Statistical Literacy over a Decade

Jane M. Watson
University of Tasmania
<Jane.Watson@utas.edu.au>

Ben A. Kelly
University of Tasmania
<Ben.Kelly@utas.edu.au>

John F. Izard
RMIT University
<john.izard@rmit.edu.au>

This study uses Rasch modelling to link student outcomes over the decade since the introduction of chance and data into the curriculum of an Australian state in 1993. Although improvement is observed over time for intact groups of students, and between grade levels in a given year, improvement across cohorts for given grades over time is not observed. The distribution of the items used in the 2003 survey across the statistical literacy variable supports earlier models of the hierarchical nature of statistical thinking obtained from a larger pool of items.

Three factors contributed to the research reported in this study. First was the introduction of chance and data into the mathematics curriculum in an Australian state in 1993, following its appearance in *A National Statement on Mathematics for Australian Schools* in 1991 (Australian Education Council [AEC], 1991; Department of Education and the Arts, 1993). Because very little research had contributed to the definition of the curriculum, research was begun into student understanding of the topics in the curriculum across the grades. This research was based on surveys and interviews, the surveys taking place in 1993, 1995, and 1997, and interviews in 1993 and 1997. A project involving teaching intervention (Watson & Kelly, 2004) made possible further survey work in different schools in 2000 and 2002, and finally in 2003 surveys were conducted again in the same schools that had participated in the 1990s.

The second factor contributing to this study was the interest in the development of student understanding over time. Because of the focus of the research on the middle school years, before the formal introduction of probability and statistics in the school curriculum, the construct of interest was conceived in terms of statistical literacy. This construct builds upon the elements of the chance and data curriculum with emphasis at higher levels on the abilities to interact with increasingly complex and less familiar contexts and to use proportional reasoning (Callingham & Watson, 2005; Watson & Callingham, 2003, 2004). The aims of statistical literacy by the time students leave school as stated by Gal (2002) are the

> (a) ability to interpret and critically evaluate statistical information, data-related arguments, or stochastic phenomena, which they may encounter in diverse contexts, and when relevant,
>
> (b) ability to discuss or communicate their reactions to such statistical information, such as their understanding of the meaning of the information, their opinions about the implications of this information, or their concerns regarding the acceptability of given conclusions. (pp. 2-3)

These are reflected in the items used in the initial surveys, with particular focus on topics from newspaper articles and students’ ability to communicate responses in language rather than numbers. The six levels of progression suggested by Watson and Callingham (2003) and supported with a different data set and related surveys by Watson and Callingham (2004) are summarised in Table 1. The items selected for the 2003 survey were chosen from a large pool and further confirmation of the statistical literacy construct could be sought.

Table 1
Levels of Statistical Literacy (Watson & Callingham, 2003)

1. Idiosyncratic	Tautologies, one-to-one counting, ability to read cell entries.
2. Informal	Intuitive non-statistical beliefs (3 is lucky), one-step calculations.
3. Inconsistent	Limited appreciation of content and context without justification; qualitative ideas.
4. Consistent Non-critical	Straight-forward engagement with context; means, simple probabilities and graphs.
5. Critical	Questioning engagement; appreciation of variation; qualitative interpretation of chance.
6. Critical Mathematical	Questioning critical engagement with context; proportional reasoning; subtle language.

The third factor supporting this study was the ability of Rasch analysis to place the students who completed surveys over the decade on a single scale. Although students in higher grades answered more items than students in lower grades, and after 1997 other items were added to the surveys, the presence of items in common across surveys allowed for links to be made and comparison across groups to occur. The presence of common students between the surveys in 1993, 1995, 1997, and 2003, allowed for longitudinal growth to be observed. Student understanding could hence be observed between grades for each cohort, across cohorts, and over time for some groups.

Research questions. Although many questions could be addressed based on the large data set, this study focuses on two. (1) Is the hierarchical nature of the statistical literacy construct first suggested by Watson and Callingham (2003) supported by analysis of the items used in the 2003 survey? (2) Using the data from one group of 13 schools in 1993, 1995, 1997, and 2003, what does Rasch analysis provide in the way of evidence for difference between grades, difference over time, and difference between cohorts?

