

June 13, 2002

MEMORANDUM TO: R. William Borchardt, Associate Director  
for Inspection and Programs  
Office of Nuclear Reactor Regulation

FROM: Bruce A. Boger, Director */RA - William M. Dean for/*  
Division of Inspection Program Management  
Office of Nuclear Reactor Regulation

SUBJECT: SUMMARY OF THE ANNUAL NRC/INPO COORDINATION MEETING  
ON TRAINING-RELATED ISSUES

On April 10, 2002, a periodic NRC/Institute of Nuclear Power Operations (INPO) coordination meeting on training-related issues was held at NRC Headquarters, Rockville, Maryland. Such meetings are conducted in accordance with the NRC/INPO Memorandum of Agreement dated December 24, 1996. The purpose of the meeting was to discuss items of mutual interest concerning INPO's training program accreditation process as well as NRC issues on operator training. Participants included representatives of the NRC's Division of Inspection Program Management and INPO's Accreditation Division, a representative from the Professional Reactor Operator Society (PROS), who participated via telephone, and a representative from the McGraw Hill publishing company. The list of meeting attendees is provided as Attachment 1. The meeting agenda is provided as Attachment 2. Attachment 3 contains a compilation of revised INPO documents, as discussed below. Attachment 4 is a background white paper, sent to industry representatives, for agenda topic discussion purposes.

A summary of the discussions related to key agenda topics covered during the meeting follows.

### **Welcome and Opening Remarks**

Introductions of the NRC/INPO personnel present were conducted. After the introductions were completed, organizational changes since the last coordination meeting, at both the NRC and INPO, were discussed.

### **Technical Specifications (TS)**

The NRC staff briefly reviewed the history of the operator license eligibility issue, including the National Academy's publication of "Guidelines for Initial Training and Qualification of Licensed Operators," the NRC's publication of Revision 3 of Regulatory Guide 1.8, "Qualification and Training of Personnel for Nuclear Power Plants," and the NRC's issuance of Regulatory Issue Summary (RIS) 2001-01, "Eligibility of Operator License Applicants," which encouraged facility licensees to review and update their licensing basis documents. The staff noted that several facility licensees have responded to the RIS by submitting TS 5.3.1 amendment requests and that the staff has been deliberating how best to revise those specifications. The staff acknowledged that the Memorandum of Agreement with INPO discourages the codification of

Academy documents and noted that, to avoid a problem in this regard, it is considering the following TS language that would permit facility licensees to incorporate the specific eligibility criteria in the facility training procedure:

- 5.3.1 Each member of the facility staff shall meet or exceed the minimum qualifications of ANSI XXX-XXXX, with the following exceptions: 1) either the senior health physics supervisor or lead health physicist, shall meet or exceed the qualifications for "Radiation Protection Manager" in Regulatory Guide 1.8, XXXX, and **2) the education and experience eligibility requirements for operator license applicants, and changes thereto, shall be approved by the NRC and described in an applicable station training procedure.**

The INPO representatives were satisfied with the draft wording of TS 5.3.1.

### **SAT/Accreditation Issues**

INPO presented and discussed recently revised documents that provided guidance on their training accreditation process. The revised documents were provisionally issued in March 2002 (See Attachment 3).

The INPO staff described each document as follows:

- 1. Key Principles of Accreditation.** This is a one page, foundation document that lists 6 Key Principles of Accreditation. These are general statements of purpose and mission of accreditation. NRC noted that **Principle #6**, states that "the accreditation process is independent of but recognized by the NRC as a means for satisfying regulatory training requirements."
- 2. Regulatory Basis for Accreditation and Timeline.** This 3-page, foundation document overviews historical events and milestones. The Timeline outlines the milestone events from 1979 (INPO Incorporation) through 1993 (Final Rule establishing accreditation as means for compliance with Federal regulations).

As a follow-up to this meeting, the NRC staff noted an inaccuracy in the third paragraph of this publication. The paragraph erroneously stated that the Final Rule established INPO's accreditation as a means of compliance with Federal regulations. In fact, as the staff pointed out, the Final Rule did not state this. INPO agreed to correct the language in this paragraph.

