# General Education Science Assessment Report 2009 – 2012
*Revised 14 August 2012 with 2010-2011 data*

The general education science assessment has been performed for the past three academic years. Each fall a sample of laboratory reports was analyzed against a rubric.

| Performance factor score | | | |
|---|---|---|---|
| **4** | **3** | **2** | **1** |
| *Metric: Scientific Procedures and reasoning* | | | |
| Accurately and efficiently used all appropriate tools and technologies to gather and analyze data | Effectively used some appropriate tools and technologies to gather and analyze data with only minor errors | Attempted to use appropriate tools and technologies but information inaccurate or incomplete | Inappropriate use of tools or technology to gather data |
| *Metric: Strategies* | | | |
| Used a sophisticated strategy and revised strategy where appropriate to complete the task; employed refined and complex reasoning and demonstrated understanding of cause and effect; applied scientific method accurately | Used a strategy that led to completion while recording all data; used effective scientific reasoning; framed or used testable questions, conducted experiment, and supported results with data | Used a strategy that led to partial completion of the task/ investigation; some evidence of scientific reasoning used; attempted but could not completely carry out testing, recording all data and stating conclusions | No evidence of procedure or scientific reasoning used; so many errors, task could not be completed |
| *Metric: Scientific communication/using data* | | | |
| Provided clear, effective explanation detailing how the task was carried out; precisely and appropriately used multiple scientific representations and notations to organize and display information; interpretation of data supported conclusions and raised new questions or was applied to new contexts; disagreements with data resolved when appropriate | Presented a clear explanation; effectively used scientific representations and notations to organize and display information; appropriately used data to support conclusions | Incomplete explanation; attempted to use appropriate scientific representations and notations, but were incomplete; conclusions not supported or were only partly supported by data | Explanation could not be understood; inappropriate use of scientific notation; conclusion unstated or data unrecorded |
| *Metric: Scientific concepts and related content* | | | |
| Precisely and appropriately used scientific terminology; provided evidence of in-depth, sophisticated understanding of relevant scientific concepts, principles or theories; revised prior misconceptions when appropriate; observable characteristics and properties of objects, organisms, and/or materials used; went beyond the task investigation to make other connections or extend thinking | Appropriately used scientific terminology; provided evidence of understanding of relevant scientific concepts, principles or theories; evidence of understanding observable characteristics and properties of objects, organisms, and/or materials used | Used some relevant scientific terminology; minimal reference to relevant scientific concepts, principles or theories; evidence of understanding observable characteristics and properties of objects, organisms, and/or materials used | Inappropriate use of scientific terminology; inappropriate references to scientific concepts, principles or theories |

The rubric generates four scores on a scale from one to four. The maximum is sixteen. Two assessors evaluated each laboratory report. The scores for both assessors are combined yielding a total possible sum of 32 per laboratory report. Each of the four subsections has a maximum possible of eight (four plus four from each grader) per laboratory report.

Over the course of three academic years, 2009-2010, 2010-2011, and 2011-2012, three different teams of assessors evaluated the reports. Three different authors then wrote up reports which are referred to as *Assessment Report Worksheet #3* at the College of Micronesia-FSM. Laying the three different reports side-by-side, one is left unable to draw comparative conclusions across the three academic years. Setting aside the issue of different teams of assessors for each of the three years, each report chooses different data, different data presentations, and different analyses.

After the completion of this report, the original data was found for 2010-2011. Prior to access to this data, *Assessment Report Worksheet #3* for 2010-2011 was used to reverse engineer the table values. The discovery of the data revealed that only one assessor was involved in marking each laboratory report. In order to generate comparable data after-the-fact, these values were doubled to produce scores out of 32.

The raw data for 2009-2010 generates results which do not match the results reported *Assessment Report Worksheet #3* for 2009-2010. Unable to generate matching data, this report chooses to use values from a re-analysis of the 2009-2010 data.

The key results are summarized by the averages in the following table. The total possible is noted in the first column.

| Category | 2009 | 2010 | 2011 |
|---|---|---|---|
| Chuuk /32 | 11.57 | | 10.28 |
| Kosrae /32 | 9.75 | 18 | 13.79 |
| National /32 | 13.77 | 20.41 | 13.91 |
| Pohnpei /32 | 7.5 | 21.64 | 12.2 |
| Yap /32 | 11.5 | 20.55 | |
| SC117 /32 | 9.69 | 22.36 | 10.04 |
| SC120 /32 | 12.58 | 18.97 | 13.69 |
| SC130 /32 | 10.6 | 21.29 | 14.83 |
| SC255 /32 | 13.57 | | |
| Sci Processes /8 | 3.18 | 5.41 | 3.74 |
| Strategies /8 | 3.16 | 5.30 | 3.28 |
| Communication/8 | 2.73 | 4.90 | 3.01 |
| Concepts/8 | 2.86 | 4.90 | 2.74 |
| Overall averages /32 | 11.92 | 21.9 | 12.77 |

The strong rise in scores for 2010-2011 is not due to any know intervention strategy and is considered to be an artifact of the changes in the assessors with each academic year. N/A refers to data that is either unavailable or which was not gathered. SC255 was dropped from the study after the first year, SC117 has been dropped for the 2012-2013 academic year. Data analysis was not reported for Chuuk in 2010, nor Yap in 2011.