Methodology

Sample. A total of 5263 student responses were used in the analyses reported here. The numbers for each grade and year are given in Table 2. The schools represented all geographical regions of the state and included rural and suburban schools, in many cases with a primary school linked to a local high school in order to follow students over time. All students completed surveys during class time (approximately 45 minutes) with teachers and members of the research team present. Younger children were sometimes assisted with reading questions but not with responses.

Analysis. The Rasch analyses (Masters, 1982; Rasch, 1960) that resulted in the data used in this study took place in three stages. The first stage, reported in Watson, Kelly, and Izard (2004) was based on data collected in 2000 from a different sample of students. The purpose of the initial analyses of data collected in 2000 was to establish anchor values for the items in common across years 1993 to 2003, so tests including these items could be calibrated on a common scale or continuum of achievement. There are a number of options in conducting analyses where each group of students attempted different subsets of items. Provided that each group attempts sufficient common items, and that these items are within the target achievement range, the simplest first option is to analyse the items in common. From this analysis an anchor file is created and used in the subsequent subsets to place remaining items on the common scale (as defined by the items in common). In 2000 there were 738 students attempting overlapping sets of items from a pool of 50 items: Grade 3 students attempted 24 items; Grade 5, 28 items; Grade 7, 45 items; and Grade 9, 46 items. This created anchor values for items in the pool.

Table 2

Sample Sizes for Rasch Analyses (numbers in parenthesis indicate students surveyed two or three times)

	1993	1995	1997	2003	Total
Grade 3	322 (147[1])	303	237 (54[2])	189	1051
Grade 5		465 (147[1])	226		691
Grade 6	311 (117[1])	337	234	174	1055
Grade 7			314 (147[1])		314
Grade 8		374 (117[1])	192		516
Grade 9	392 (117[2])	371 (51[2])	105	251 (54[2])	1119
Grade 10			297 (117[1])		297
Grade 11		118 (117[2])	51 (51[2])		169
Total	1025	1968	1656	614	5263

[1]Students surveyed three times. [2]Students surveyed twice.

The second stage in the analysis was of data collected in 2003. The survey in 2003 did not include all of the 2000 items. Twelve items had anchor values from the previous analyses of 2000 data. An initial analysis was conducted (without anchoring) on the 13 items common to all 614 data cases to check that these items-in-common were internally consistent. Subsequent analyses involved the 189 Grade 3 cases and anchor values for 9 items from the 2000 analyses, and 175 Grade 6 cases used anchor values for 8 items from the 2000 analyses and 4 items obtained from an earlier run. The analysis established anchor values for the remaining items in the 23 attempted by Grade 6. The analysis for the 251 Grade 9 cases used anchor values for items from the 2000 analyses and the anchor values for items obtained from earlier runs to establish anchor values for the remaining items in the 31 attempted by Grade 9. The statistics associated with output of the Rasch analysis were acceptable and are reported in Appendix A. From the 2003 data for 37 items anchored on 2000 results, 6 were deleted, leaving 31 items for further use in subsequent analyses including those for the 1993-1997 data.

The third stage in the analysis was of data collected in 1993, 1995, and 1997. The analyses were based on anchor values for 8 items available from previous analyses. The analysis established anchor values for the other items in the first 13. These anchor values were used to calibrate the first 24 items. Subsequent analyses used these anchor values and values for some later items available from intermediate analyses. Analyses of 40 items involved 902 cases.

From the 1993-1997 data for 40 items anchored on 2000 and 2003 results, 4 were deleted, leaving 36 items, 16 surviving from the previous 2003 analyses and 20 from the 1993-1995-1997 analyses for further use in subsequent analyses. The summary of statistics for the final run for the 1993-1997 data is given in Appendix B.

Because the item difficulties were now anchored it was possible to interpret scaled scores on subsets of these items on a common scale. These scaled scores were used in subsequent analyses to determine changes over time for individual students and groups where longitudinal data and cohort data by grade were available. Mean scores were compared for groups using Cohen’s (1969) effect size methodology and accompanied by

Cohen’s descriptors (Cohen, 1969; Izard, 2004). Positive differences reflect positive change.

Results

Research Question 1: Confirmation of the Statistical Literacy Construct

Thirty-one items were used in the final Rasch analysis that produced a variable map for the 2003 data set of 614 students across Grades 3, 6, and 9. The content of items was the same or similar to items presented in Watson and Callingham (2003) or Callingham and Watson (2005). Although there was some movement of items relative to each other, the overall structure was similar. The partial credit coding of responses resulted in 80 criteria for the 31 items, further detailing progress across the construct. Examples of the criteria for the tasks are presented.

