IB Research Notes

Information for the IB research community

Volume 3, Issue 3 July 2003

In this issue:

Feature Article

A Comparison of Student Assessment within the IB Diploma Programme (DP) and Advanced Placement (AP) Chemistry Courses

John Hare

Research Noticeboard

- Journal of Research in International Education
- Research literature
- International education research database
- IBO public web site
- Online curriculum centre

Welcome to *IB Research Notes*, Volume 3 Issue 3.

This issue's feature article is by John Hare. John was prompted to write the article after reading the National Research Council's report, *Learning and Understanding: Improving Advanced Study of Mathematics and Science in US High Schools*, prepared by the council's Committee on Programs for Advanced Study of Mathematics and Science in American High Schools. John's article focuses on particular aspects of the report, especially matters related to assessment. Two commentaries also accompany this article. The first is by Professor Arden Zipp, former chief examiner for chemistry for the IBO as well as chief reader for chemistry in the Advanced Placement (AP) programme, the second commentary is provided by Kristen Jones an AP chemistry teacher from Texas.

We would be interested in hearing from other authors who are interested in commenting on other aspects of the National Research Council report or other articles related to any field of international education. They can e-mail ibru@ibo.org to discuss their ideas and suggestions for articles.

IB teachers can access *IB Research Notes* via the research pages of the online curriculum centre and in the research section of the IBO public web pages.

Roger Brown
Head of research support and development

IB Research Notes is published four times a year and is a joint publication of the International Baccalaureate Research Unit (IBRU) and the International Baccalaureate Curriculum and Assessment Centre (IBCA).

Contact details: IBRU
Department of Education
University of Bath
Claverton Down, Bath, England
GB BA2 7AY
UNITED KINGDOM
Fax: +44 1225 323 277
E-mail: ibru@ibo.org

IBO

Feature Article

John Hare began his career as a research biochemist and then moved into human resources management. He entered teaching as a mature entrant and has taught science in the UK and USA. John is now a postgraduate student at the University of Bath. His particular interest is the contribution that science education makes within the holistic nature of the MYP.

A Comparison of Student Assessment within the IB Diploma Programme (DP) and Advanced Placement (AP) Chemistry Courses

John Hare, postgraduate student, University of Bath

Abstract

This paper focuses on the assessment of DP and AP courses as described by the chemistry review panel for the NRC report. Both chemistry courses utilize a written, summative examination for student assessment and the DP also includes an internal assessment component. The NRC chemistry review panel questioned whether the DP or AP chemistry examinations assessed deep understanding and contextual appreciation of chemistry. This paper seeks to challenge some of the findings of the NRC report and provides suggestions for improving assessment within both programmes.

Introduction

The report, *Learning and Understanding: Improving Advanced Study of Mathematics and Science in US High Schools* (National Research Council, 2002), was the outcome of a two-year study by a committee of the US National Research Council (NRC) and sought to:

> *…explore the current status of high school mathematics and science education by means of an in-depth look at programs designed for advanced students, such as the AP and IB programs. (National Research Council 2002a p.17)*

The remit of the committee was wide and examined courses of study in mathematics and science at an advanced level in US high schools. The study focused on the International Baccalaureate Diploma Programme (DP) and the Advanced Placement (AP) programme since these are the most widely adopted courses of study within the USA.

This paper focuses on the assessment of chemistry students within USA high schools as described by the chemistry review panel for the NRC report. Comparisons are made of the assessment processes that support the DP and AP chemistry courses, and areas for further study and evaluation are recommended.

Student assessment within the AP and DP chemistry courses

The summative examination

Both chemistry courses utilize a written, summative examination for student assessment. However, the DP uses a combination of this and ongoing, teacher-led assessment in determining student capability and understanding of the subject material. The DP chemistry examination consists of three externally assessed, written examination papers that constitute 76% of the final student assessment mark. One of these papers contains multiple-choice questions. By comparison, the AP course relies solely on a written examination, consisting of a multiple-choice and a free-response element.

The chemistry review panel reviewed the style and content of the two examination processes and made the following general comment.

