

## ACS Exams and Assessing Student Learning in Chemistry Classes

**Thomas Holme**, University of Wisconsin-Milwaukee, Director, ACS Examinations Institute

### Abstract

The ACS Exams Institute enjoys a history of more than 70 years of producing nationally normed exams for chemistry classes. Currently, ACS Exams exist for a wide range of classes at the college and high school level. This paper will discuss the process by which exams are developed, released and secured as assessment tools for chemistry educators currently. It will also look to the development of electronic delivery methods for ACS Exams and talk about the importance of security in an electronic educational environment and the challenges that confront ACS Exams in maintaining security on exams so that they remain valid assessment instruments. Presuming an ability to address the security concerns, the opportunities provided by electronic delivery of exams will also be noted in this paper.

### Paper

The ACS Exams Institute can trace its roots back over 70 years to the Cooperative Exams produced by committees of chemical educators as far back as 1934. At that time, many disciplines produced these cooperative exams, but Chemistry appears to be the only one that has maintained this service over the decades. This fact is a testament to not only the active role taken by [the American](#) Chemical Society in education, but also the skill and tenacity of specific leaders such as Prof. Ted Ashford and Prof. Dwaine Eubanks.

ACS Exams have always been produced as multiple-choice, paper and pencil exams. Over the years the number of such exams grew from originally including only General Chemistry to now covering high school and undergraduate courses throughout the entire curriculum. Still, at present our exams are all paper-and-pencil instruments, but we are moving rapidly towards providing electronic delivery of exams. In this paper we will look at the current exams, how they are produced, copyrighted and normed. Then we'll look at some of the work we're currently undertaking to advance the utility of exams, most of which are planned for launch in the near future.

### The Tradition of Producing ACS Exams

Our current procedures for producing exams have been established for many years and enjoyed the participation of many hundreds of dedicated volunteers from the Chemical Education community. Exam development in most cases takes roughly 2-3 years. The Exams Institute appoints a chair for a new exam committee, usually from the membership of the previous committee, and then works with that chair to identify people to serve as members. Every effort is made to establish a committee with a diverse membership along several vectors, including institution type (research, comprehensive, liberal arts and community colleges where appropriate) and size; geographical balance as well as more traditional measures of diversity

such as gender and ethnic background. Regardless of background, all committee members are expected to be actively teaching the course for which the exam is being developed.

Once appointed, committee members do much of their work independently (with regular email communication) but 2-3 face-to-face meetings are also held normally in conjunction with ACS National Meetings, the Biennial Conference on Chemical Education or the ChemEd meeting (this final venue often hosts meetings only for High School exam committees).

Most committees hold organizational meetings at which basic design parameters for the exam are decided. While logistical issues such as number of items to be included in the exam and constraints on students (such as access to calculators or mathematical equations) are included in these discussions, the most important dialogue occurs in terms of content coverage to be included in the exam. For some exams, such as General Chemistry, the content coverage tends to remain relatively stable, the fact that this conversation occurs for each committee in some form provides an important assurance of content validity for ACS Exams. If users of our exams have contacted us with content coverage concerns for a previously released exam, these communications are forwarded to the chair and are considered during the content discussion at the outset of the new exam writing process.

Once the content has been agreed upon, the committee agrees to a distribution of item (question) writing assignments for all committee members. Committee members do their initial writing of items, and for some committees conduct editing as well, as individuals or small groups communicating by email. At some point, either the chair or an appointed secretary, collects and collates the items being proposed for the exam.

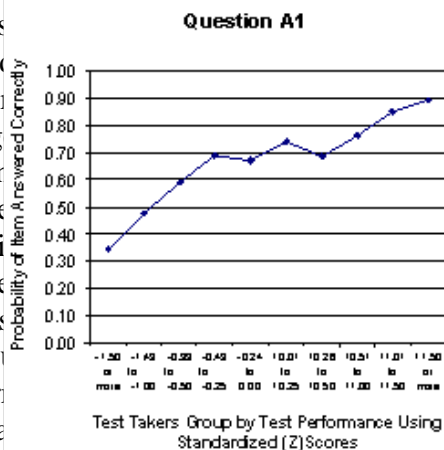
The next committee meeting is arguably the most important. The editing and choosing of items to be included on the trial versions of the exam is accomplished at this meeting. Most committees take one and a half days of effort for this step, and sometimes more. At this point items are vetted with the collective expertise of the members. This process looks for both content issues and the construction of the items themselves. So, while some items are edited (or even discarded) based on difficulties associated with the chemistry inherent in the question, others are determined to provide too many hints for the unknowledgeable student (construct concerns). Virtually all items undergo some level of editing at this stage and many items ultimately look very little like the question initially written by a single member.