The above tabular data is presented graphically below.



Average gen ed sci performance by site
Fall term laboratory report assessment
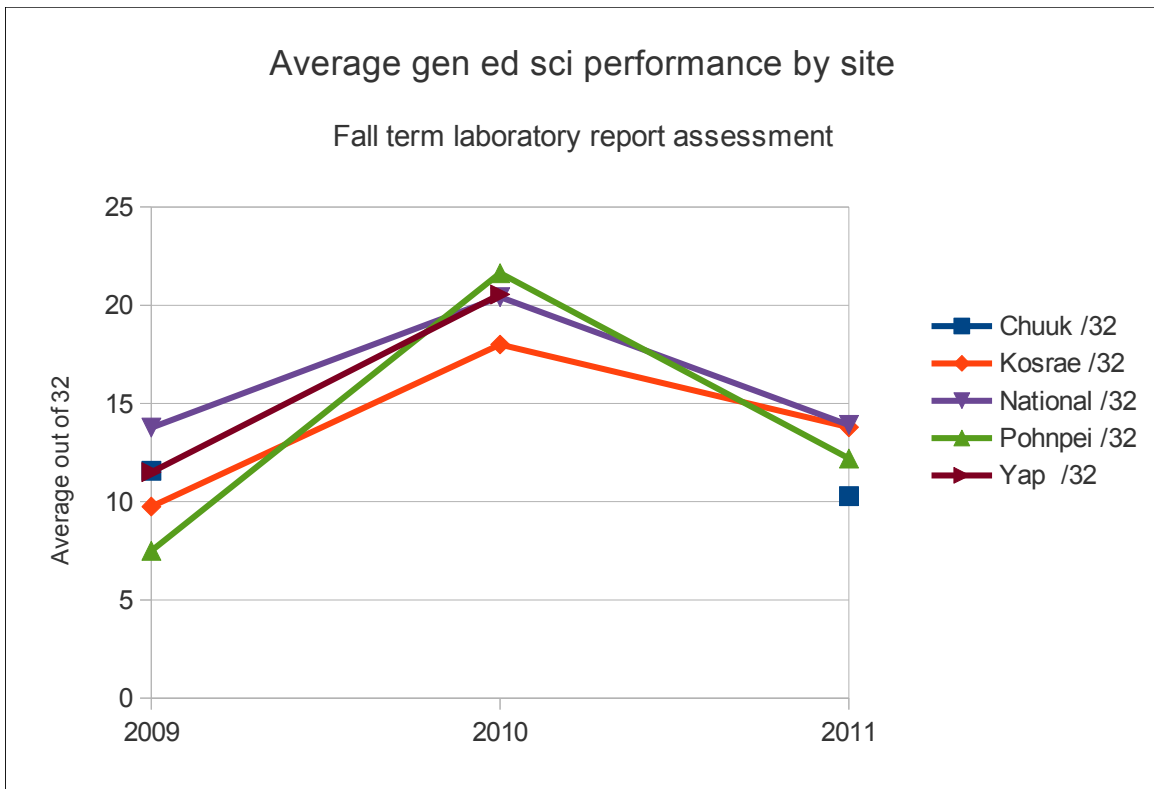
In the following discussion academic year 2010-2011 is generally set aside as being a marking standards issue. Thus there may be small improvements 2009-2010 to 2011-2012 for Pohnpei and Yap sites. The National site held at a constant level and first rank in 2009-2010 and 2011-2012. The small sample sizes, reported in the table below, for this study mean that small differences in average performance are not statistically significant.