Level 1 (Idiosyncratic) consisted of 16 partial credit item codes, reflecting the ability to read frequencies from a pictograph and cell values from a two-way table. Reasons for chance outcomes reflected idiosyncratic reasoning, such as “It’s the way I roll the die” or “6 is always easier to get than a 1.” For a media article, non-statistical beliefs were expressed such as “if students have guns everyone could get shot.” In attempting to read a stacked dot plot, responses used elements of the graph but inappropriately for the task set.

Level 2 (Informal) consisted of 12 partial credit item codes reflecting the ability to compare two values in a table, regard average as “normal,” recognise a qualitative colloquial interpretation of “15% chance” (e.g., “good” chance), and suggest that outcomes for a die depend on how it is thrown.

Level 3 (Inconsistent) contained 17 partial credit item codes reflecting understanding of the purpose of conducting a survey but an inability to detect inappropriate methods. For a task to draw a graph of the association of two variables, only a single aspect of the task (e.g., a single variable) was shown. Qualitative descriptions were given for dice outcomes.

Level 4 (Consistent non-critical) included 20 partial credit item codes reflecting appropriate appraisal of many situations where critical analysis was not required. These included the ability to show appropriate variation in predicting outcomes for 60 rolls of a 6-sided die and explaining the variation, the ability to order seven newspaper headlines involving chance appropriately on a 0-1 number line, and the ability to decide and justify that a scaled stacked dot plot tells a data story better than an unscaled plot.

Level 5 (Critical) included 8 partial credit item codes suggesting the ability to analyse contexts critically but without high level proportional reasoning. Likely outcomes were that quantitative values for simple chance events were given, a graph representing the association of two variables was successfully drawn, questions about research design considered cause and effect, an error in a pie chart was discovered, and an integrated definition of variation was provided.

At Level 6 (Critical mathematical) there were 7 partial credit item codes that reflected either more sophisticated mathematical reasoning or more subtlety of argument. Tasks involved using proportional reasoning to interpret a two-way data table, criticising non-appropriate methods of sample selection, recognising the possibility of outliers in choosing the median over the mean in a social context, and expressing uncertainty when reaching statistical decisions.

Research Question 2: Student Performance across Grades and over Time

The Rasch analysis put all students on the same scale with respect to statistical literacy. The measure employed was the logit, the logarithm of the odds of success. Table 3 provides means and other information to assess the effect size for each comparison of successive pairs of grades. As can be seen there are large differences in the three-year spreads for each of the four years (with one medium difference in 1997). These data for 12 different cohorts of students provide the backdrop for later comparisons.

Table 3
Comparison of Successive Grades in 1993, 1995, 1997, and 2003

	Grades 3 and 6	Grades 6 and 9
1993	*G3/6 Mean, SD*: -0.78, 0.67 / -0.14, 0.52 *Mean Diff, Effect Size, SE*: 0.64, 1.06, 0.08 *Descriptor*: *Large*	*G6/9 Mean, SD*: -0.14, 0.52 / 0.64, 0.61 *Mean Diff, Effect Size, SE*: 0.78, 1.36, 0.08 *Descriptor*: *Large*
1995	*G3/6 Mean, SD*: -0.86, 0.61 / -0.15, 0.55 *Mean Diff, Effect Size, SE*: 0.71, 1.23, 0.09 *Descriptor*: *Large*	*G6/9 Mean, SD*: -0.15, 0.55 / 0.32, 0.69 *Mean Diff, Effect Size, SE*: 0.47, 0.75, 0.08 *Descriptor*: *Large*
1997	*G3/6 Mean, SD*: -0.51, 0.76 / -0.13, 0.60 *Mean Diff, Effect Size, SE*: 0.38, 0.55, 0.09 *Descriptor*: *Medium*	*G6/9 Mean, SD*: -0.13, 0.60 / 0.35, 0.53 *Mean Diff, Effect Size, SE*: 0.48, 0.83, 0.12 *Descriptor*: *Large*
2003	*G3/6 Mean, SD*: -1.25, 0.74 / -0.47, 0.62 *Mean Diff, Effect Size, SE*: 0.78, 1.14, 0.11 *Descriptor*: *Large*	*G6/9 Mean, SD*: -0.47, 0.62 / 0.18, 0.58 *Mean Diff, Effect Size, SE*: 0.65, 1.09, 0.11 *Descriptor*: *Large*

For the subsets of Grade 3 and 6 students who were followed from 1993 to 1995 to 1997, in two-year grade jumps, Table 4 provides the information for making effect size comparisons and these are “medium” or “large.”