By the end of the editing meeting, a set of items roughly twice the number as will be needed in the final version of the exam has been agreed to by the committee. In some cases, items are explicitly paired, so that a particular content point is asked in two different ways. Such items appear on different versions of the trial exams so that the construct validity can be determined (as it is for all items.) Construct validity in this sense is determined mostly by statistical means the item statistics including difficulty and discrimination are measured by utilizing the trial exams with students from around the country.

Before the trial tests are administered, however, the Exams Institute staff takes the items from the committees and sets them in two column format with reasonable illustrations (though not always the final form of the illustrations.) Several proofreading cycles are undertaken at this

stage, and the Institute staff is simultaneously lining up volunteers to use the Trial Tests in their classrooms.

Ideally enough students take the trial tests in one finals season to allow for statistical measures of validity to be calculated. For large enrollment courses like General Chemistry or Organic Chemistry this goal is nearly always attained. Upper division classes often require two or even more testing cycles to have enough student performances for the Institute calculates several statistical measures. The fraction of students who answer the item correctly (so it appear on released ACS Exams fall in the difficulty range of this range can be included to allow for appropriate correction of the discrimination which is calculated by determining the quarter of the pool who answer correctly minus the fraction who do. Ideally, a discriminating item would be answered correctly by the high performing students and incorrectly by the low performing students. The worst case scenario would be an item with no discrimination (which would indicate that the low performing students often than the high performing student.) Items with negative discrimination are not included in the released version of an ACS Exam.



Recently, the Exams Institute has provided another tool to committees to discern the construct validity of items. The Item Characteristic Curve (ICC) as shown in Figure 1, provides a visual representation that augments the difficulty and discrimination index calculations. To calculate the ICC, student performances are binned. The bin cutoffs are established by the mean ( $\bar{x}$ ) and the standard deviation ( $\sigma$ ), as noted in Table 1.

**Figure 1.**  
Example of an Item Characteristic Curve

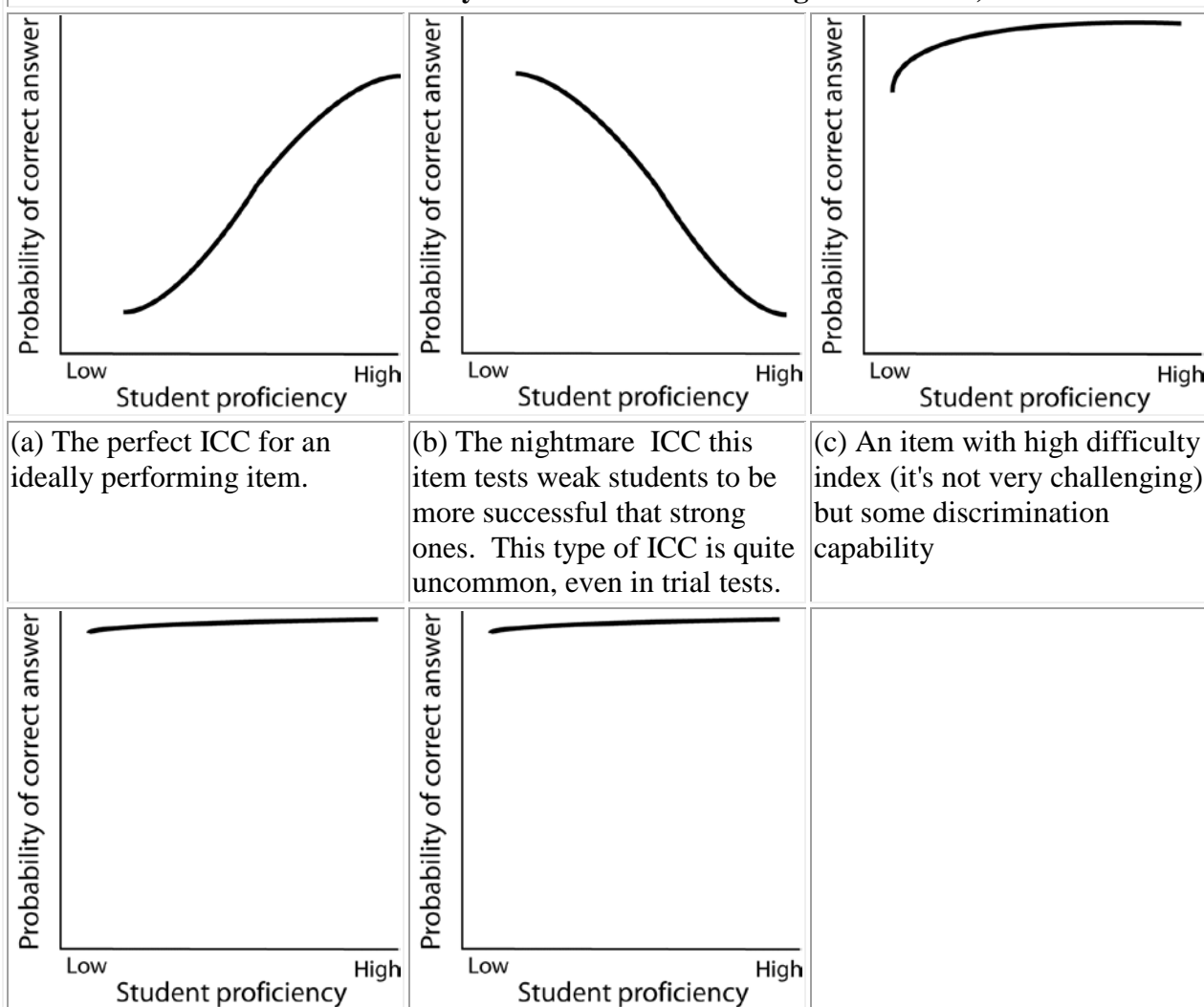
Table 1: Cutoff points for bins used in calculation of ICC curves.

