

Wānangatia te Putanga Taurā
National Monitoring Study
of Student Achievement

Health and Physical Education

Achievement Findings 2022

NMSSA • CYCLE 2

NMSSA Report 28: HEALTH AND PHYSICAL EDUCATION 2022 – ACHIEVEMENT FINDINGS

CYCLE 2
NMSSA Report 28

Wānangatia te Putanga Tauira
National Monitoring Study
of Student Achievement

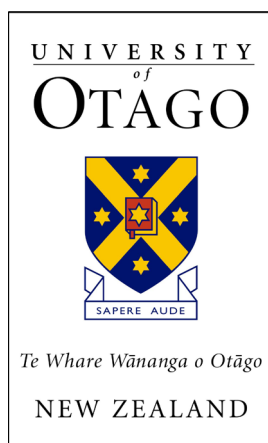
Health and Physical Education 2022

Achievement Findings

Educational Assessment Research Unit
and
New Zealand Council for Educational Research



© 2023 Ministry of Education, New Zealand



Key reports for Health and Physical Education 2022

(all available online at <http://nmssa.otago.ac.nz/reports-and-resources>)

- 28: Health and Physical Education 2022 – Achievement Findings
- 29: Health and Physical Education 2022 – Contextual Findings
- 32: Technical Information 2022



National Monitoring Study of Student Achievement Report 28: Health and Physical Education 2022 – Achievement Findings
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: (Print) 2350-3254
ISSN: (Online) 2350-3238

ISBN: (Print) 978-1-927286-75-3
ISBN: (Online) 978-1-927286-76-0

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

Acknowledgements	5
Student Achievement in Health and Physical Education	1
1. The 2022 study	1
2. Changes in achievement from 2013 to 2022	2
3. Overall achievement on the 2022 CT assessment	5
4. Achievement against the curriculum on the CT assessment	6
5. Achievement on the CT assessment by student-level variables	7
6. Achievement by school-level variables on the CT assessment	8
Appendix 1: The scale description	9
Appendix 2: Summary statistics	11

Appendices Tables:

Achievement in the HPE learning area	13
Table A2.1. Achievement on the CT scale: Summary statistics for Year 4 students	13
Table A2.2. Achievement on the CT scale: Summary statistics for Year 8 students	14
Table A2.3. Curriculum levels for CT: Year 4 and Year 8 students	15

Health & PE Project Team – 2022	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 curriculum advisory panel and reference groups for health and physical education
- 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.

Student Achievement in Health and Physical Education

In 2022, the National Monitoring Study of Student Achievement (NMSSA) assessed achievement in the health and physical education (HPE) learning area of the New Zealand Curriculum. Two assessments were used: one focussed on critical thinking in health and physical education and the other on learning through movement. This document summarises the key achievement findings from the two assessments. Within this summary, any reported differences in achievement between groups are statistically significant unless stated otherwise¹. Tables of results related to the reporting in this document are available in Appendix 2. Statistical information can also be found on a data window available on the NMSSA website (<https://nmssa.otago.ac.nz>).

1. The 2022 study

The 2022 NMSSA HPE study involved two nationally representative samples of students at English medium state and state-integrated schools—a Year 4 sample and a Year 8 sample. Each sample was made up of about 1,800 students from 100 schools.

The samples were constructed using a two-stage sampling design. In the first stage, a stratified random sampling approach based on school decile, geographical region, and school size was used to select 100 schools at each year level. In the second stage, up to 27 students were randomly selected from each school at the appropriate year level to be assessed. While all selected students were involved in the overall assessment programme in each school, up to 20 participated in the critical thinking in HPE assessment and a subset of these students also took part in the learning through movement assessment. Table 1 shows the number of students involved in each assessment, by year level.

Table 1: Number of schools and students involved in the 2022 NMSSA health and physical education study

Year level	Critical thinking in HPE (CT)	Learning through movement (LTM)
4	1850	621
8	1826	623

The 2022 HPE assessment programme

As noted above, the 2022 HPE assessment programme involved two different assessments: the Critical Thinking in HPE (CT) assessment and the Learning Through Movement (LTM) assessment. Teacher Assessors (TAs) administered both assessments as part of a two-and-a-half-day visit to each school in Term 3, 2022². Questionnaires were also used as part of the 2022 assessment programme to collect contextual data related to teaching and learning HPE from students, teachers, and school leaders. Analysis of data collected using the questionnaires is reported separately.

Critical thinking in HPE

The CT assessment focussed on three important aspects of learning in HPE: 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. Students recorded their answers in writing using an answer booklet. The second part of the assessment involved a

¹ Some plots use error bars to indicate a 95 percent confidence interval around a plotted statistic, such as a group mean. The intervals provide a range within which we can be fairly sure the population value for the reported statistic lies. The confidence intervals have been adjusted (widened) to account for any design effect associated with NMSSA's sampling approach (i.e. sampling schools and then sampling students).

² The visit also included the administration of assessments in the mathematics and statistics learning area².

series of tasks incorporated into one-to-one interviews with a teacher assessor. Some of the tasks contained material that was used at both Year 4 and Year 8.

