacy	2	1
	Probability	3	5

The In-depth component

The In-depth component involved a series of open-ended tasks. Most of these tasks were designed to explore students' developing understandings involving fractions and algebra. These tasks also provided some insight into how students used mathematical processes, including communication skills such as explaining thinking and constructing a reasoned argument. Most of the In-depth fraction and algebra tasks were presented as part of a one-to-one interview with a teacher assessor (TA) and some were done independently in small groups.

The In-depth component also included two focus group interviews with students. One of these focussed on the connections ākonga Māori and Pacific learners made between their cultures and learning mathematics at school. The second looked at students' perspectives related to learning opportunities in mathematics. The two focus group activities were designed to generate insights into learning and were not assessed for achievement.

The students' responses to the tasks were coded using specially prepared rubrics which considered students explanations and reasoning The in-depth tasks were trialled in March of 2022 and the responses were used to fine-tune the rubrics, questions, and administration instructions.

Table A7.4 lists the tasks and their emphasis.

Table A7.4 Task titles and emphasis for the in-depth component of the mathematics and statistics assessment programme

Task title	Emphasis
Fractions	Understanding fractions
True or False	Algebra (Equations)
Chocolate bar	Understanding the whole
Shaded circle	Understanding fractions
Closest to 3/5	Size of fractions
Making decorations	Fractions as quotients
What's the symbol	Algebra the equals sign
Missing number	Algebra number sentences
Tukutuku	Algebra patterns
Are they the same	Fractions equivalence
Fractions on a number line	Fractions as numbers
Greater, less, equal	Comparing fractions
½ and 3/4	Size of fractions
Coloured Counters	Size of fractions
Sharing a cookie	Fractions as quotients
Adding fractions	Adding fractions

.

Appendix 8:
Assessment Framework for Health and Physical Education 2022

Contents:

Tables:

Figures:

Introduction

This appendix describes the assessment approach that the National Monitoring Study of Student Achievement (NMSSA) took to assess health and PE in 2022. It describes the health and PE learning area of the NZC and outlines the conceptual framework that guided the development of the Critical Thinking (CT) and Learning Through Movement (LTM) assessments used by NMSSA to assess health and PE. In 2022 NMSSA assessed health and physical education (HPE) for the third time. It was previously assessed in 2013, and 2017.

1. Health and PE in the New Zealand Curriculum

The focus of the health and PE learning area is on 'the well-being of the students themselves, of other people and of society through learning in health-related and movement contexts' (NZC, p. 22). Four underlying and interdependent concepts are at the heart of this learning area: hauora, attitudes and values, a socio-ecological perspective and health promotion. Learning activities in health and PE arise from the integration of these four concepts with four strands (and their achievement objectives) and seven key learning areas.

The four strands are:

- personal health and physical development
- movement concepts and motor skills
- relationships with other people
- healthy communities and environments.

The seven key areas of learning are: mental health, sexuality education, food and nutrition, body care and physical safety, physical activity, sports studies and outdoor education. HPE encompasses three different but related subjects: health education, physical education, and home economics.

The NZC (p. 23) states:

> In health education, students develop their understanding of the factors that influence the health of individuals, groups and society: lifestyle, economic, social, cultural, political, and environmental factors.
>
> In physical education, the focus is on movement and its contribution to the development of individuals and communities. By learning in, through and about movement, students gain an understanding that movement is integral to human expression and that it can contribute to people's pleasure and can enhance their lives.

2. The health and PE assessment (Critical thinking in health and PE)

The 2022 the NMSSA assessment programme was focused around two achievement measures: Critical Thinking in health and PE (CT) and Learning Through Movement (LTM).

LTM was assessed as a separate component of health and PE, however, some items from the movement tasks contributed to the CT scale.

The Critical Thinking in health and PE construct

The CT assessment encompasses the three areas of thinking important to health and PE: critical thinking, critical action and creative thinking.

Critical thinking includes thinking about:

- *self and others*: understanding different perspectives and points of view relating to health and well-being, (including inclusiveness and diversity), justifying one's opinions and attitudes
- *information*: examining, analysing, critiquing and challenging information
- *society*: understanding the impacts of the (social, environmental, political, cultural) determinants on well-being.