1	2	3	4	5	6	7	8	9	10
< $\bar{x} - 1.5\sigma$	$\bar{x} - 1.5\sigma$	$\bar{x} - 1\sigma$	$\bar{x} - 0.5\sigma$	$\bar{x}$	$\bar{x} + 0.25\sigma$	$\bar{x} + 0.25\sigma$	$\bar{x} + 0.5\sigma$	$\bar{x} + \sigma$	> $\bar{x} + 1.5\sigma$
	to	to	to	to	to	to	to	to	
	$\bar{x} - 1\sigma$	$\bar{x}$	$\bar{x}$	$\bar{x} + 0.25\sigma$	$\bar{x} + 0.25\sigma$	$\bar{x} + 0.5\sigma$	$\bar{x} + \sigma$	$\bar{x} + \sigma$	

		$0.5\sigma$	$0.25\sigma$	$x$	$0.25\sigma$	$x + 0.5\sigma$		$1.5\sigma$	
--	--	-------------	--------------	-----	--------------	-----------------	--	-------------	--

In our experience, this strategy provides bins with roughly the same number of students in each one (because our exam results are normally Gaussian). The ICC curve then plots the fraction correct within each bin and connects them with a smooth line. There are several possible curves available, as shown with comments in Figure 2. This level of analysis was provided for the first time to the committee that is developing the 2007 General Chemistry exam. To be able to calculate a meaningful ICC curve requires a larger data set of student performances than to calculate the difficulty and discrimination. The Institute is considering how to provide ICC's for released exams as well, presumably in an on-line fashion.

**Figure 2. Prototypes of Item Characteristic Curves (based on work from Dr. Larry Meyers for statistical analysis on the California Diagnostic Exam.)**



(d) An item with high difficulty index (not challenging) and little discrimination as well.	(e) An item that has a medium difficulty, so it might look fine in terms of the numbers, but it doesn't really discriminate particularly well.	
---	--	--

With items chosen based on consideration of both content and importantly well constructed items (construct validity) the set of items that will compose the released exam is forwarded to the Institute. Now, illustrations and spectra are improved and either a single form or two forms (for large enrollment exams such as High School and General Chemistry) are produced. Norms are calculated based on volunteer return of student scores from faculty members who have used the exam. A relatively new feature on our web site provides nearly instant feedback for instructors using new exams, comparing the performance of their students with those previously submitted, even when insufficient data has been collected to calculate the formal norm for the exam.