Item Response Theory (IRT) was used to locate achievement on the CT assessment in 2022 on the same measurement scale developed for use in NMSSA assessments of HPE in 2013 and 2017 (the CT scale). The inclusion of tasks that had been used in previous years made this possible.

Appendix 1 provides a description of the knowledge and skills associated with different score ranges on the CT scale. Readers are encouraged to refer to this when considering the meaning of the CT scale scores that are provided throughout the report.

Learning through movement

The LTM assessment was designed to assess students' ability to use and show an understanding of movement skills, movement sequences, and strategy. The assessment was made up of two types of tasks. The first type was completed in pairs or groups of four and involved game-type situations that required the students to make use of a range of movement skills, sequences, and strategic decisions. The second type of tasks involved follow-up activities administered in one-to-one interviews between the students and a teacher assessor. During the interviews, students were asked to do things such as: identify and evaluate their own or opponents' strategies and suggest adaptations to the game activities to make them more inclusive. All the student responses to the LTM tasks were videoed. The videos were scored by trained markers using carefully prepared coding schedules.

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 based 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.

2. Changes in achievement from 2013 to 2022

Changes in Critical thinking in Health and PE

As noted above, NMSSA has used the CT assessment to assess achievement in HPE in 2013, 2017, and 2022. The use of a measurement scale linked through common items used in each cycle of assessment makes it possible to compare achievement over that time. Figure 1 shows how the average scale scores have changed at Year 4 and 8 on the CT scale between 2013 and 2022.

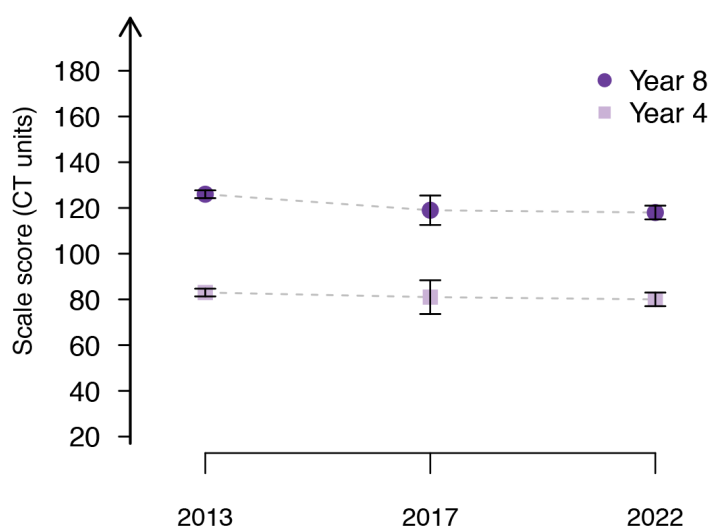


Figure 1: Change in average achievement in Critical Thinking in HPE from 2013 to 2022

Levels of achievement in 2022 on the CT assessment were similar to what they were in 2017

Any change in the average score on the CT assessment between 2017 and 2022 was not statistically significant at Year 4 or Year 8. Between 2013 and 2017 there was a statistically significant decline in the average scale score for students in Year 8.

There were no statistically significant changes in the average CT scores by gender, ethnic group, or school decile band between 2018 and 2022.

Table 2 shows the changes in average scale score on the HPE and statistics scale between 2017 and 2022 for all students and across gender, ethnic, and decile band³ groupings at Year 4 and Year 8. None of these changes were statistically significant.

Table 2: Change in average score on the CT scale between 2017 and 2022 by year level, gender, ethnic group, and school decile band

	Year 4		Year 8		
	Difference in average scores 2017 to 2022 (CT units)*	95 percent confidence interval (CT units)	Difference in average scores 2017 to 2022 (CT units)*	95 percent confidence interval (CT units)	
All students	-1.0	(-4.3, 2.2)	All students	-1.1	(-4.3, 2.2)
Girls	-1.5	(-4.9, 2.5)	Girls	-1.2	(3.7, -4.9)
Boys	-1.0	(-4.5, 2.9)	Boys	-0.8	(3.7, -4.5)
NZE	-2.4	(-5.9, 1.2)	NZE	-1.4	(-4.9, 2.1)
Māori	-2.2	(-6.7, 2.4)	Māori	-1.6	(-6.1, 2.9)
Pacific	4.1	(-1.9, 10.0)	Pacific	-2.9	(-8.9, 3.1)
Asian	2.3	(-3.0, 7.5)	Asian	3.2	(-2.5, 8.8)
Low decile	-2.2	(-6.8, 2.4)	Low decile	-3.4	(-8.2, 1.5)
Mid decile	0.4	(-3.4, 4.2)	Mid decile	1.0	(-2.7, 4.7)
High decile	-2.9	(-6.7, 0.9)	High decile	0.1	(-3.7, 3.9)

* Bolded numbers indicate the difference is statistically significant ($p < 0.05$)

Changes in Learning through movement

Figures 2 and 3 use dumbbell plots to show the change in average percentage scores on questions used as part of the LTM assessment in both 2017 and 2022. In the plots, the green dots are used to show the average percentage scores on the questions in 2022 and the blue dots to show the average percentage scores in 2017. As an example, the task Stepping Patterns shown at the top of each plot incorporated four questions used in 2022 and 2017. The figure for Year 4 shows that the average percentage score on two of these decreased between 2017 and 2022 and the other two increased. For three of the four questions, the size of the change was relatively small.