> *[The] kinds and levels of questions that appear on both the AP and IB examinations reinforce[s] the emphasis on broad but shallow coverage of topics. (National Research Council 2002b p.17)*

This criticism of the DP examinations is not supported by evidence presented by the chemistry review panel. A preliminary inspection of the 1999 and 2000 DP chemistry examination papers confirmed that these papers include questions that go beyond recall of facts, and questions that draw together various aspects of the chemistry course. It also confirmed that the questions help to develop skills of data interpretation and application. Indeed, by way of agreement and apparent contradiction, the review panel commented as follows.

> *AP examinations ask less than IB exams about the application of concepts, especially with respect to new contexts or chemical systems. (National Research Council 2002b p.21)*

A brief review of the 2000, 2001 and 2002 AP chemistry papers confirmed a frequency of questions directed towards a recall of data and facts. Furthermore, there is a similarity in question style between one year and another. Undoubtedly, there are questions on the DP chemistry examination that also require factual recall, and the issue within any examination paper is always that of balance. There will invariably be questions that require factual recall and knowledge of principles while other questions should lead a student, through the application of this knowledge, to make deductions and conclusions on previously unseen information.

The chemistry review panel, on questioning whether the DP or AP chemistry examinations assess a deep understanding and contextual appreciation of chemistry by the student, made the following comment.

> *A striking inadequacy of the AP and IB programmes is the lack of detailed research about what their examinations actually measure, including the kinds of thinking that the examinations elicit. (National Research Council 2002a p.10)*

Clearly there is a need to evaluate what exactly is being assessed in both these programmes' summative examinations and where improvements could be made. A thorough evaluation of what the AP and the DP examinations actually assess awaits completion by the author. There may never be a perfect student-evaluation process, however, it seems apparent that the single, written examination with no teacher-led input to assess student capability cannot be the preferred approach.

The role of multiple-choice questions as a means of assessment

As identified above, both chemistry courses have written summative examinations that include multiple-choice questions as a compulsory element.

The chemistry review panel does not comment upon the role of multiple-choice questions within either of the examination processes or, indeed, their value and validity in assessing an understanding of the subject matter. Standardized tests such as multiple-choice questions, that treat all candidates the same, might be an attractive evaluative tool but they are not aligned to current thoughts on assessment as illustrated in the following quotation.

> *In the context of testing, equity requires us to ensure that human judgement is not overrun or made obsolete by an efficient, mechanical scoring system. Externally designed and externally mandated tests are dangerously immune to the possibility that a student might legitimately need to have a question rephrased or might deserve the opportunity to defend an unexpected or "incorrect" answer when the questions are well-structured and the answers are multiple choice. (Wiggins 1989 in Gipps, p155)*

It is possible that multiple-choice questions, which follow a familiar format, may encourage rote-learning and teaching to the examination. It is also difficult with this assessment paradigm to envisage how a student can demonstrate their ability to use their evaluative skills in a previously unseen area in chemistry.

Internal assessment within the DP chemistry course

Internal assessment (IA) or teacher-led assessment of student work accounts for 24% of the final student mark in the DP chemistry course. The IBO externally moderates this work to ensure consistency in assessment between teachers and accredited schools. Common IA criteria are used for DP students regardless of their location and guidance is given on the application of these criteria to develop consistency. The IA includes investigative laboratory studies, a cross-discipline experimental sciences project and an "extended essay" which may go beyond the immediate confines of the subject material on the syllabus. The assessment framework for the investigative chemistry is clearly described within the chemistry programme guidelines, as are the detailed assessment criteria.

Teacher-led assessment is also a feature of the DP biology and physics courses and is a part of the assessment paradigms of the UK high school science programmes.

The IA component is absent from the AP chemistry course and the chemistry review panel make no comment about its educational value or the benefits it may bring to the AP student assessment regime.

The value of teacher-led assessment

The purpose of teacher-led assessment is to provide additional opportunities for student assessment and to give a wider perspective of student ability over the entire course than that given by a single written examination (Gipps 1994, p.124).

The NRC report makes little comment on the value of the DP theory of knowledge (TOK) course, group 4 interdisciplinary project and extended essay. Each of these has a significant role to play in assessing a student in terms of their data search and analysis skills; their team-working, interpretative and evaluative skills; their attitudes; and their aptitudes. All of these are significant undergraduate skills and may be indicators of future college success. For example, the extended essay asks

students to work independently to find relevant information, evaluate it and comment on their findings; much as an undergraduate would be required to do. None of these elements appear in the AP chemistry course.