Critical action includes action for:

- *self*: an understanding of strategies and the ability to manage healthy lifestyles and relationships, risk and resilience
- *others*: the ability to plan and engage in health promotion to bring about change as individuals and collectively.

Creative thinking supports and enhances well-being for oneself and others and includes:

- an understanding of visioning and big picture thinking
- the ability to engage in problem solving and finding solutions
- an ability to express oneself through movement and to interpret the movement of others[13].

Table A8.1, sets out the indicators of student achievement in relation to the three areas of thinking developed by the NMSSA team to assess the achievement objectives at curriculum levels 1 to 4 of the health and PE learning area. The indicators were developed for the 2017 assessment programme[14].

Table A8.1 Indicators of student achievement in three areas of thinking in HPE at levels 1 to 4 of the NZC

	Critical thinking	Critical action	Creative thinking
	Students can:	*Students can:*	*Students can:*
NZC Level 1	• Use personal knowledge • Locate/retrieve simple information from a single source • Communicate ideas using everyday language • Describe a personal feeling or idea • Describe changes to self and others	• Use personal knowledge/ experience to inform decision- making • Recognise issues of personal significance: suggest possible actions • Relate to others	• Convey an imaginative idea about how to solve a problem, but with little relationship to efficacy
NZC Level 2	• Locate/ retrieve basic information from a single source and align it with prior knowledge to show a more developed understanding • Communicate ideas using everyday language to describe objects and events • Describe benefits to well-being/hauora • Express an opinion and elaborate with simple reasons • Describe different values and viewpoints • Identify a message and make inferences • Identify main ideas and some details • Recognise factors that influence choices	• Decide on and justify an action to address an issue; identify some possible positive and negative impacts of proposed actions • Consider and demonstrate respect, manaakitanga, aroha and responsibility • Suggest strategies to support others	• Offer solutions to health-related problems and consider how to convey these

[13] NMSSA Report 3: Health and Physical Education 2013, p. 13

14 NMSSA, Technical Information Report, 2017, p 57

	Critical thinking	Critical action	Creative thinking
	Students can:	*Students can:*	*Students can:*
NZC Level 3	• Make inferences and provide evidence • Identify another's point of view • Look at a proposition from a range of perspectives • Agree / disagree with a view and provide a convincing justification • Describe the impact of social and cultural determinants on well-being/hauora; understand and describe models of well-being/hauora • Recognise discrimination and assumptions e.g. gender stereotypes and body image messages • Recognise media and consumer influences e.g. persuasive messages, target audiences	• Compare and demonstrate ways of establishing and managing relationships • Identify and affirm the feelings and beliefs of self and others • Decide on and justify an action to address an issue; identify some possible positive and negative impacts of proposed actions • Propose possible actions to mitigate discrimination • Identify risks and plan safety strategies	• Accommodate big picture issues – combine prior knowledge, new knowledge and imaginative thinking to come up with tentative solutions to problems. Ideas are practical and are built on logical reasoning • Describe personal strategies for enhancing well-being/hauora, and coping with social and physical changes e.g. managing competition
NZC Level 4	• Describe the complexities of an issue and possible impacts of actions e.g. changing relationships, discrimination • Reflect on social, cultural, environmental, and economic factors that impact on the well-being of self, others and society • Recognise that people can be deliberately positioned and analyse how that has been developed • Explore and identify a range of cultural perspectives • Critique the influence of the media on people's lives e.g. gender stereotypes, relationships, body image, discrimination	• Decide on and justify an action to address an issue and effect change; identify and evaluate positive and negative impacts of actions • Access and use information to make and action safe choices • Identify and demonstrate positive and supportive relationships • Recognise ways to manage healthy lifestyles • Plan strategies to support self and others in a range of environments e.g. online • Recognise how and plan to take individual and collective action to promote community well-being	• Accommodate big picture solutions – combine prior knowledge, new knowledge and imaginative thinking to come up with tentative solutions to problems. Ideas have merit and are rationally justified and evaluated • Transfer learning to other situations

3. Curriculum coverage in the health and PE assessments (including CT and LTM)

Table A8.2 presents the curriculum coverage matrix for the CT and LTM assessments by strand and component. The shaded cells in the table indicate that aspects of the task (represented in each row) were associated with the component represented in the column.