³ When the NMSSA samples were drawn for the 2022 study, deciles were used as a stratifying variable. They have therefore been used for reporting purposes. The low decile band comprised students in decile 1 to decile 3 schools, the mid band comprised students in decile 4 to decile 7 schools, and the high band comprised students in decile 8 to decile 10 schools.

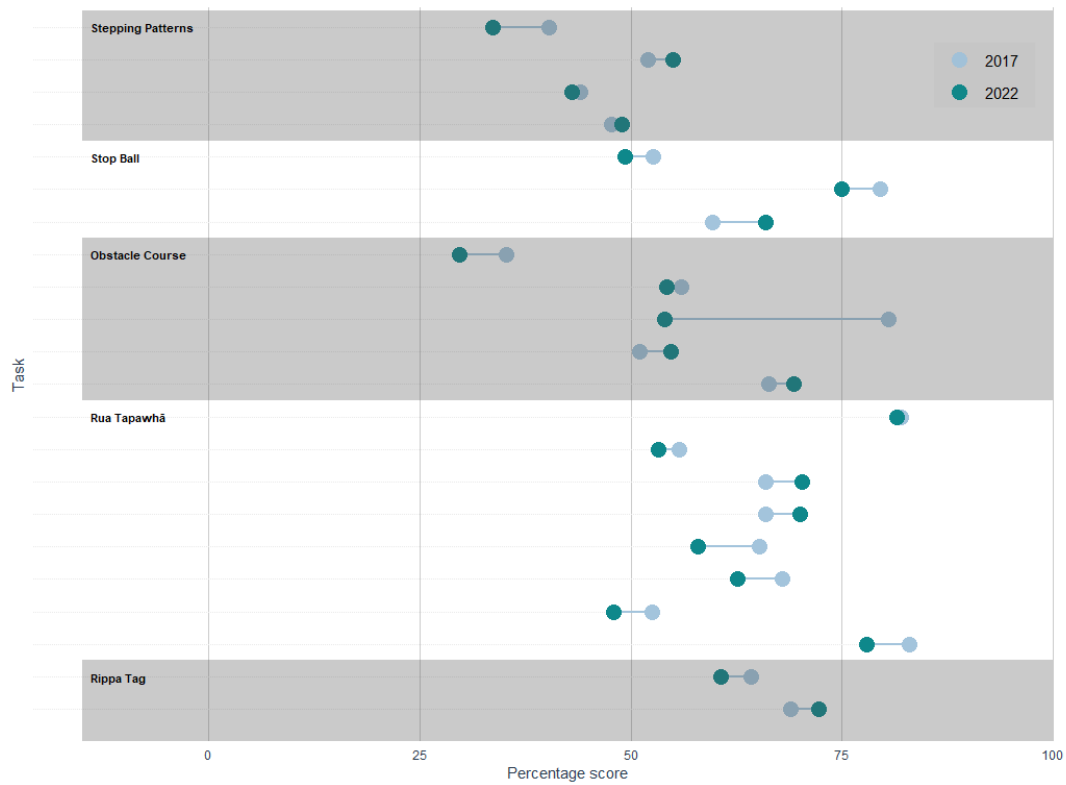


Figure 2: Average percentage scores for Year 4 students on common items from the Learning Through Movement (LTM) assessment used in 2017 and 2022