Grading students within the DP and AP chemistry courses

The final grade awarded to DP chemistry students is criterion-referenced to standards. If the student can demonstrate skills and knowledge at a certain predetermined level then the student is awarded that grade. There is no maximum percentage of students that can be awarded a particular grade.

The AP chemistry student is graded differently. The AP grade is determined by reference to several factors including how well others performed in the AP chemistry test over previous years and how well college students performed in the test. The chemistry review panel commented that: "…the final grade received in an AP chemistry course is not specifically linked with the student's performance on the AP examination" (National Research Council 2002b p.20).

In summary, the final grade for an AP chemistry student is determined not by what they know but what others know as well. This would suggest that students from different years but of the same ability could be graded differently. This approach may lead to questions concerning the comparable merit of a particular grade depending on the year the student passed the AP chemistry examination and, by inference, the ability of that student.

Recommendations

The chemistry review panel emphasized that chemistry education should move closer to the constructivist model of learning which is described in the NRC report and elsewhere. It is surprising therefore, that the panel made no reference to the value of teacher-led student assessment in this context as it has been linked with this constructivist model (Gipps 1994b). Undoubtedly, teaching methods may benefit from the adoption of a constructivist approach and this would be an ambitious and worthy target to pursue. The question remains however, whether it would be more appropriate to improve the student assessment processes before implementing a constructivist model.

The following recommendations are made for both programmes.

Examinations and their value

It would be advantageous to review the AP and DP chemistry examination papers in more detail over an extended period. A comparative review of the following points would be informative:

- the style of the questions asked
- the frequency and similarity of questions from year to year
- the balance of the topics that are examined between various years
- the purpose behind the questions and the response that they generated from the students
- whether the questions elicit an appreciation of the depth of understanding of the subject matter
- what other approaches may assess this depth of understanding.

An evaluation is required of whether multiple-choice examinations are a valid method for assessing a deep understanding and appreciation of the concepts within a chemistry course and whether they have a place in the examination regimen. They may, for example, have a role in questions that examine the capacity to evaluate and deduce from given data. The possibility exists nevertheless, that a student may obtain a mark through luck rather than deep understanding. From this data, it is believed that a satisfactory evaluation can be made of what the summative examination processes are actually assessing and whether the process can be enhanced.

Teacher-led assessment of student ability

A review and critique of the internal assessment criteria used within the DP chemistry course is recommended and could further enhance this substantial area of difference between the AP and DP schemes.

It would be valuable to examine how teacher-led assessment can be further utilized in the student evaluation process. Such examination may lead to a move away from the current, significant bias on the formal written examination process. Teacher-led assessments may be a better indicator of whether a student actually understands the subject matter in depth and can apply the material in an appropriate manner. It is appreciated that such assessment could lead to an increased administrative burden on teaching staff and this should be minimized.

The contribution of the DP's theory of knowledge, group 4 project and extended essay components as student assessment tools should also be defined.

Concluding remarks

This paper has focused on the student assessment methods used in the AP and DP chemistry courses. Both are preparative courses for potential undergraduates in their final year of high school education. It is accepted that this paper presents a very narrow comparison of the two programmes. However, the comparisons do raise valid questions.

To give a broader comparison of the two programmes leading to enhancements in one or both courses of study, detailed comparisons of other aspects of the chemistry programmes are necessary. These should include:

- detailed comparisons of the content and balance of the courses
- the interrelationships of the chemistry courses with other academic courses that are pursued simultaneously
- the links that are made between the chemistry courses and other science disciplines to present a holistic view of science as a discipline
- the role and contribution of experimental investigative studies as an integral part of the chemistry course.

With such additional information, a broad and detailed comparison of these programmes can be made, which may indicate the preferred route for students as they embark on their undergraduate science studies.

Acknowledgments

Valuable discussions with Lloyd Hacker (head of science) and Kevin Skeoch (head, senior school) of the International School of the Basel Region, AG, are gratefully acknowledged.