Table 4
Successive Two-year Comparison of Students Surveyed Longitudinally Twice

	1993-1995	1995-1997
Grade 3,5,7 (n = 147)	*93/95 Mean, SD*: -0.80, 0.66 / -0.25, 0.48 *Mean Diff, Effect Size, SE*: 0.55, 0.95, 0.12 *Descriptor*: *Large*	*95/97 Mean, SD*: -0.25, 0.48 / 0.14, 0.48 *Mean Diff, Effect Size, SE*: 0.39, 0.81, 0.12 *Descriptor*: *Large*
Grade 6,8,10 (n = 117)	*93/95 Mean, SD*: -0.10, 0.53 / 0.25, 0.54 *Mean Diff, Effect Size, SE*: 0.35, 0.65, 0.13 *Descriptor*: *Medium*	*95/97 Mean, SD*: 0.25, 0.54 / 0.73, 0.66 *Mean Diff, Effect Size, SE*: 0.48, 0.80, 0.14 *Descriptor*: *Large*

The effect size data provided in Table 5 considers Grade 9 two years after 1993 and 1995. In these cases the changes between Grades 9 and 11, although showing improvement over time, are only small.

Table 5
Comparisons for Grade 9 Students Surveyed Longitudinally Once

	1993-1995 (n = 117)	1995-1997 (n = 51)
Grade 9,11	*93/95 Mean, SD*: 0.84, 0.62 / 0.97, 0.65 *Mean Diff, Effect Size, SE*: -0.13, -0.20, 0.13 *Descriptor*: *Small*	*95/97 Mean, SD*: 0.67, 0.65 / 0.86, 0.80 *Mean Diff, Effect Size, SE*: -0.19, -0.26, 0.20 *Descriptor*: *Small*

Table 6 contains data for the single six-year comparison between 1997 and 2000 from Grades 3 to 9. From the information presented the change from Grades 3 to 9 was large.

Table 6

Comparison of Grade 3 Students Surveyed Longitudinally after Six Years, 1997-2003

Grade 3,9	*97/03 Mean, SD*: -0.28, 0.78 / 0.28, 0.62
(n = 54)	*Mean Diff, Effect Size, SE*: -0.56, -0.79, 0.20 *Descriptor*: *Large*

Using the data from Table 3, Table 7 provides information on positive or negative differences in favour of a subsequent year. In the 1990s there were two positive medium differences at Grade 3 level, but comparisons for 2003 showed three negative medium or large differences. At Grade 6 level, in the 1990s there were no positive medium differences, but comparisons for 2003 showed three negative medium differences. At the Grade 9 level in the 1990s there were two negative medium differences while comparisons for 2003 showed two negative large differences.