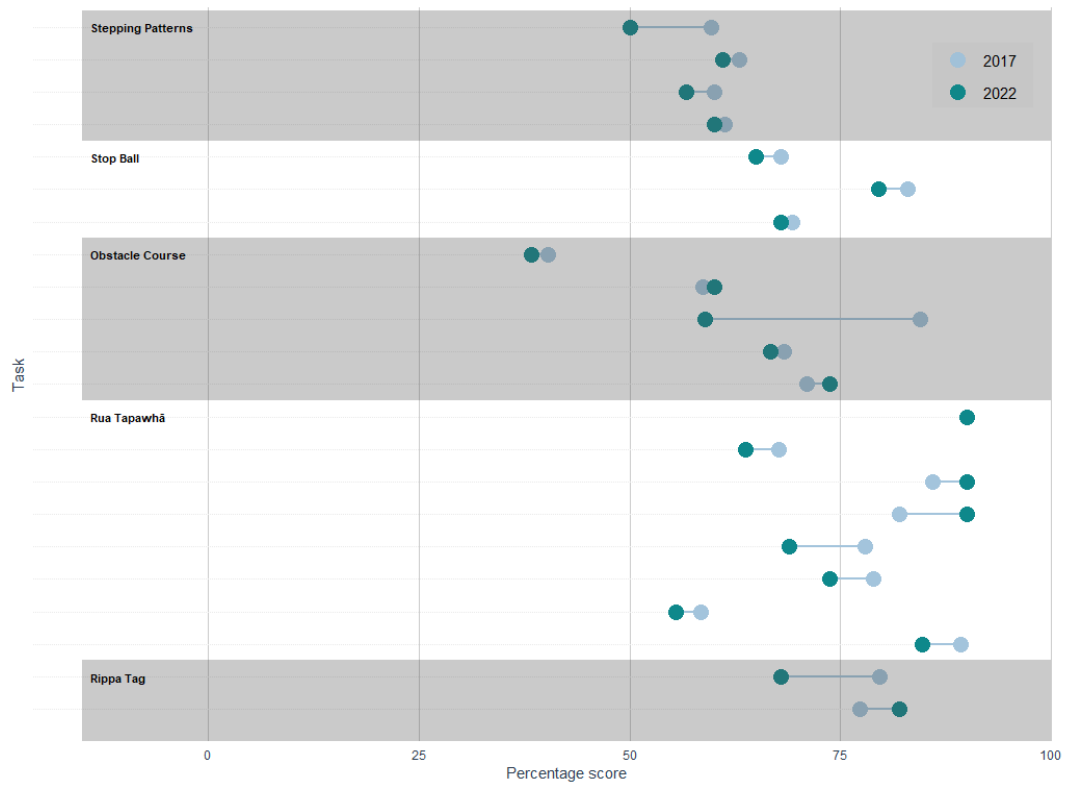


Figure 3: Average percentage scores for Year 8 students on common items from the Learning Through Movement (LTM) assessment used in 2017 and 2022

Overall, there was little change in achievement on questions used in both 2017 and 2022 from the Learning Through Movement assessment

The scoring patterns on the questions from the LTM assessment that were used in both 2017 and 2022 were similar at both time points. At both Year 4 and 8, the median change in the average percentage score for each question was a decrease of about 2 percentage points. Very few of the changes were statistically significant. One question from the Obstacle Course task did show a large decrease in the average percentage score at both year levels.

3. Overall achievement on the 2022 CT assessment

Figures 4 and 5 show the distribution of achievement on the CT scale for Year 4 and Year 8 in 2022, respectively. The difference in the average scores for the two year levels was 38 scale score units. This represents an annualised difference of 9.5 units and equates to an effect size of 0.45. The annualised difference indicates the amount of ‘progress’ associated with one year of schooling and can be helpful when interpreting score differences on the CT scale between groups.