References

IBO. 2001. *Diploma Programme Chemistry Guide* (for first examinations in 2003). Geneva. International Baccalaureate Organization.

Gipps, CV. 1994a. in *Beyond Testing; Towards a Theory of Educational Assessment*. London. RoutledgeFalmer.

National Research Council. 2002a. *Learning and Understanding: Improving Advanced Study of Mathematics and Science in US High Schools, Committee on Programs for Advanced Study of Mathematics and Science in American High Schools*. Gollub, JP, Bertenthal, MW, Labov, JB and Curtis, PC (Eds). Washington DC. National Academy of Sciences.

National Research Council. 2002b. *Learning and Understanding: Improving Advanced Study of Mathematics and Science in U.S. High Schools: Report of the Content Panel for Chemistry*. National Research Council. Retrieved 6 June, 2003, from the World Wide Web: http://books.nap.edu/books/NI000403/html/

Wiggins, G. 1989. "Teaching to the (authentic) test" *Educational Leadership*, 46, (7), pp 41–7.

Response to John Hare's article: 1

Arden Zipp, former chief examiner for IB Diploma Programme chemistry and former chief reader in the AP chemistry programme.

John Hare's paper raises a number of interesting issues that need further discussion, however of main interest to me was that neither his paper nor the original NRC report recognized the largest single difference between the IB Diploma Programme (DP) and Advanced Placement (AP): the syllabus details of the two programmes.

The DP chemistry syllabus is very detailed and is broken down according to topics, sub-topics, and assessment statements. Accordingly, the DP exams are constructed to explicitly address the assessment statements, and each question can be linked to one (or more) assessment statement. Therefore, DP teachers know exactly what they should teach for the exam. The AP syllabus, on the other hand, is far less detailed, consisting only of a list of topics and sub-topics. When AP exams are assembled, there are no assessment statements to provide a limit for questions, which often require students to apply their knowledge of a topic to new situations.

There are differences in both syllabus coverage and emphasis between the various examination papers in the two programmes. In the DP the core is covered by the multiple-choice paper and one of the free-response papers, while the other free-response paper examines the "options". In contrast, the multiple-choice and free-response AP papers examine the same block of material.

In my view John Hare has misinterpreted the awarding of grades in both programmes. As someone who has participated in the establishment of grade boundaries for both AP chemistry (serving as chief reader, CR) and DP chemistry (chief examiner, CE), I believe both programmes take great care to award each student the grade (s)he deserves. The DP achieves this by having members of a grade award team (including the CE) read individual scripts and establish boundaries for the grades 1–7. In AP the CR calls upon her/his expertise as a chemistry educator to fix the divisions for the grades from 1 to 5 (based on the skills demonstrated by the students in that sample). The CR is also provided with data about the performance of the current group of students (relative to a previous population) on a subset of the questions on the multiple-choice paper. The purpose of this data is not to generate a standard distribution but to ensure that students are being assessed by the same standards used in previous years. As with the DP programme assessment, **there is no maximum percentage of students that can be awarded a particular grade** in AP. However, it should not be surprising that a comparison of the performances of a large number of individuals in both programmes results in a fairly standard distribution.

The paper concludes by recommending that AP and DP chemistry examinations should be evaluated over a longer period of time with regard to the types and purposes of the questions asked as well as the responses generated by them. It also raises the question of "whether multiple-choice examinations are a valid method for assessing a deep understanding and appreciation of the concepts within a chemistry course, and whether they have a place in the examination regimen."

Presumably, the goal of any assessment tool is to determine candidates' knowledge of a body of information as set forth in the syllabus, and her/his ability to apply that knowledge to new situations. This is a goal that multiple-choice questions can achieve with a high degree of reliability. Even if such questions are not good at "assessing a deep understanding and appreciation of the concepts...", they are still useful for the tasks they can perform. There is probably no single assessment tool that can serve all ends, and the best overall assessment requires many different tools.