Table 7

Comparisons of Grade Cohorts over the Survey Years

	Grade 3	Grade 6	Grade 9
1993-1995	*Mean Difference*: -0.08 *Effect Size, SE*: -0.12, 0.08 *Descriptor*: *Very Small*	*Mean Difference*: -0.01 *Effect Size, SE*: -0.02, 0.08 *Descriptor*: *Very Small*	*Mean Difference*: -0.32 *Effect Size, SE*: -0.49, 0.07 *Descriptor*: *Medium*
1993-1997	*Mean Difference*: 0.27 *Effect Size, SE*: 0.38, 0.09 *Descriptor*: *Medium*	*Mean Difference*: 0.01 *Effect Size, SE*: 0.02, 0.09 *Descriptor*: *Very Small*	*Mean Difference*: -0.29 *Effect Size, SE*: -0.49, 0.11 *Descriptor*: *Medium*
1993-2003	*Mean Difference*: -0.47 *Effect Size, SE*: -0.67, 0.09 *Descriptor*: *Medium*	*Mean Difference*: -0.33 *Effect Size, SE*: -0.59, 0.10 *Descriptor*: *Medium*	*Mean Difference*: -0.46 *Effect Size, SE*: -0.77, 0.08 *Descriptor*: *Large*
1995-1997	*Mean Difference*: 0.35 *Effect Size, SE*: 0.51, 0.09 *Descriptor*: *Medium*	*Mean Difference*: 0.02 *Effect Size, SE*: 0.04, 0.09 *Descriptor*: *Very Small*	*Mean Difference*: 0.03 *Effect Size, SE*: 0.05, 0.11 *Descriptor*: *Very Small*
1995-2003	*Mean Difference*: -0.39 *Effect Size, SE*: -0.59, 0.09 *Descriptor*: *Medium*	*Mean Difference*: -0.32 *Effect Size, SE*: -0.56, 0.09 *Descriptor*: *Medium*	*Mean Difference*: -0.14 *Effect Size, SE*: -0.22, 0.08 *Descriptor*: *Small*
1997-2003	*Mean Difference*: -0.74 *Effect Size, SE*: -0.99, 0.10 *Descriptor*: *Large*	*Mean Difference*: -0.34 *Effect Size, SE*: -0.56, 0.10 *Descriptor*: *Medium*	*Mean Difference*: -0.17 *Effect Size, SE*: -0.30, 0.12 *Descriptor*: *Small*

Discussion

The hierarchical nature of the statistical literacy construct suggested by Watson and Callingham (2003) is supported by the ordering of the partial credit item codes in relation to the abilities of the 614 students in Grades 3, 6, and 9 who completed the statistical literacy survey in 2003. Although there were not as many items requiring proportional reasoning, there was still a number requiring critical thinking and mathematical subtlety at the highest level of the construct. Descriptions of the requirements of other levels agreed well with those of Watson and Callingham.

Where the same individuals have been re-tested in later years there is clear evidence of positive longitudinal change, particularly in the primary school classes. Differences observed across grades within each cohort show similar differences, suggesting that changes observed for different cohorts across grades represent more than differences between cohorts. The small improvement noted from Grades 9 to 11 (see Table 5) may be a result of students leaving school or changing schools in the intervening two years and perhaps choosing to focus on school subjects that were non-mathematically based. There was no attempt to follow mathematically talented students to Grade 11. Whereas all students were enrolled in a mathematics course at Grade 9, in Grade 11 mathematics was an optional subject.

In comparing the same grades in successive years when the survey was administered, the 1997 students overall demonstrated higher ability levels than their grade cohorts in the other years, with the 1993 students demonstrating the next highest ability levels comparatively. Although many of the comparisons across years were associated with very small effect sizes, the lowest level of performance overall was by students in 2003. Explanation for the drop in performance 10 years after the introduction of the Chance and Data curriculum is difficult to explain. Although the sample size in 2003 was somewhat smaller than earlier years, all of the original schools were represented. The research projects that collected the data presented here did not intervene in the schools over the decade and although there is anecdotal evidence of professional development within the state over the decade, there are no records of numbers of sessions or teachers attending. It may be that over the decade the emphasis on professional learning in this part of the curriculum diminished and that this resulted in the relatively poorer performance in 2003. Another related influence may be curriculum change taking place within the state in the years from 2000. The introduction of the *Essential Learnings Framework* (Department of Education, Tasmania, 2002) reduced emphasis on the discipline of mathematics, focussing instead on the Essential Elements of “Being Numerate” and “Inquiry.” This may have resulted in less emphasis being placed on Chance and Data by teachers feeling the need to address new curriculum issues. Particularly in the primary grades, a further emphasis on basic number sense and skills in the light of national benchmarking, may have further reduced the emphasis on Chance and Data.

Acknowledgement. This research was funded by the Australian Research Council, Discovery Grant No. DP0208607.

References

Australian Education Council. (1991). *A national statement on mathematics for Australian schools*. Melbourne: Author.

Callingham, R.A., & Watson, J.M. (2005). Measuring statistical literacy. *Journal of Applied Measurement*, *6*(1), 19-47.

Cohen, J. (1969). *Statistical power analysis for the behavioural sciences*. New York: Academic Press.

Department of Education and the Arts. (1993). *Mathematics guidelines K-8*. Hobart: Curriculum Services Branch.

Department of Education, Tasmania. (2002). *Essential learnings framework 1*. Hobart: Author.