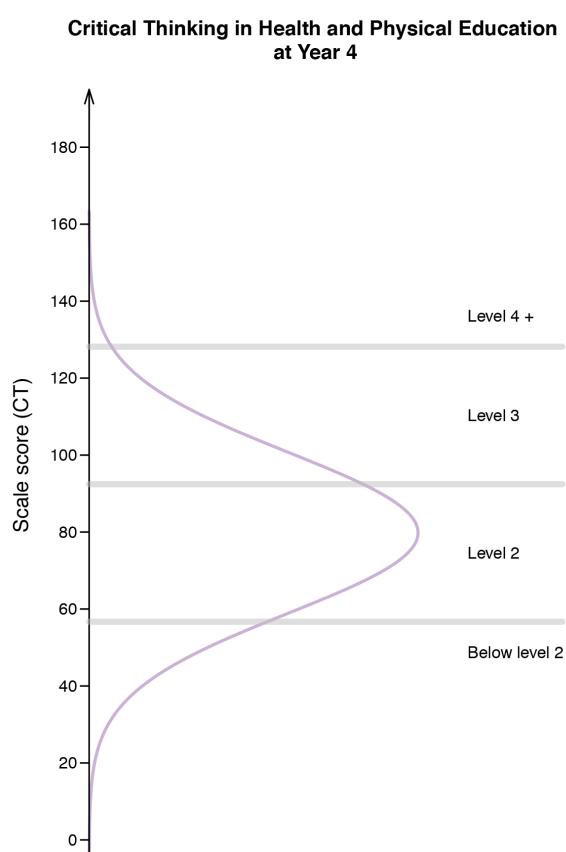


Figure 4: Distribution of scores for Year 4 students on the CT assessment

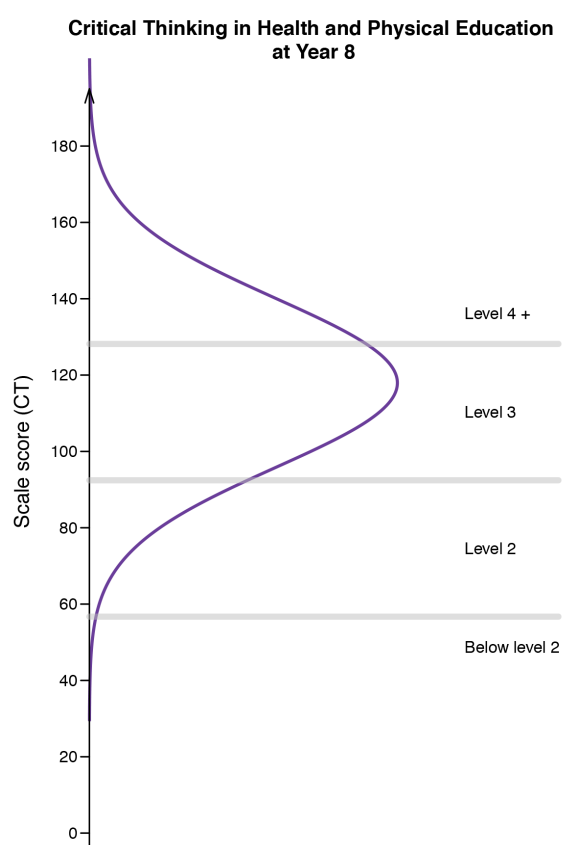


Figure 5: Distribution of scores for Year 8 students on the CT assessment

4. Achievement against the curriculum on the CT assessment

In 2013 a curriculum alignment exercise identified score ranges on the CT scale associated with meeting achievement expectations at different curriculum levels. Because achievement for 2022 has been located using the same scale, data from the 2022 CT assessment can also be interpreted against curriculum expectations (see Figures 4 and 5). At the end of Year 4, the curriculum expectation is for most students to have achieved at level 2. At the end of Year 8, most students are expected to have achieved at level 4. The expected score gain between the level 2 and level 4 cut points is 71 scale score units.

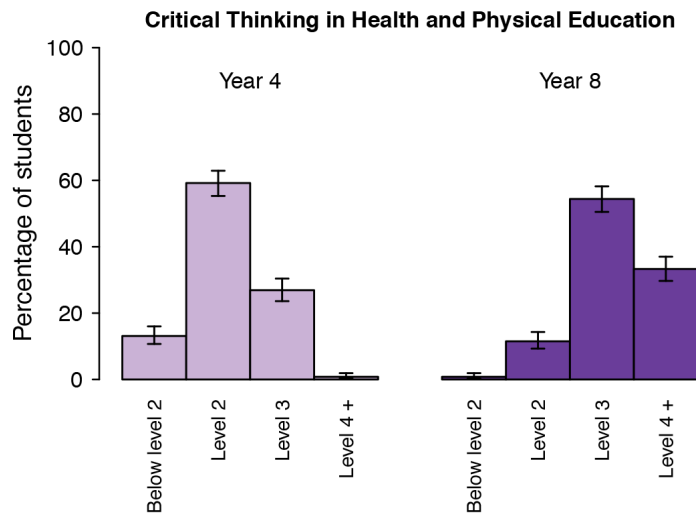


Figure 6: Achievement against the curriculum on the CT assessment, by year level

Fewer Year 8 students are meeting curriculum expectations in HPE compared with Year 4 students.

On the CT assessment, the proportion of students meeting curriculum expectations was markedly lower at Year 8 than at Year 4. At Year 4, 87 percent of students achieved at or above the minimum score on the CT scale associated with achieving at curriculum level 2. At Year 8, 33 percent of students achieved at or above the minimum score associated with achieving at level 4.

5. Achievement on the CT assessment by student-level variables

Figures 7 and 8 show the score distributions for all students, and by gender and ethnic group on the CT scale at Year 4 and Year 8, respectively.

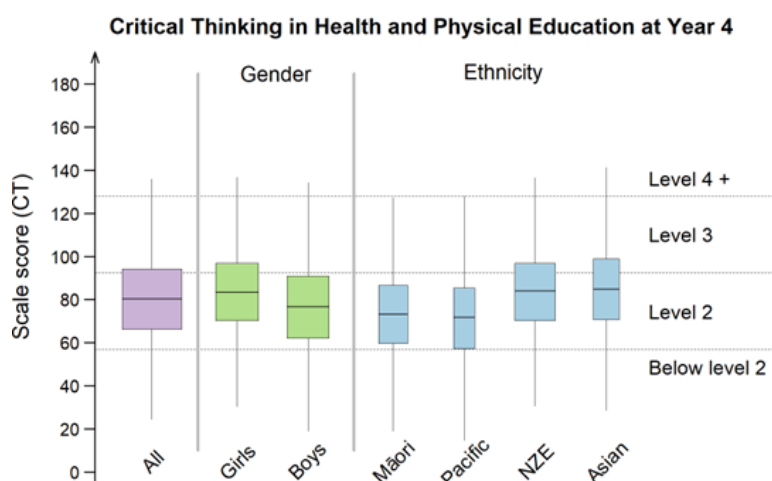


Figure 7: Distribution of scores for Year 4 students on the Critical Thinking in HPE (CT) scale, by gender and ethnicity (NZE = New Zealand European)