Table A8.3 shows the same tasks aligned with components of the CT and LTM constructs.

Tasks which were developed for the 2022 assessment are marked with an asterix. Task names in italics are group administered (GAT) tasks.

Table A8.2 Coverage matrix for the HPE assessment by strand, and component NMSSA constructs

	Strand A Personal Health and Physical Development				Strand B Movement Concepts and Motor Skills				Strand C Relationships with other people			Strand D Healthy Communities and environments		
TASK TITLE	Personal growth and development	Regular physical activity	Safety management	Personal identity	Movement skills	Positive attitudes	Science and technology	Challenges and social and cultural factors	Relationships	Identity, sensitivity and respect	Interpersonal skills	Societal attitudes and values	Community resources	Rights, responsibilities and laws/ people and the environment
An important message														
*Belonging to groups**														
Charlotte's letter														
*Crash Pad Game**														
Fair Play														
Fit Bit														
Gaming														
Hauora														
*Health message**														
Kai														
MM Play ground														
New School														
Play or breaktime*														
*Poster**														
Powerade														
School gardens*														
*Social media **														
*Te whare tapa whā **														
Tough Boris														
Water safety *														
Obstacle course														
Poi rākau*														
Rippa Tag														
Rua tapawhā														
Stepping patterns														
Stop Ball														

Table A8.3 Coverage matrix for the HPE assessment by NMSSA constructs

	Construct Critical thinking in HPE			Construct Learning through movement					Underlying concept – hauora		Literacy and numeracy	
TASK TITLE	Critical thinking	Critical action	Creative thinking	Locomotion/ object control	Interpreting the movement of others	Strategies and tactics	Adaptability/ creativity	Team work	Understanding a model	Understanding wellbeing	Literacy	Numeracy
An important message	■											
*Belonging to groups**	■	■	■									
Charlotte's letter	■	■	■							■		
*Crash Pad Game**		■			■						■	
Fair Play	■											
Fit Bit	■	■										■
Gaming	■		■									
Hauora	■								■	■		
*Health message**	■	■									■	
Kai	■	■	■								■	
MM Playground	■	■										
New School	■	■	■									
Play or breaktime*	■	■	■							■		
*Poster**	■									■		
Powerade	■											
School gardens*	■	■	■							■		
*Social media **	■	■	■								■	
*Te whare tapa whā**									■	■		
Tough Boris	■	■								■	■	
Water safety *	■	■	■									
Obstacle course								■				
Poi rākau*	■		■	■		■	■	■		■		
Rippa Tag				■		■						
Rua tapawhā	■	■		■		■						
Stepping patterns	■		■	■			■			■		
Stop Ball				■		■						

4. The CT assessment

The CT assessment comprised 20 tasks. Year 4 students responded to combinations from 18 of the tasks; Year 8 students responded to combinations from the whole bank of 20. Each task included a set of items based on one theme or idea. Descriptive criteria were used to mark each item. These were scored dichotomously (0 or 1) or using partial credit scales that ranged from 0 to 2, 0 to 3 or 0 to 4.

Twenty students per school responded to four or five tasks requiring written responses, and a subset of up to eight of these students participated in four or five tasks in one-to-one interviews.

Example of CT assessment task

The following task is group administered task (GAT) from the health and PE CT assessment. The main features of the task are shown (the curriculum strand/s, health and PE key areas of learning, and task stimulus material). The task consists of several items. Examples of the questions students responded to, the scoring guide and possible student responses are illustrated.

Task: *Social Media*

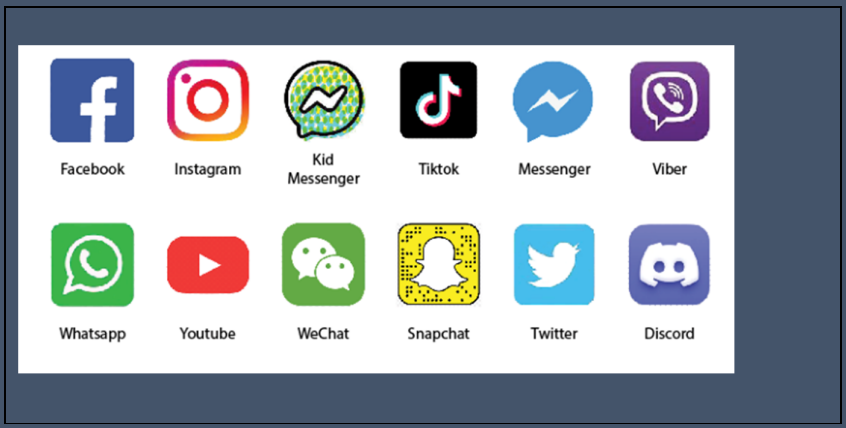
The Social Media task was developed for Year 8 students only, for the 2022 NMSSA study of health and PE achievement. It is a task for Year 8 students only. In the task, students identified the social media apps they used, and were asked to consider impacts of social media, and strategies they might take to mitigate a negative impact.