Gal, I. (2002). Adults’ statistical literacy: Meanings, components, responsibilities. *International Statistical Review, 70*, 1-51.

Izard, J.F. (2004, March). *Best practice in assessment for learning*. Paper presented at the Third Conference of the Association of Commonwealth Examinations and Accreditation Bodies on Redefining the Roles of Educational Assessment, South Pacific Board for Educational Assessment, Nadi, Fiji.

Masters, G.N. (1982). A Rasch model for partial credit scoring. *Psychometrika*, *47*, 149-174.

Rasch, G. (1960). *Probabilistic models for some intelligence and attainment tests*. Copenhagen: Danish Institute for Educational Research.

Watson, J.M., & Callingham, R.A. (2003). Statistical literacy: A complex hierarchical construct. *Statistics Education Research Journal*, *2*(2), 3-46.

Watson, J.M., & Callingham, R.A. (2004, June). *Statistical literacy: From idiosyncratic to critical thinking*. Paper presented at the International Association for Statistics Education Roundtable on “Curricular Development in Statistics Education,” Lund, Sweden.

Watson, J.M., & Kelly, B.A. (2004). A two-year study of students' appreciation of variation in the chance and data curriculum. In I. Putt, R. Faragher & M. McLean (Eds.), *Mathematics education for the third millennium: Towards 2010* (Proceedings of the 27th Annual Conference of the Mathematics Education Research Group of Australasia, Townsville, Vol. 2, pp. 573-580). Sydney, NSW: MERGA.

Watson, J.M., Kelly, B.A., & Izard, J.F. (2004, December). *Student change in understanding of statistical variation after instruction and after two years: An application of Rasch analysis*. Refereed paper presented at the annual conference of the Australian Association for Research in Education. Available at http://www.aare.edu.au/04pap/alpha04.htm

Appendix A

Summary Results: 2003 data anchored on 2000 results

```
Items 1to37 2003 data (Run No 8)
Item Estimates (Thresholds) all on all      Case Estimates all on all
(N = 614 L = 31 Probability Level=0.50)     (N = 614 L = 31 Probability Level=0.50)
Summary of item Estimates                   Summary of case Estimates
Mean                          -0.41         Mean                          -0.42
SD                             1.29         SD                             0.89
SD (adjusted)                  1.29         SD (adjusted)                  0.83
Reliability of estimate        1.00         Reliability of estimate        0.87
Fit Statistics                              Fit Statistics
 Infit Mean Square      Outfit Mean Square   Infit Mean Square      Outfit Mean Square
    Mean    0.91           Mean    0.92         Mean    0.94           Mean    0.93
    SD      0.14           SD      0.21         SD      0.37           SD      0.58
      Infit t                Outfit t             Infit t                Outfit t
    Mean   -1.30           Mean   -0.74         Mean   -1.17           Mean   -0.09
    SD      1.79           SD      1.81         SD      1.03           SD      0.77
   0 items with zero scores                    0 case with zero scores
   0 items with perfect scores                 0 case with perfect scores
```

Appendix B. Summary Results: 1993-1997 data anchored on 2000 and 2003 data

```
JW Items 1-40 Grades 3-10 1993+ Initial Test only (Run 6)
Item Estimates (Thresholds) all on all      Case Estimates all on all
(N = 902 L = 36 Probability Level=0.50)     (N = 902 L = 36 Probability Level=0.50)
Summary of item Estimates                   Summary of case Estimates
Mean                           0.06         Mean                           0.47
SD                             1.23         SD                             0.63
SD (adjusted)                  1.22         SD (adjusted)                  0.57
Reliability of estimate        1.00         Reliability of estimate        0.83
Fit Statistics                              Fit Statistics
 Infit Mean Square      Outfit Mean Square   Infit Mean Square      Outfit Mean Square
    Mean    0.89           Mean    0.88         Mean    0.95           Mean    0.88
    SD      0.22           SD      0.24         SD      0.30           SD      0.35
      Infit t                Outfit t             Infit t                Outfit t
    Mean   -1.77           Mean   -1.27         Mean   -0.25           Mean   -0.18
    SD      3.71           SD      2.69         SD      1.16           SD      0.66
   0 items with zero scores                    0 case with zero scores
   0 items with perfect scores                 0 case with perfect scores
```