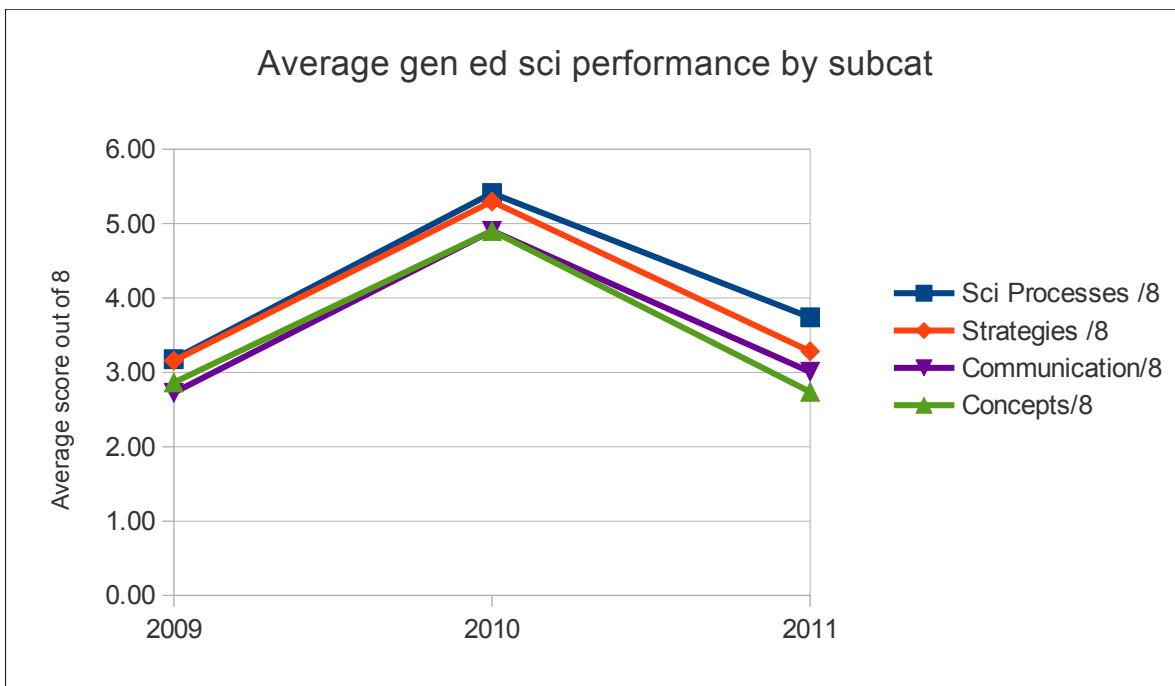| Sample n | 2009 | 2010 | 2011 |
| --- | --- | --- | --- |
| Chuuk | 7 | | 32 |
| Kosrae | 8 | 9 | 19 |
| National | 26 | 29 | 65 |
| Pohnpei | 6 | 22 | 25 |
| Yap | 4 | 11 | |

Note too that sample sizes are not consistent across the years.
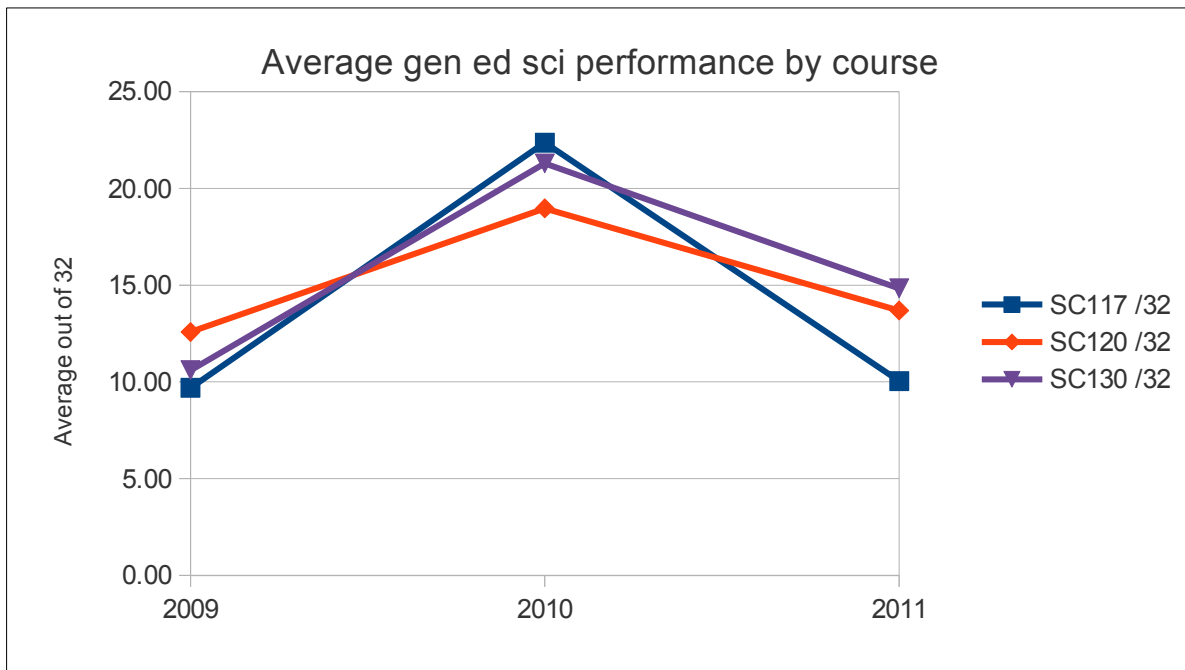


The results as reported by course, setting aside 2010-2011, suggests some improvement for SC 130 Physical Science and SC 120 Biology. An instructor for SC 117 noted in an August 2012 conference that the SC 117 does not produce laboratory reports which are aligned to the particular rubric being used. The instructor felt that SC 117 should be removed from the general education science program assessment in the future.