Response to John Hare's article: 2

Kristen Jones, AP chemistry teacher, A&M Consolidated HS, College Station, Texas

While addressing the criticism made in the NRC report that many questions on both the AP and DP chemistry exams emphasize "broad but shallow coverage of topics", the author states that this is not supported by inspection of recent DP examinations. He does, however, believe that the recent AP exams emphasize recall of data and facts. As a teacher of AP chemistry for the past 14 years, I feel that the mathematical sections of the AP test often include questions that only require simple calculations. In recent tests, this is more apparent as the test writers have broken problems down into smaller sections to speed up and improve the grading of over 60,000 tests. One of the strengths and challenges of these mathematical questions is that they combine concepts and calculations from all parts of the chemistry curriculum into one problem. Rarely do textbooks (and college courses) ever expect a student to combine stoichiometry, gas laws, equilibrium, thermochemistry and electrochemistry calculations in one problem. This is not uncommon on an AP chemistry test. The non-mathematical "essay" questions frequently expect the students to "justify your answer", "explain your reasoning" and "explain why". Simply knowing facts or trends will not give the student much credit. This tests the depth of a student's knowledge and their ability to apply this knowledge in a variety of situations.

The author mentions the NRC criticism of the "predictability" of AP test questions. There is some validity to this criticism and there have been discussions that may lead to changes in the AP test format. Traditionally, the first mathematical free-response problem has always been a multi-part, multi-concept equilibrium problem. One of the strengths of the AP programme is the depth to which various types of equilibria must be covered. For most students, this is the most challenging part of the course and can easily involve 20–25% of the curriculum. Knowing that questions on this topic are a considerable part of the test helps to ensure that teachers emphasize mastery of this important concept.

An area of criticism in the NRC report that was not mentioned by the author involves the testing of laboratory knowledge. The report stated that this has only been tested recently on the AP examination. Multiple-choice and free-response laboratory questions have always been a part of the AP examination. In recent years, a required free-response question has been devoted to one of the 22 recommended AP labs. Parts of other questions also require application of laboratory knowledge. On the 1999 multiple-choice section, lab safety, spectroscopy, flame test results, qualitative analysis and other lab procedures were tested. Other questions tested real-life applications of chemical knowledge that required more than simple recall. Because AP and DP chemistry courses vary widely in the amount of time spent in the laboratory, colleges may want to examine a student's laboratory notebook before giving laboratory credit.

The author misinterpreted a quote concerning the relationship between the final grade in an AP chemistry course and the AP examination score. The NRC report stated that the AP test results were not returned to the school until after the close of the school year and could therefore not be factored into the student grade. This has nothing to do with the determination of the AP test score that is reported to the colleges. The author suggests that students from different years but of the same ability could be graded differently. When the numerical AP test results are being evaluated to determine the AP score break points, results on multiple-choice test questions given in multiple years are compared in an attempt to ensure that a student with a given knowledge would be graded consistently from year to year.

This article has looked at the evaluation section of the NRC report. It would be beneficial to look also at other sections to evaluate statements involving curriculum and teacher preparation. While we may not agree with parts of the document, there is much we can learn from this "outside view" of our programmes.

Research Noticeboard

Journal of Research in International Education

Information about this journal can be found at: http://www.sagepub.co.uk

Research literature

Cambridge, J. (2003). "Identifying globalist and internationalist missions of international schools". *International Schools Journal, XXII* (2), 54–58.

Hodgson, A. & Spors, K. (Eds). (2003). *Beyond A-levels: Curriculum 2000 and the Reform of 14–19 Qualifications*. London, Kogan Page.

Pearce, R. (2003). "Cultural values of international schools". *International Schools Journal, XXII* (2), 59–65.

Phillips, G. & Pound, T. (Eds). (2003). *The Baccalaureate: A Model for Curriculum Reform*. London, Kogan Page.

Stables, A. (2003). *Education for diversity: making differences*. Aldershot, Ashgate Publishing Ltd.

International education research database

This database has now been launched and currently contains over 1500 research articles on international education along with research articles on the International Baccalaureate programmes. The international education research database can be accessed at www.ibo.org. Access the shortcuts menu to go to the research pages which provide a link to the searchable research database.

IBO public web site

The IBO's main web site (http://www.ibo.org) provides general information about the organization and its programmes.

Online curriculum centre

The online curriculum centre (http://online.ibo.org) is available to all teachers in IB schools that subscribe to the site. The online curriculum centre is a valuable source of information for those considering research related to the IB programmes.