The Social Media task contained five items. Item one was a tick box repose for those who had not used social media. Items 2 and 3 required students to report positive and negative aspects of social media (Figure A8.1). The fourth item required students to identify a negative outcome of social media, and to describe how that might impact the receiver. (Figure A8.2). The fifth and sixth items asked for suggestions of strategies, and the final item asked for an opinion on the value of age restrictions on App use. (Figures A8.3 – A8.5).

Curriculum elements: Personal health; Relationships with other people; Healthy environment

Item 1. Here are some icons for social media apps.

Circle the ones that you use. If you don't use social media, then put a tick in the box. (Not marked)



Items 2 and 3

What do you think is good about social media?

What do you think is not so good about social media ?

Component: Identifies different perspectives (Critical Thinking)

Scoring category	Example responses
0: Limited response	"The icons are colourful"; "I'm not allowed to use it"
1: Simple description of a positive or negative effect	"It's entertaining"; "People can be mean"
2: Explanation connects to the effect on people's wellbeing/ feelings of self-worth	"Can help people feel connected"; "share viewpoints"; OR "increases social anxiety"

Figure A8.1 Items 1, 2 and 3 of the CT assessment Social Media (Year 8)

Item 4. Sometimes when people use social media, they might have negative experiences that are not good for their hauora. Give one example people might have when using social media. How might this experience affect them?	
Component: Safety (Personal health) – identify risk in an online environment	
Scoring category	**Example responses**
0: Inappropriate response	"Sore eyes"
1: Identifies a negative experience and provides a simple explanation of how if might affect people	"Putting an awful photo of someone on line – the person might be hurt/sad"
2: Identifies a negative experience and provides a deeper explanation of how if might affect people's wellbeing	"Putting an awful photo of someone on line – the person might not feel they look good enough and that affects their self-esteem"

Figure A8.2 Item 4 of the CT task Social Media (Year8)

Item 5. If this experience happened to some one you know, what could they do about it?	
Component: Safety management - Describe/ plan safe practices to manage online risks	
Scoring category	**Example responses**
0: Inappropriate/limited response	"Be mean back"
1: Can provide one strategy	"Talk to someone you trust"
2: Can provide two different strategies	"Contact Netsafe to find out what to do"; "Block the person"

Figure A8.3 Item 5 of the CT task Social Media (Year 8)

Item 6. What do people need to think about carefully before they post or share something on social media	
Component: Safety management; Personal responsibility	
Scoring category	**Example responses**
0: Inappropriate/ limited response	"Does the photo fit the page?"
1: Gives one idea	"What if my family sees it?" "Do I have permission to post this?" etc
2.Gives two different ideas	"Is it respectful to others and self?" "Am I happy for it to be online for ever?"

Figure A8.4 Item 6 of the CT task Social Media (Year 8)

Item 7. Most social media apps have an age restriction. 1. Do you think they should have an age restriction? YES MAYBE NO (circle response) 2. Why do you think that? Justify your ideas and make them as clear as possible.	
Component: Considers current law related to online use. (Rights, responsibilities and laws)	
Scoring category	**Example responses**
0: Limited response. States an opinion with limited or no justification.	"No – because I want to see it."
1: States an opinion and provides a simple justification for what/why not…	"Yes because there can be bad things for children there."
2. Opinion is thoughtful and clearly stated. Considers impact on others in a deeper response.	"Yes/ to protect young people's wellbeing and vulnerability"; "Maybe/ depends on the maturity of the viewer and their awareness of the consequences." "No - it's not up to other people to decide what we see – we have individual rights."