- 3. The Objectives and Criteria for Accreditation of Training in the Nuclear Power Industry, ACAD 02-001.** This document provides the basis for conducting the self-evaluation and accreditation team reviews. This document was revised on the basis of extensive interviews and surveys conducted with industry executives, including Chief Executive Officers, and all INPO member training organizations. The purpose was to solicit viewpoints from all the varying levels within nuclear organizations. There are six (6) broad objectives, each containing multiple criteria to be met during the accreditation. The document also contains an Appendix that outlines the 5-phase Systematic Approach to Training (SAT) process.

4. *The Process for Accreditation of Training in the Nuclear Power Industry, ACAD-002*

This document is an overview of the accreditation process. It describes:

- purposes of accreditation,
- areas (operational and technical) to be accredited,
- accreditation process elements,
- the visitation process,
- self assessment,
- attaining and maintaining accreditation.
- withdrawing accreditation

INPO discussed its graded approach to the SAT process. This was described as a more flexible means toward allowing plants to make training program changes and modifications without necessarily conducting and documenting a comprehensive SAT process. NRC agreed that such a common-sense approach to SAT changes made sense, but cautioned that INPO and industry should be wary on what training situations merited a graded approach from those warranting an in-depth or comprehensive approach.

INPO stated it has eight accreditation teams that are set up to conduct accreditation visits using the revised procedures. Each plant has a scheduled visit once every four years per program area. To supplement both the utilities' requirement to perform a self-evaluation as well as the team visit itself, INPO has assigned a senior training representative to each plant. This INPO assignee will serve as an advisor who will assist plants in identifying and resolving training related issues during the four year self-evaluation phase and prior to the scheduled INPO team visit.

INPO informed NRC that it is using a new coupled approach to the accreditation visits; this approach couples the accreditation team visit to occur four months following the plant evaluation, permitting the accreditation team to review and incorporate any pertinent plant evaluation findings.

INPO stated that in its accreditation team visits in programs in 2001, it noticed the following problem areas:

1. Selection of continuing training program content
2. Ineffective use of simulators in training
3. Ineffective use of on-the-job-training (OJT) to reinforce management-identified training needs

INPO plans to review these areas during its accreditation visits.

### **Initial Examination Process**

INPO had assumed project responsibility for establishing a licensed operator examination test bank. Such an examination bank, as a repository of validated test items from past NRC license examinations, is an intended tool designed to support examination developers in the preparation of future NRC initial license examinations. Presently, INPO has approximately

19,000 test items inputted into the test bank. Only test items from 1998 through the present and beyond will be considered for future inclusion. INPO stated that it has encountered some difficulties which have included problems in downloading examination files from the NRC's ADAMS system because of software compatibility issues, incomplete or missing data from some examinations, and INPO resource limitations to perform the tasks. NRC agreed to support INPO in this project in any ways possible and noted that Region II NRC staff were working with INPO staff to resolve problems and seek ways to expedite the development, user-friendliness, and maintenance of the bank.

### **Requalification Issues**

NRC reintroduced a topic raised in the 2001 NRC/INPO Coordination Meeting on Training-related Issues: excessive test item overlap in biennial comprehensive requalification examinations among crews being tested in the same requalification cycle.

In preparation for this discussion, NRC had forwarded a white paper on this topic to INPO and to the PROS member for background (Refer to Attachment 4). NRC stated that because there have been instances of excessive test item overlap among comprehensive requalification examinations among crews who have undergone the same training program, there could be the *potential* for leaks of test content that might compromise the integrity of such examinations. Thus, NRC believes that examination integrity and validity could be better maintained in the development and administration of comprehensive requalification examinations if the number of test items that could be duplicated on successive comprehensive requalification examinations were limited. Specifically, NRC discussed the need to limit test item overlap in such comprehensive requalification examinations to no greater than 50 percent from items that appeared in previous comprehensive requalification examinations during the training cycle. NRC plans to seek stakeholder comments and subsequently issue guidance in a written venue yet to be determined. INPO expressed some question over the need to provide such guidance because there is no clear evidence that any such examination compromises occurred, and moreover, applicants could be trusted to maintain secrecy over examination content if they are similarly to be trusted to operate power plants.

Both the NEI and PROS representatives agreed that NRC should convey information to the training industry on the topic of examination integrity, validity, and compromise in examination development as a forum for educating the industry in examination "good practices."