The following table provides the sample size for each course in each academic year.

| Sample n | 2009 | 2010 | 2011 |
|----------|------|------|------|
| SC117    | 13   | 11   | 47   |
| SC120    | 19   | 29   | 58   |
| SC130    | 5    | 31   | 36   |
| SC255    | 14   |      |      |

**Average gen ed sci performance by subcat**

Performance by the specific subcategory is out of eight possible rather than the 32 possible when all four sections are combined. Thus the above is the average score per laboratory out of eight possible. Excluding the 2010-2011, no subcategory has shown a strong movement system-wide. There may be a small up tick in performance on scientific processes and reasoning skills.

The above chart is "rotated" below to match a chart presented in the 2010-2011 report.



**Average gen ed sci performance**

By subcategory, rotated
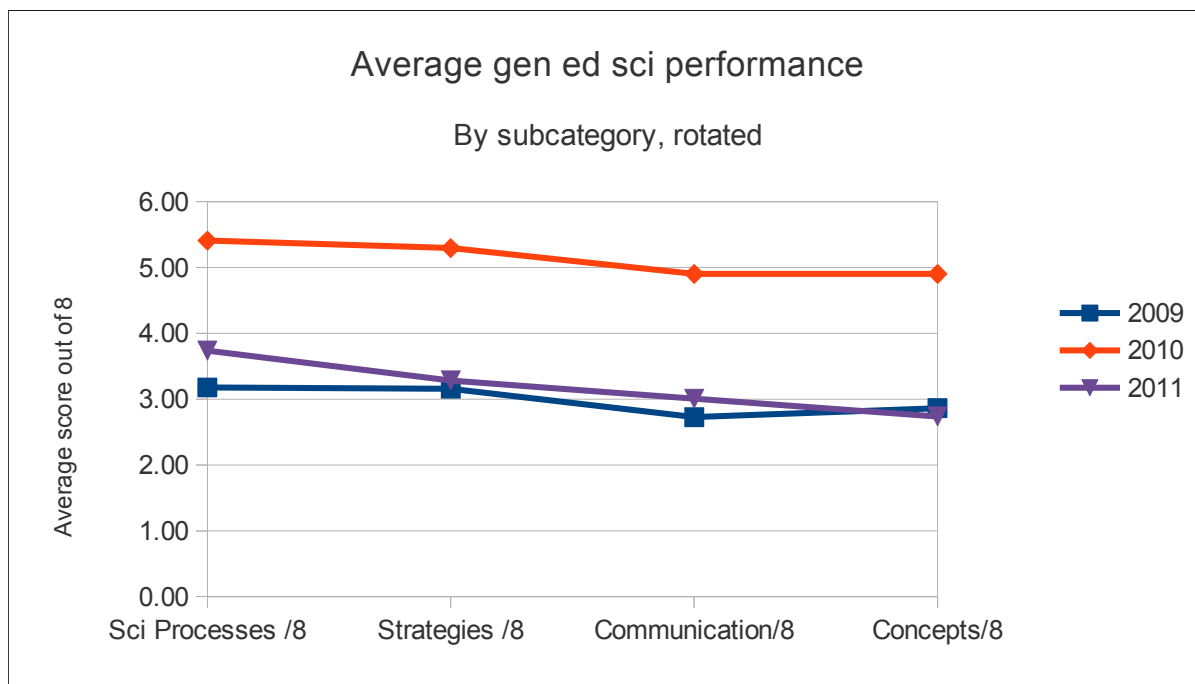
This chart appears different from that seen in the 2010-2011 *Assessment Report Worksheet #3* primarily due to the inability of this author to generate the numbers seen in the report

from 2009-2010.

Overall there may be a small statistical improvement from 11.92 in 2009-2010 to 12.77 in 2011-2012.

For the 2012-2013 academic year the assessment team has recommended resampling and remarking laboratory reports from across all four academic fall terms (2009, 2010, 2011, 2012). This effort is scheduled to occur early in 2013. The hope is that this will provide a set of comparable scores.