Figure A8.5 Item 7 of the CT task Social Media (Year 8)

5. The Learning Through Movement assessment (LTM)

The LTM assessment focused primarily on the strand: Movement concepts and motor skills and used authentic movement contexts (games) to assess students' ability to do things such as:

- develop and carry out complex movement sequences
- move in a range of ways
- strategise, communicate and cooperate
- think creatively – express themselves through movement, and interpret the movement of others
- express social and cultural practices through movement[15].

Five of the tasks used in 2017 were part of the LTM assessment in 2022.

The LTM construct

Table A8.4 sets out indicators of student achievement in relation to technique; movement dynamics; tactics; perceptiveness; adaptability; and team work. These indicators were synthesised from the work of New Zealand researchers and were based on the framework proposed by Ovens and Smith (2006)[16].

Table A8.4 Movement skills and indicators for the Learning Through Movement (LTM) assessment

Students can:	Skills	Indicators:
• play together in positive ways • engage in games and physical activities and include others	Locomotion (run/step/jump/dodge/evade/hop/land) and object control (catching, passing)	• Posture appropriate to the movement purpose • Movement dynamics are efficient, fluid and balanced • Ready position, follow through
	Movement dynamics/ effectiveness of actions	• Controlled • Balanced and stable • Accurate • Quick
	Strategies/tactics/follow rules	• Deliberate problem-solving decisions • Actions maximise performance • Identify, describe and justify game strategies – own and opponents • Follow rules of a game
	Creativity/adaptability	• Create novel movements or movement sequences fluidly/rhythmically and successfully – with and without equipment • Think creatively about physical activity, physical resources, and physical activity environments • Adapt games for inclusivity
	Team work	• Work collaboratively • Express and accept ideas • Communicate well • Take direction • Show leadership • Be inclusive • Critique and analyse e.g., give feedback • Work with other people in physical activity contexts Accept others' ideas about movement • Communicate movement based ideas (including critique and analysis)
	Perceptiveness	• Perceive opportunities for actions that the opposition or environment allows • React and respond to game play • Anticipate behaviour of opposition

[15] NMSSA Report xxx : Health and Physical Education 2017, p. 66.

[16] Ovens, A & Smith, W (2006) Skill: making sense of a complex concept. *Journal of Physical Education in New Zealand*, 39(1) , 72-82.

Students' skills were categorised according to high; mid or low range achievement (Table A8.5).

Table A8.5 Achievement range

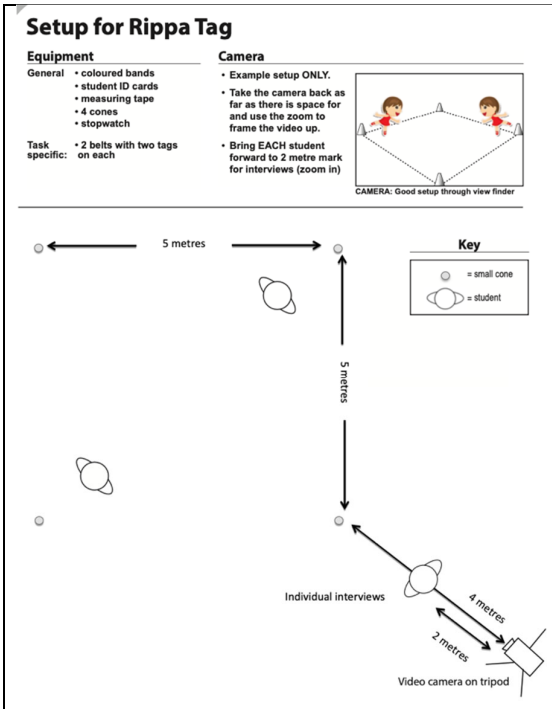
	Locomotion (Running, Dodging, Evading)	
		Working analysis terms
	High-range	Consistently good posture (leans in direction of desired movement, arms and legs in opposition) Technique consistently appropriate and quick (consistently on balls of feet, backwards and sideways movement) Movement dynamics consistently efficient, fluid and balanced
	Mid-range	Mostly good posture (occasionally on balls of feet when moving) Technique mostly appropriate and quick Movement dynamics mostly efficient, fluid and balanced (occasional extra movements – side stepping)
	Low-range	Generally poor posture Technique usually slow Movement dynamics not efficient, fluid or balanced (jerky movement, frequently overbalances)
	Insufficient	Insufficient movement displayed

Example of LTM task

The LTM assessment comprised six games. Eight students per school participated in different combinations of two games each in two groups of four. Each of the players wore a colour band for later identification. They were scored separately on their movement and strategising skills using a 0- 3 scale for most items.