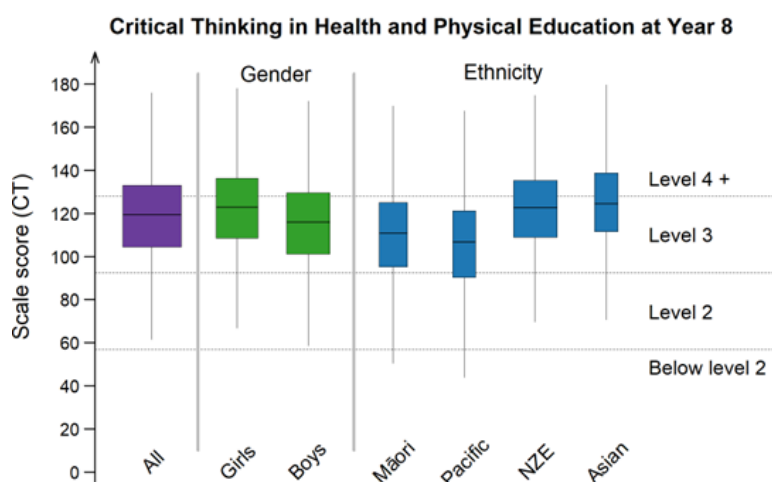


Figure 8: Distribution of scores for Year 8 students on the Critical Thinking in HPE (CT) scale, by gender and ethnicity (NZE = New Zealand European)

There were gender differences associated with achievement.

On average, girls scored higher than boys at both year levels on the CT assessment. The difference was about 7 scale score units at both year levels.

There were differences in achievement related to ethnic group.

On average, non-Māori students scored higher than Māori students and non-Pacific students scored higher than Pacific students on the CT assessment at both year levels. The greatest difference was between Pacific and non-Pacific learners at Year 8 (15 scale score units). The other differences ranged from 8 to 11 scale score units.

Asian students scored higher than non-Asian students on the CT assessments at Year 4 and Year 8 (by 6 and 8 units, respectively).

6. Achievement by school-level variables on the CT assessment

Figures 9 and 10 show the achievement of students according to school decile band and school type for the CT assessment for Year 4 and Year 8 students, respectively.

Students attending high decile schools scored higher, on average, than those attending mid or low decile schools.

At both year levels, students from high decile schools scored higher, on average, than those from mid and low decile schools. These differences were larger at Year 8 than at Year 4, with the difference in achievement at Year 8 between high and low decile schools representing an effect size greater than 1.

Students attending full primary schools achieved higher, on average, than those who attended intermediate schools

Students from full primary schools achieved higher, on average, than students who were attending intermediate schools. The difference represented an effect size of about 0.15.

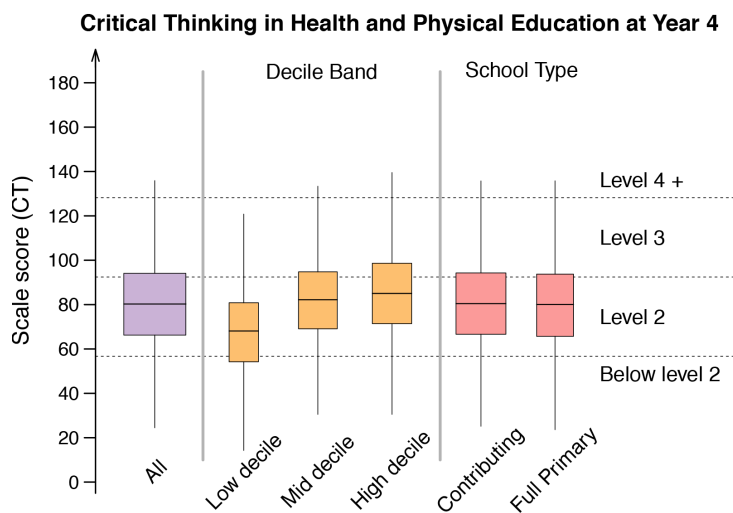


Figure 9: Distribution of scores for Year 4 students on the CT scale, by school decile band and school type

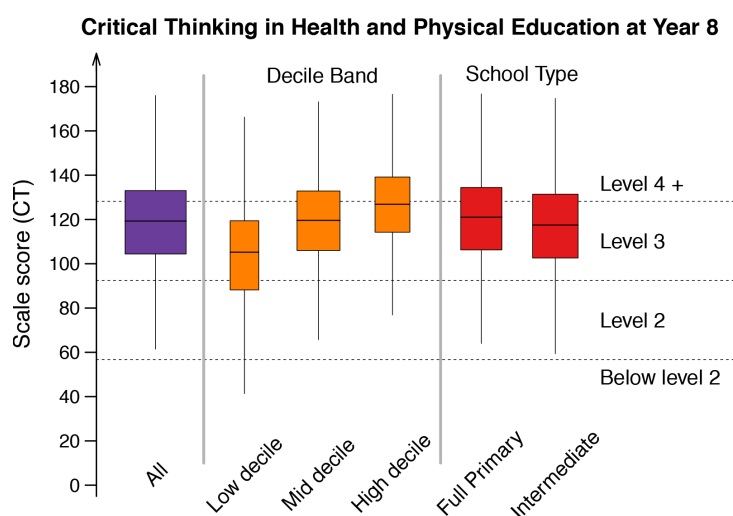


Figure 10: Distribution of scores for Year 8 students on the CT scale, by school decile band and school type

Appendix 1: The scale description

Figure A1.1. provides a description of the HPE knowledge and skills measured by the CT scale. In both figures, the descriptors provide examples of the kinds of skills that students, who achieved at different parts of the scale, were typically able to display⁴.