### **Miscellaneous Issues**

1. The NRC discussed the simulator rule and stated that some plants were using the simulator, as allowable, for reactivity manipulations.
2. The NRC discussed several changes it is considering to the structure of the initial license examination. Such changes had been discussed earlier with the NRC/NEI Initial Operator Licensing task force during its last semi-annual meeting in February, 2002. Among the changes is a proposal to reduce the initial license Reactor Operator (RO) written examination from 100 to 75 test items while proportionately sampling knowledge in the

required systems, procedures, and administrative procedures at the current sample plan percentages. The Senior Reactor Operator (SRO) exam would include a separate, additional SRO-only exam component of 25 items and would be scored as a combined 100 [75+25=100] point examination.

Operating test changes include the following: Use of Job Performance measures only in Category A (Admin) of the exam with an 80 percent passing score for that section; or a combining of Categories A and B (Systems) in conjunction with changes to the grading and scoring structure of Category C simulator scenarios.

The NRC asked INPO whether it saw any need to update the Knowledge and Abilities Catalogs, NUREGS 1122 and 1123, which contain knowledge topics from which examination test samples are drawn. INPO stated that it was not aware of any need to change the Catalogs and thus felt no basis to make any such recommendations.

No other items were discussed. The meeting was adjourned at approximately 3:30 P.M.

Attachments: As stated

required systems, procedures, and administrative procedures at the current sample plan percentages. The Senior Reactor Operator (SRO) exam would include a separate, additional SRO-only exam component of 25 items and would be scored as a combined 100 [75+25=100] point examination.

Operating test changes include the following: Use of Job Performance measures only in Category A (Admin) of the exam with an 80 percent passing score for that section; or a combining of Category A and B (Systems) in conjunction with changes to the grading and scoring structure of Category C simulator scenarios.

The NRC asked INPO whether it saw any need to update the Knowledge and Abilities Catalogs, NUREGS 1122 and 1123, which contain knowledge topics from which examination test samples are drawn. INPO stated that it was not aware of any need to change the Catalogs and thus felt no basis to make any such recommendations.

No other items were discussed. The meeting was adjourned at approximately 3:30 P.M.

Attachments: As stated

ACCESSION #: ML021160697

<b>OFC</b>	<b>IOHS/IOLB/DIPM</b>	<b>IOHS/IOLB/DIPM</b>	<b>IOLB/DIPM/NRR</b>	<b>DIR/DIPM/NRR</b>
<b>NAME</b>	GUsova	DTrimble	TQuay	BBoger
<b>DATE</b>	06 /06 /02	06 /06 /02	06 /13/02	06/ 13/02

**OFFICIAL RECORD COPY**

## 2002 NRC/INPO TRAINING COORDINATION MEETING LIST

APRIL 10<sup>TH</sup> 2002

ATTENDEE NAME	ORGANIZATION	PHONE	E-MAIL
George M. Usova	NRR/DIPM/IEHB	301-415-1064	GMU@NRC.gov
Ted Quay	NRR/DIPM/IEHB	301-415-2402	TRQ@NRC.gov
David Trimble	NRR/DIPM/IEHB	301-415-2942	DCT@NRC.gov
Richard Conte	NRC/Region I	610-337-5183	RJC@nrc.gov
Robert C. Evans	NEI	202-739-8101	RCE@NEI.org
Fred Guenther	NRR/DIPM/IEHB	301-415-1056	SXG@NRC.gov
Brian O'Donnell	INPO	770-644-8628	O'DonnellBC@INPO.org
Kent W. Hamlin	INPO	770-644-8503	HamlinKW@INPO.org
Dale Powers	NRC/Region IV	817-860-8195	DEP@NRC.gov
Samuel L. Newton	INPO	770-644-8390	newtonsl@INPU.org
Philip N. McCullough	INPO	770-644-8213	mcculloughpn@inpo.org
John Munro	NRR/DIPM/IEHB	301-415-1097	JFM@NRC.gov
Bill Dean	NRR/DIPM	301- 415-1004	WMD@NRC.gov
Gregory Twachtman	McGraw-Hill	202-383-2166	gregory_twachtman@plahs.com
Bob Meyer	PROS	815-406-2172	robert.meyer@exeloncorp.com