### **Exam Security and Copyright**

As noted earlier, ACS Exams carry a proud heritage with an origin in the cooperative exams of the 1930's. In part because they are the result of the cooperative efforts of many educators still, the secure nature of the exams was handled in a spirit of collegiality for literally decades. While breaches of security occasionally occurred they were accidental and localized. The emergence of the Internet has changed this equation dramatically and the Institute has changed its posture accordingly.

In spring of 2002, two of the Organic Chemistry exams were posted on the faculty web-site at Yeshiva University along with their keys. The postings were ultimately reported by a faculty member in Texas in late November of 2002 when students there downloaded the exams. The Institute considered both exams fully compromised by this event, engaged in an immediate replacement program for all Departments who had purchased the compromised exams and a new Organic Exam was rapidly developed (from start to finish in under 14 months.) The Institute engaged the assistance of the corporate council of the ACS and ultimately of a Washington, DC law firm that specializes in Intellectual Property (IP) law. Suit was filed against Yeshiva and after two years of court proceedings (including depositions of the faculty member involved and the Exams Director, for example) a court ordered mediation process resulted in a Consent Judgment in favor of the Exams Institute that included a substantial cash award.

Since this case, the Institute has retained the services of our IP lawyers in several additional cases, including one this past spring when an Instructor at the University of South Florida posted 25 out of 60 items from the most recent Biochemistry Exam on an electronic reserve site. Again, the Institute has contacted all users of this exam and they will be replaced by a new exam, currently under expedited development.

These examples point to the importance of exam security. The Institute has taken the position that it is prepared to take any necessary steps to protect the security of exams so that chemistry educators can use them with confidence. The legal bills for these efforts have been significant,

but both the community and particularly the volunteers who spend many hours creating the exams, deserve to have the expectation that the Institute is vigilant.

To this effect, we now purchase extensive web searching services. A large number of keywords and keyword combinations are continuously searched by the services and the results monitored by the Exams Institute staff. As an example of the efficacy of this approach, this past spring a student blog site provided significant information about the nature of a symmetry item on the Inorganic Exam. Our web search found the blog site roughly 7 hours after it was posted. A faculty member from the student's college was contacted. That faculty member contacted the student and the blog entry was removed less than 30 hours after it was posted.

It's also important to note some specific details about the nature of the copyright protection of ACS Exams. Over the years, stories have been told, some of which were rooted in a lack of understanding at any level of the nature of the secure copyright. Today, all ACS Exams are copyrighted with a secure copyright. This means that the paperwork submitted contains no publication date because ACS Exams are NOT PUBLISHED when they are released. This is the same status that a secure exam like the GRE or SAT has. Thus, if you've looked at this document, you may have noticed I have always referred to exams as released, because that is the accurate term.

When two new versions of an exam have been produced, the exam is retired. When an exam is retired it is then considered published by the Institute. The Institute still holds the copyright for that exam it is most certainly not public domain but it no longer warrants the security of the exam. We do encourage our customers to maintain security on such retired, published exams but the idea of showing items from such an exam in a study session, for example, represents a viable usage (and fair use copyright principles of published works would corroborate this usage.) Note that publishing a complete published work on the internet would not constitute fair use under copyright law, so placing a PDF of a retired exam on the internet would still not be allowed.

We understand that many chemical educators are interested in study materials for students for ACS Exams. At this time, we have study guides available for General Chemistry and for Organic Chemistry. A new Physical Chemistry study guide is in advanced editing at this time and we are interested in working with educators to build study guides for Analytical Chemistry, Biochemistry and Inorganic Chemistry.

### **The Future of ACS Exams**

There are a number of ways that ACS Exams is looking to provide enhanced services for the future.

First consider new exam development. The tried and true formula of exam development can and has been engaged to produce wholly new exams. Thus, we now have an exam for the first-term of the 2-term Organic Chemistry sequence that will be released this fall. We have Spanish versions of both General Chemistry and Organic Chemistry nearly completed. We're always interested in exploring places where a new exam could be developed. Not all courses lend themselves to a standardized exam, however. One example is the one-semester Organic Survey

course that is taught on many campuses, but does not appear to have a sufficiently standard content coverage to allow us to construct an exam.