⁴ The descriptors located at each part of the scale were generated by examining the kinds of questions that students with achievement levels located at the same parts of the scale were able to answer successfully about 70 percent of the time.

NMSSA Critical Thinking in Health and Physical Education (CT) Scale

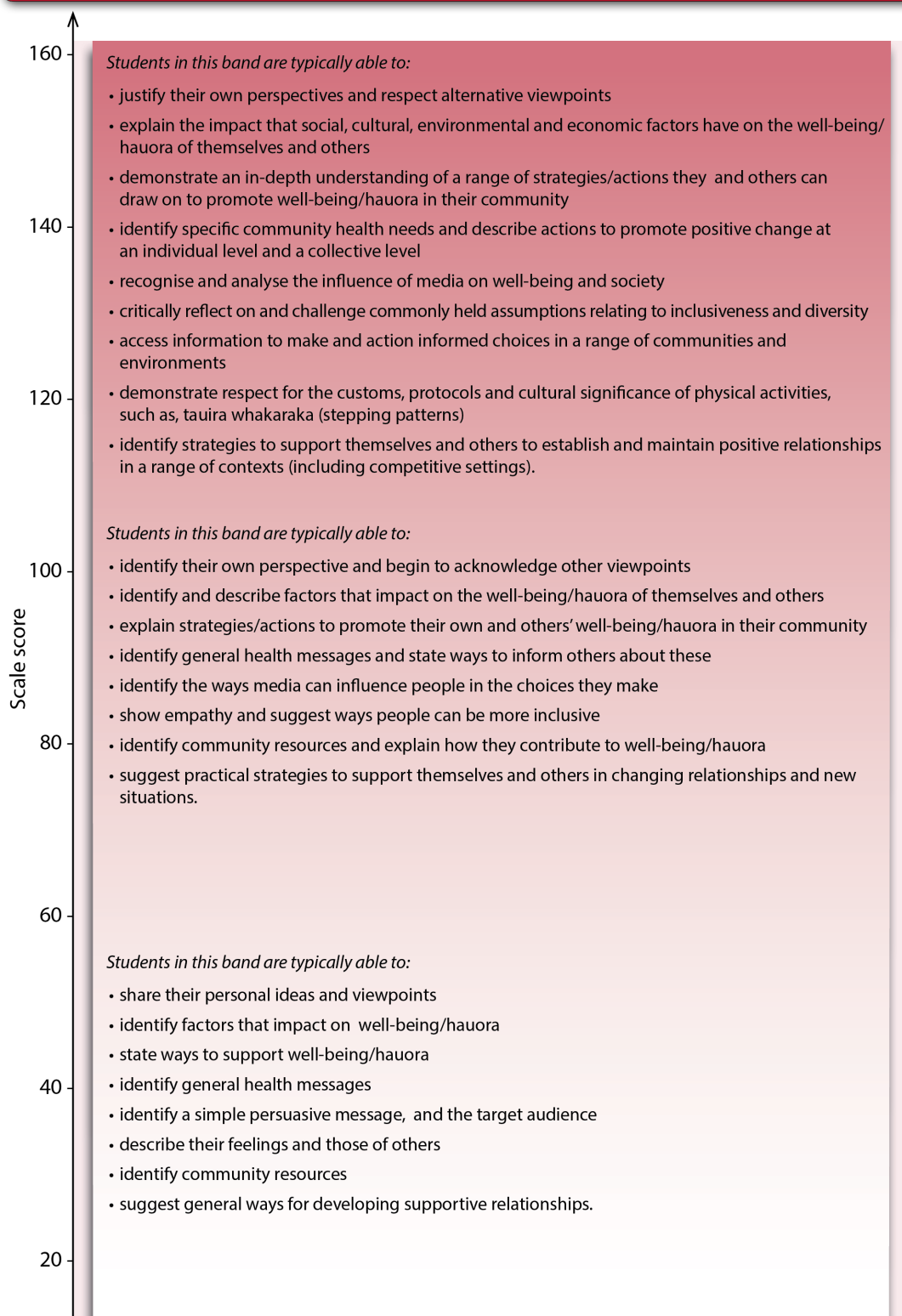


Figure A1.1. Description of the Critical Thinking in Health and Physical Education (CT) scale

Appendix 2: Summary statistics

Tables:

Achievement in the HPE learning area	13
Table A2.1. Achievement on the CT scale: Summary statistics for Year 4 students	13
Table A2.2. Achievement on the CT scale: Summary statistics for Year 8 students	14
Table A2.3. Curriculum levels for CT: Year 4 and Year 8 students	15

Reporting of statistics

The following tables report a range of statistics associated with the 2022 NMSSA HPE study.