ATTACHMENT 1

## 2002 NRC/INPO Training Coordination Meeting

### Agenda

TIME	TOPIC	INPO/NRC
8:00 AM	<b>Introductions and Organizational Updates</b>	INPO and NRC
8:30 AM	<b>Technical Specifications</b>  1. ACAD guidance documents referenced in Licensing Basis documents	NRC and INPO
9:30 AM	<b>SAT/Accreditation Issues</b>  1. Update on the Accreditation Improvement Initiative 2. INPO plans related to the coordination of accreditation visits and plant evaluations 3. Training warning flags and recurring accreditation weaknesses	INPO INPO INPO
10:30 AM	<b>Initial Examination Process</b>  1. Licensed operator test bank	NRC and INPO
11:00 AM	<b>Requalification Issues</b>  1. Recent problems noted in licensed operator requalification 2. Experience with the operator requalification significance determination process 3. Guidance on examination overlap of test items, including limits on bank use, among crews	NRC NRC NRC and INPO
12:00	<b>LUNCH</b>	
1:00 PM	<b>Miscellaneous Issues</b>  1. Implementation of the simulator rule 2. NRC update on the Generic Fundamentals Examination 3. INPO role in the nuclear power plant fatigue survey and NRC concerns with the scope of the survey (as early as possible in the meeting) 4. Efforts to further improve the initial exam process. 5. Interest in reviewing/updating the licensed operator job/task analysis in order to update the knowledge and abilities catalog.	NRC NRC NRC and INPO NRC NRC and INPO
4:00 PM	<b>Summary and wrap-up</b>	INPO and NRC

**ATTACHMENT 3**

INPO DOCUMENTS

SEE ADAMS DOCUMENT: ACCESSION NO. ML 021190005

March 26, 2002

NOTE TO: D. Trimble

FROM: G.M. Usova

SUBJECT: **WHITE PAPER: PROMOTING INTEGRITY AND VALIDITY IN NRC  
REQUALIFICATION EXAMINATIONS**

## **Introduction**

This paper addresses and discusses *test and measurement* principles and practices that promote the integrity and validity of requalification examinations in general. More recently, concerns have risen over the number of test items that can be repeated from exam to exam among crews undergoing the same requalification training program. In particular, this paper outlines the issues involved and sets forth a recommended policy.

## **Requalification Examinations: Promoting Integrity and Validity**

The fundamental purpose of the biennial written examination and annual operating examination during the requalification training programs, as required in 10 CFR 55.59, is to assure continuation of safe operator knowledge and performance. The biennial written examination is utility-developed and is predicated upon the specific learning needs identified within each respective utility SAT-based program. Requalification training is focused in nature, aimed at the continuing training needs identified within the utility SAT program; moreover, its scope of training is narrower than that of initial license training and is constrained by a shorter time frame (nominally @ six weeks) allowed by such requalification training.

## **Validity and Integrity**

Both the NRC initial license and requalification examinations measure operator safety important knowledge and performance, and, as with good examination practices, both should be valid measures of knowledge and performance. For initial and requalification examinations to achieve validity, both must be developed and administered in accordance with established psychometric procedures; one such compilation of these procedures is those contained in the Examination Standards, NUREG-1021, Revision 8. More specifically, these Examination Standards provide explanation and guidance for attaining examination validity --- a cornerstone attribute for examinations. The three components of validity are (1) content, (2) operational, and (3) discriminant. It is the latter level of validity --- discriminant validity--- that is most often the primary issue of concern in making a license decision.

Examination instruments (written and operating tests) should be developed to incorporate discriminant validity so as to discern which applicants(s) do and which do **not** have the minimum level of safety significant knowledge and performance required to pass the examination. Passing the requalification examination permits the inference to be made that those who do pass will operate the plant in a safe manner, and conversely, those who do not pass, will not operate the plant in a safe manner. Therefore, it is important that any NRC license examination

be constructed and administered such that confident safety significant discriminations (validity) can be made for each individual taking the examination so as to permit a confident licensing decision to be made.