Perhaps the most exciting new exam being developed is the ACS DUCK, for Diagnostic of Undergraduate Chemistry Knowledge. As program assessment has become more important (often in response to regional accreditation expectations) the call for an exam to be used at the end of the undergraduate curriculum has become a more common one. Our DUCK will be an integrative instrument where items will follow reading passages describing experimental or other factual chemistry circumstances. The basic premise of this approach is that it makes no sense to organize an end of curriculum exam around the sub-disciplines. Departments could use our regular exams for that purpose. Instead, we started from the big ideas enumerated at the Exploring the Molecular Vision conference held by the ACS Society Committee on Education (SOCED) and are building the scenarios in the DUCK to make sure they are all elaborated. Our initial set of big ideas from the EMV are enumerated in Table 2, but this list could be extended as the DUCK itself is developed.

Number	Big Idea Concept
1	Atomic Theory and Structure
2	Chemical Bonding
3	Molecular structure and structure determination
4	Intermolecular forces and properties
5	Chemical reactions and synthesis
6	Thermodynamics and thermochemistry
7	Reaction kinetics and mechanisms
8	Chemical equilibrium
9	Experiments, safety and data handling

In addition to this form of new exam development, the Institute is also working to enhance the scholarly basis for the assessment instruments we produce. Thus, we are carrying out studies of things such as item order effects where possible. We are working to establish consensus level standards that are inherent for the content in our exams and then conducting alignment research for exams with these consensus content standards. We are also looking at additional cognitive analyses. For example, we are now working to establish measures of cognitive load for items on ACS Exams. Once completed, this analysis will complement the traditional construct validity measures of difficulty and discrimination for items on ACS Exams so instructors will be able to better analyze the performance of their students.

Finally, the Institute is moving towards electronic delivery of assessment materials. It is the premise of the Institute that this process should be driven by one of three potential advantages. Electronic delivery should (a) increase efficiency of measuring what we have always measured, or (b) increase the instructor flexibility of assessment, or (c) measure things that could not be measured with traditional exam formats.

The main mechanism for increased efficiency of measurement will be to introduce adaptive testing methods. Adaptive methods can be expected to measure student proficiency with better accuracy with fewer items (and therefore less time) than traditional methods. Our main target audience for the initial launch of adaptive methods will be for college entrance exams. Many campuses have testing centers that are already equipped to offer electronic testing, so this new product being built in conjunction with colleagues at Brigham Young University will face fewer technology based barriers to becoming useful in the community. We anticipate moving to beta-level testing of the first adaptive placement exam from the Institute this spring.

Adding flexibility holds it's greatest promise in those fields where the content coverage is affected by an overloaded curriculum. The prime example is Physical Chemistry where constraints in the academic calendar often force instructors to make decisions about where to include whole topics (such as where does kinetics get taught in a two-semester sequence.) Thus, our initial attempts at providing electronic delivery for enhanced flexibility will focus on Physical Chemistry as the test-bed. We are still working on the details of the delivery platform, but we expect to work with the team of computer science faculty at UMass-Amherst that has built the OWL program over the years. It's important to realize that what we are working on is the platform development, not the chemistry content of OWL. The Institute has it's own content (exams that have been developed and validated) so what we needed was expertise in delivering content electronically and that is where our partnership with UMass will focus.

Finally, we intend to use electronic delivered materials to measure things that have not been on the radar screen before. One idea that is routinely mentioned is to use the enhanced visualization capacity of computers to ask new forms of exam questions and we are certainly looking to conduct the appropriate scholarship (establishing reliability and validity of such items) to be able to do this. The project we are particularly excited about, however, lies in measuring problem solving skills using the IMMEX system. This web-delivered system provides students with complex problems solving tasks, captures student actions into a data base, and then uses data base mining techniques to establish the change in problem solving strategies used. It would take an entire additional paper to fully elaborate on this project, but the goal is to be able to measure student problem solving strategies on related concepts (such as structure-function concepts) as they progress through the undergraduate curriculum. Pilot studies of our ability to do this are currently underway under the auspices of an NSF Grant with the Institute, Clemson University and IMMEX at UCLA.

#### **Conclusion:**

With all these things happening, new exams under development, legal challenges for the Institute, the emergence of electronic delivery options in the near future, it really is an exciting time for the Exams Institute. We always need dedicated volunteers to work on projects (particularly exam writing) so I hope that some participants in this conference will consider adding their names to our list of potential volunteers. I look forward to the opportunity to discuss this paper when the time arrives.

---



**CONFICHEM** on-line conferences are organized by the ACS Division of Chemical Education's Committee on Computers in Chemical Education ([CCCE](#)).

---