95 percent confidence intervals

The tables show the 95 percent confidence intervals associated with the mean scores and percentages of students scoring at or above curriculum benchmarks reported in the tables. The intervals provide a range within which we can be fairly sure the population value for the reported statistic lies. The confidence intervals have been adjusted (widened) to account for any design effect associated with NMSSA's sampling approach (i.e. sampling schools and then sampling students).

Decile band

For the purposes of this report, the low decile band comprised students in decile 1 to decile 3 schools, the mid band comprised students in decile 4 to decile 7 schools, and the high band comprised students in decile 8 to decile 10 schools.

School type

A composite school combines students from different year levels that are typically found in separate primary or secondary schools. A restricted composite, sometimes known as a middle school, caters for Years 7 to 10. A contributing school caters for Years 1 to 6 of schooling. A full primary school caters for Years 1 to 8 of schooling. Secondary schools cater for Years 7 to 15 of schooling, although many cater for Years 9 to 15 only. An intermediate school caters for Years 7 and 8 of schooling. Please note, the numbers of students in the study representing school types other than those classified as either contributing, full primary or intermediate are low. Care should be taken when interpreting the statistics for these school types.

Achievement in the HPE learning area

Table A2.1. Achievement on the CT scale: Summary statistics for Year 4 students

Group	Sample size	Mean	Confidence interval for the mean	Standard deviation
All	1850	79.8	(78.7, 80.9)	20.9
Gender				
Girls	952	83.4	(81.9, 84.9)	19.8
Boys	898	76	(74.3, 77.7)	21.3
Ethnicity				
Māori	421	72.8	(70.4, 75.2)	21.1
Pacific	229	71.2	(67.9, 74.5)	21.4
Asian	366	84.6	(82.1, 87.1)	20.7
NZE	866	83.4	(81.8, 85.0)	19.5
Decile band				
Low decile	403	67.4	(65.0, 69.8)	20.4
Mid decile	722	81.7	(80.0, 83.4)	19
High decile	725	84.8	(83.0, 86.6)	20.2
School type				
Contributing	1188	80	(78.6, 81.4)	20.7
Full primary	649	79.4	(77.4, 81.4)	21.2
Composite (Year 1-15)	13	80.6	(62.6, 98.6)	23.6

Table A2.2. Achievement on the CT scale: Summary statistics for Year 8 students

Group	Sample size	Mean	Confidence interval for the mean	Standard deviation
All	1826	118	(116.8, 119.2)	22.1
Gender				
Girls	891	121.5	(119.8, 123.2)	22
Boys	935	114.6	(112.9, 116.3)	21.6
Ethnicity				
Māori	405	109.4	(106.8, 112.0)	21.9
Pacific	263	105.1	(101.8, 108.4)	22.8
Asian	272	124.7	(121.8, 127.6)	20.5
NZE	973	121.6	(120.1, 123.1)	20.3
Decile band				
Low decile	385	103.2	(100.5, 105.9)	22.9
Mid decile	876	119	(117.4, 120.6)	20.3
High decile	565	126.3	(124.4, 128.2)	19
School type				
Composite	5	120.6	(79.7, 161.5)	17.8
Full primary	790	119.6	(117.7, 121.5)	22.2
Intermediate	973	116.2	(114.5, 117.9)	22
Restricted composite	19	120.7	(108.5, 132.9)	20.5
Secondary	39	126.7	(119.3, 134.1)	18.9

Table A2.3. Curriculum levels for CT: Year 4 and Year 8 students

Group	Year 4			Year 8		
	Sample size	Level 2+ (%)	Confidence interval	Sample size	Level 4+ (%)	Confidence interval
All	1850	86.9%	(84.9, 88.6)	1826	33.3%	(30.7, 35.9)
Gender						
Girls	952	91.4%	(89.0, 93.3)	891	39.5%	(35.8, 43.4)
Boys	898	82.1%	(78.9, 84.9)	935	27.3%	(24.0, 30.9)
Ethnicity						
Māori	421	78.8%	(73.8, 83.1)	405	20.4%	(16.1, 25.5)
Pacific	229	75.4%	(68.2, 81.4)	263	15.5%	(10.9, 21.4)
Asian	366	91.0%	(86.8, 93.9)	272	43.2%	(36.4, 50.3)
NZE	866	91.3%	(88.8, 93.3)	973	38.8%	(35.2, 42.5)
Decile band						
Low decile	403	71.1%	(65.6, 76.1)	385	12.8%	(9.3, 17.3)
Mid decile	722	90.6%	(87.8, 92.9)	876	33.3%	(29.6, 37.1)
High decile	725	91.9%	(89.2, 94.0)	565	47.3%	(42.4, 52.2)
School type						
Contributing	1188	87.3%	(84.8, 89.4)			
Full primary	649	86.2%	(82.7, 89.1)	790	36.1%	(32.2, 40.2)
Composite (Year 1-15)	13	84.5%	(52.0, 96.5)	5	31.6%	(6.2, 76.2)
Intermediate				973	30.4%	(27.1, 34.0)
Restricted composite				19	37.7%	(17.3, 63.6)
Secondary (Year 7-15)				39	46.3%	(29.2, 64.2)