In this regard, examination validity is inextricably linked with examination integrity -- an essential attribute toward making a valid licensing decision. Examination integrity is the attribute that conveys the *overall soundness* of the developed examination to achieve its intended purpose. That is to say, the examination instrument is constructed under an array of essential measurement development criteria that promote its integrity (e.g., sample plan development, psychometrics, consistent levels of knowledge and levels of difficulty, test bank use, validation procedures, etc.). Thus, the integrity of the examination encompasses more than testing the required body of knowledge (content validity) in a manner that permits an effective discrimination (operational and discriminant validity) but also includes adhering to the many other essential test and measurement development criteria noted above. In sum, the integrity of any examination is inseparable from its validity. When integrity is threatened or compromised, validity must similarly be questioned and analyzed to ensure it remains intact.

#### **Examination Integrity and Potential for Compromise: 10CFR55.49**

For examinations to be valid measures of knowledge and operating performance, no activities should occur that would compromise the integrity of the examination; thus, it is essential that individual applicants receive no assistance or aid from any external sources for any test items or tasks *prior to or during* the development or administration of the examination.

#### **Discussion**

In the Statement of Considerations, dated April 30, 1999, the Commission concluded that facility licensees may not be aware that the requirements of 10 CFR55.49 covers more than just those activities directly involving the physical control and administration of the examination. In this regard, the Commission considered the integrity of an examination or test to be compromised if any activity occurs that could affect the equitable and consistent administration of the examination or test, regardless of whether the activity takes place before, during or after the administration of the exam or test. Thus, 10 CFR55.49 encompasses not only activities such as lapses in physical security but also activities that compromise the integrity or validity of the exam itself during its development and construction (e.g., biased or improper sample plan, unallowable test bank use, excessive psychometric flaws (e.g., as stated in ES-401), or improper validation procedures).

Although the above context applied to the initial examination, sound measurement practices apply equally to the requalification examination as well. In this connection, NRC has provided additional NUREG 1021 guidance on:

- Use of facility test item banks
- Development of new test items at higher cognitive levels
- Development of an integrated exam outline, i.e., using a systematic and random sampling process when preparing the written exam outline reduces the chance that the written exam is biased in favor of topics that were emphasized during training.

- Use of old exams, quizzes, JPMs and simulator scenarios
- Moreover, NUREG-1021, Rev. 8, ES 201, Attachment 1, item 2, "Other Considerations," states that license applicants should not be able to predict or narrow the possible scope or content of the licensing exam based on the facility licensee's exam practices (other than those authorized by this NUREG or in writing by the NRC).

For example, SECY 96-206, dated September 25<sup>th</sup>, 1996, stated that overuse of bank items limits and reduces discrimination between competent and deficient license applicants and reduces exam integrity. Use of the same exam outline can also jeopardize exam integrity by increasing exam content predictability. For instance, if the license applicants can limit the pool from which test items are drawn, they are likely to study only from that pool and could exclude items covering other important job skills and knowledge. In other words, when the broader area of content is precluded from the test sample, then the validity inferences that would normally be made upon test performance, in fact, could not longer be confidently made.

### **Proper Testing Principles and Practices**

The use of proper testing development and administration practices can contribute to sound, valid examinations; similarly, improper practices can threaten or compromise examination integrity. Therefore, examination developers should engage in good acceptable test design, development, and administration processes so as to avoid any practices that call integrity or validity into question.

It is well-recognized that in a SAT-based requalification training program, learners should expect to be tested, in large part, upon the content covered during that training; however, those learners should not know nor be informed of, with undue specificity, the actual items that will be asked of them on the exam. Examples of activities that would adversely affect examination integrity are the following:

1. Telegraphing or announcing actual test content that would appear or **not** appear on the examination *prior* to the examination.
2. Directly instructing applicants (coaching) on specific test items content that knowingly will appear on the examination.
3. Improper validation procedures.

In particular, applicants (students) in the requalification training cycle should **not** be exposed to any test items nor play any role in validations *before* they take their own examination. Validation is a process of improving the test item(s) *before* administration to ensure that the items are technically, psychometrically, and editorially sound. Validation typically involves a tryout -- a process where samples of representative personnel actually take the exam and review it for technical and psychometric soundness. Validation experts agree that the tryout is to be done using a "sample" of a comparable group to that taking the test (Plumlee (1974) Office of Personnel Management). Also, Mallory (1987) in the Training and Development Handbook, states that validation or "tryout uses a limited number of trainees who are typical of the target population." However, in *no instance* should members

of the actual test population, itself, be used for pre-validation purposes; exposure to the very items or similar test items that a validating crew, itself, would take would, in turn, distort and compromise the integrity and validity of the examination results for that validating crew.