Task: *Rippa tag.*

The following example shows the layout and dimensions for one of the LTM assessment tasks. This game was used in 2013, 2017 and again in 2022. The example shows the set up for the game, the optimal cameral position for videotaping; it includes a description of the game, and the marking rubric.



Rippa Tag is a tag game played by pairs in a restricted area.

In the game, the students were positioned diagonally in opposite corners of a square, marked out using spots. Each student wore a belt with Velcro-ed material tags. The object of the game was for one of the students to remove the tags from the other, whose job it was to avoid that by moving away or around in as many ways as they could. Holding on to the tags was not a valid strategy.

The students took turns at being the antagonist. They were given multiple opportunities to play, with time between to consider specific strategies they might employ.

In a short interview after the game, students were asked to explain the strategies they used in the game when attacking and defending, and to identify and evaluate the strategies used by their opponent.

The students were videoed as they competed in the game, and participated in the interview.

The students were assessed on two different aspects: their movement skills; and their ability to describe the use of strategy in the game. The tables that follow show how the students achieved on the two components.

Figure A8.6 Layout of the Rippa Tag game

Curriculum Strands: Movement Concepts and Motor Skills (Locomotion, Strategies and Tactics)

Part one of the LTM task Rippa tag (Year 4 and 8)

Item 1. Play the game	
Component: Locomotion (Running, Dodging/ Evading)	
Scoring category	**Examples**
0: Student displays insufficient movement	Student does not participate actively in the game
1: Displays low range movements	Infrequently leans in the direction of desired movement Jerky or stiff movements, overbalances, fixed arms Flat footed when moving – heels
2: Displays mainly mid-range movements	Occasionally on balls of feet when moving Occasionally leans in direction of desired movement, arms and legs mostly in opposition Movements usually fluid and quick, occasional extra movement – sidestepping
3: Student displays all/almost all high-range movements	Consistently moves on balls of feet (toes/mid foot strike pattern when moving) Consistently leads with leg closest to direction headed, arms legs in opposition Movements consistently fluid/ looks balanced, movements efficient Demonstrates a variety of avoidance or attacking movements

Figure A8.7 Item 1 of the LTM task Rippa Tag (Year 4 and 8)

Item 2. What strategies did you use when you were trying to grab the tags? Did they work? Why/Why not? **What strategies did you use to stop the other person grabbing your tags? Did they work? Why/why not?** **What strategies did the other person (your opponent) use? .**	
Component: Strategies and tactics	
Scoring category	**Example responses**
0: Not able to identify a strategy with respect to game play	"I didn't get the tags."
1: Able to identify general strategies (own actions) No or limited justification	Run fast; dodge; side step; twist "I lost"; "The tags were easy to get off"; "They were slower."
2: Able to identify ONE specific strategy with respect to game play, and justify clearly why it did or didn't work (Corner opposition/ blocking/ directional changes/ outrun)	"I kept spinning so that it makes it hard for the opponent to grab the tag"; "I confused them by dodging"
3: Able to identify ONE specific game strategy and evaluate effectiveness, AND identify ONE specific strategy the opposition has used	"After she grabbed one tag I changed direction quickly and she couldn't grab the other one – I kept my side where the tag was away from her. It worked because she couldn't get close enough. She tried to surprise me by going around the other way but I was faster and more agile."

Figure A8.8 Item 2 of the LTM task Rippa Tag (Year 4 and 8)

NMSSA • CYCLE 2

NMSSA Report 32 : TECHNICAL INFORMATION 2022 - Mathematics and Statistics; Health and Physical Education

Te Mahau

Te Tāhuhu o te Mātauranga
Ministry of Education

University of Otago
Te Whare Wānanga o Otāgo
NEW ZEALAND

NZCER
Rangahau Mātauranga o Aotearoa

ISSN: 2350-3238 (Online)

978-1-927286-83-8 (Online)