When this occurs, the examination's ability to test operator knowledge could become biased by short-term knowledge gained through the examination validation. In sum, the examination instrument could lose its ability to measure operator knowledge independently and objectively, resulting in falsely skewed positive results for those operators who were exposed to similar items that appeared on their own examinations.

Summarily, although the requalification program is more finite and narrow in its scope of content coverage than that training content covered in the initial examination, the requalification exam should, nonetheless, intend to discriminate between those who *have* from those who *have not* mastered the requalification program content. When test applicants are exposed to actual or similar items --- prior to the exam --- that they will be tested upon, then their scores will likely be higher than they would have been in the absence of that exposure, and thus, discriminant validity is somewhat reduced.

4. Excessive test item duplication between examinations and among crews undergoing the same training and testing processes OR excessive test bank use

A test bank is a collection of past validated test items which can be used as a basis for drawing items to reappear on future examinations. A policy that allows some percentage of test bank items to be used on future examinations is acceptable. Allowing applicants some limited predictability of test content coverage within the available test bank permits a learning value to accrue through test bank study, thereby promoting the retention of testable content which may improve subsequent job performance. However, such a policy should *never* allow applicants to know *which* specific items will appear on the examination. The deliberate uncertainty in knowing which or any actual test items that will appear on any given examination will "compel" the motivated applicant to study the entire test bank domain in order to benefit from this limited predictability. Moreover, when applicants know that some items on an examination will come from sources beyond the test bank, then applicants will likely strive to prepare from the broader area of knowledge discussed above.

What constitutes excessive test item duplication and what determines excessive test bank use, i.e., the number of past test bank items that can be reproduced on any given future examination, are matters of judgement, which can be influenced by the bank size, frequency of use of the test bank during a period of time, and examination development cost considerations.

### **Problems with Excessive Test Item Duplication and Overlap**

When high percentages of test item duplication takes place (e.g., > than 50 percent), the integrity of the examination may be under question. Discriminant validity, a part of the examination's integrity, is lessened; specifically, test item discrimination, as an essential element of validity, is in jeopardy when an excessively high percentage of items are duplicated from previous examinations.

Empirical data supports the 50 percent item duplication limit. For instance, data gathered from the Generic Fundamental Examination from 1992 - 2000 and from over its 2,800 test items showed that a 50 percent limit on test bank usage coupled with a mix of modified and new item development yielded a discriminating examination. Thus, the mix of (1) bank items previously seen (2) modified, similar to previously seen items and (3) new, not previously seen items were key elements toward maintaining the overall integrity of the examination.

In no instance should the entirety of the bank ever be used as the sole basis for developing the examination since this would establish and guarantee the domain of test content coverage and likely limit applicant exam preparation exclusively to the domain of the test bank. Arguably, there should always be some percentage of examination change or novelty via new item development so as to "compel" applicants to study beyond the finite realm of the test bank. Thus, infusion of new and modified items on any examination contributes to that examination's integrity and discriminant validity.

Moreover, successive administrations of the same or closely similar biennial requalification examination to different crews undergoing the same requalification training program raise the potential for compromising examination integrity (CFR 55.49). When examination integrity is in question, then the validity of the examination to achieve its intended purpose is similarly in question. In general and for all tests, to assume that the capability for items within an examination to discriminate, over time, in the same manner as those items discriminated on the first and second examinations are naive. Item repeatability or duplication that exceeds 50 percent of its items from examination to examination or from past examinations is unacceptable measurement practice.

### **Potential for Examination Leaks**

Examination leaks, i.e., communicating test content among candidates between examinations over time, is a potential occurrence. The potential for examination "leaks," that might occur between exam administrations to different crews will lessen the discriminant validity of the exam over time. Of course, if all examinees were administered the same examination at the same time, this type of security "leak" would be moot. Signed security agreements or security measures that only entail verbal caveats not to divulge or discuss the examination contents, do not guarantee examination security. It is naive to expect full compliance with either approach. Educational research points to an overwhelming body of evidence showing that the education and training community, at-large, have and continue to experience instances of examination breeches even when security measures are taken. Also, national studies conducted over 20 years have revealed strong evidence that academic dishonesty is widespread, e.g., 80-90 percent of test takers report engaging in dishonest practices when the pressures to perform well on examinations are high (Haines, Diekhoff, LaBeff, and Clark (1986); Jendreck (1989)).

Whether test banks are open or closed is largely immaterial. The difference in item exposure is one of degree--not kind. Closed banks have had some exposure since those items were likely seen in either training quizzes or examinations at some earlier point in time. Some individuals hold the misperception that it is acceptable to readminister an identical examination to separate crews that have been separated by more than one week in time. This view and practice is seemingly based upon two assumptions --- (1) that leaks would not occur among crews and (2) that human memory is fleeting and cannot recall information beyond a two-week time period. The former assumption has been addressed above and the later assumption is

unsupported by educational and psychological research. In fact, educational research points to human memory that can retain and recall certain test items over an indeterminate amount of time. Particularly in the case of requalification examinations in which the period of time is relatively short, measured in weeks, makes the recall of information easier and the potential for leaks possible. As Anastasia (1979) has stated regarding tests and retests:

If the interval between retests is fairly short, the examinees may recall many of their former responses. In other words, the pattern of right responses is likely to recur through sheer memory.

So as to mitigate the “recency effect,” i.e., the improved recall of information most recently perceived, guidance in NUREG 1021 states:

When items are repeated, they should be selected in a distributed manner and approximately equally over all previous exams so as to reduce predictability ( ES-601. E.3.b).

Moreover, in this connection, INPO’s document 86-025 states: “Sufficient questions should be developed so no two groups of individuals are given the same examination (p.13).”

### **Recommended Policy**

Regarding requalification program examination development and test item repeatability among crews undertaking the same requalification program, *no more than 50* percent of items appearing on any single examination should be repeated from any previous comprehensive or combination of comprehensive requalification exams within the same requalification training cycle (either one or two year cycle); conversely, this also means that any exam administered to any single crew should consist of 50 percent new or different items. Utilities are encouraged to develop test banks so that items comporting with requalification training program learning objectives can be exchanged within examinations. When bank items are to be repeated, they should be distributed across multiple past examinations approximately equally.

Moreover, administering any **same** biennial requalification examination to different crews undergoing the **same** requalification training is considered unacceptable. Identically administered examinations should not be readministered to any subsequent separate crew; this would seriously question the discriminant validity of the exam.

Regarding the requalification retake examinations, similar criteria used for examination development should apply: the retake examination should consist of at least 50 percent new or different items and the balance of the exam should be a drawn from and distributed across all previous examinations approximately equally.

### **Summary**

1. Utilities should engage in positive testing practices when developing and administering requalification examinations.

2. Recommended Policy - Examinations should not repeat more than 50 percent of test items from any comprehensive or combination of comprehensive requalification exams within the same requalification cycle (either one or two year cycles). This means, therefore, that any examination must consist of 50 percent new or different items.
3. Administering the same examination to more applicants simultaneously reduces the need to develop additional exams and saves resources. For example, utilities could develop two examinations, administering one examination to crews 1, 2, and 3 and the second to the remaining crews, or some other similar combination, to gain resource efficiencies and maintain exam integrity.

## References

Allen, Mary and Yen, Wendy (1979), Introduction to measurement theory. Monterey, CA: Brooks/Cole Publishing Company.

Alreck, Pamela and Settle, Robert (1985), The survey research handbook. Homewood, Illinois: Richard Irwin, Inc.

Anastasi, Anne (1976), Psychological testing, 4th. ed. New York, NY: Macmillan Publishing Company.

Copperud, Carol (1979), The test design handbook, Englewood Cliffs, NJ: Educational Technology Publications.

Isaac, Stephen and William, Michael (1990), Handbook in research and evaluation. San Diego, CA: EDITS Publishers.

Oppenheim, A.N. (1992), Questionnaire design, interviewing, and attitude measurement. Saint Martin's Press, Inc.

Haines, V. J., Diekhoff, G. M., LaBeff, E. E., & Clark, R. E. (1986). College cheating: Immaturity, lack of commitment, and the neutralizing attitude. *Research in Higher Education*, 25, 342-354.

Jendreck, M. P. (1989). Faculty reactions to academic dishonesty. *Journal of College Student Development*, 30, 401-406.