U.S.NRC

UNITED STATES NUCLEAR REGULATORY COMMISSION

Protecting People and the Environment

DRO-ISG-2023-01

Operator Licensing Programs

Interim Staff Guidance

10 CFR Part 53

March 2026

INTERIM STAFF GUIDANCE

OPERATOR LICENSING PROGRAMS

10 CFR Part 53

DRO-ISG-2023-01

PURPOSE

The U.S. Nuclear Regulatory Commission (NRC) staff is providing this interim staff guidance (ISG) to facilitate staff reviews of applications for an operating license (OL) or combined license (COL) under Title 10 of the *Code of Federal Regulations* (10 CFR) Part 53, "Risk-Informed, Technology-Inclusive Regulatory Framework for Commercial Nuclear Plants" (Part 53). The guidance in this ISG supports the staff review of the portion of such applications related to the operator licensing examination program. This ISG provides guidance for reviews of examination programs that are tailored to the specific role of operating personnel at commercial nuclear plants other than research and test reactors (RTRs). Specifically, it addresses the review and approval of both initial and requalification examination programs for senior reactor operators (SROs), reactor operators (ROs), and generally licensed reactor operators (GLROs), as provided for by the requirements in Subpart F, "Requirements for Operation," of Part 53. This ISG also addresses proficiency programs for SROs and ROs under Part 53. Because operator licensing programs might need to begin in advance of facility licensing, this guidance is written to address both the applicants for and holders of facility OLs and COLs under Part 53; these are referred to as "applicants" and "licensees," respectively, throughout this document for the sake of brevity.

This guidance may also facilitate the NRC staff review of OL applications for non-large light-water power reactors (non-LLWRs) under 10 CFR Part 50, "Domestic Licensing of Production and Utilization Facilities," or of COL applications for non-LLWRs under 10 CFR Part 52, "Licenses, Certifications, and Approvals for Nuclear Power Plants," as appropriate to the design challenges posed by new or novel facility designs. Establishing such an operator licensing examination program may require the applicant to request exemptions from certain requirements of 10 CFR Part 50, 10 CFR Part 52, or 10 CFR Part 55, "Operators' Licenses," that the staff would then need to review against the applicable exemption criteria of 10 CFR 50.12, 10 CFR 52.7, or 10 CFR 55.11, "Specific exemptions," respectively. The NRC staff will use the guidance in this ISG to inform its consideration of (1) requests by 10 CFR Part 50 and 10 CFR Part 52 applicants and licensees for exemptions from the requirements in 10 CFR Parts 50, 52, and 55 or (2) proposals to use alternative methods to those described in NUREG-1021, "Operator Licensing Examination Standards for Power Reactors," or NUREG-1478, "Operator Licensing Examiner Standards for Research and Test Reactors," by applicants and licensees under 10 CFR Part 50 or 10 CFR Part 52.

BACKGROUND

The guidance in this ISG was developed alongside the Part 53 rulemaking. Key rulemaking documents, including the final rule language and stakeholder comments, can be found at Regulations.gov under Docket ID NRC-2019-0062.

RATIONALE

This ISG was developed to meet near-term guidance needs to support the NRC staff's review of commercial nuclear plant applications with respect to risk-informed, technology-inclusive operator licensing examination programs. NUREG-1021 implements operator licensing standards and regulations as specified in 10 CFR 55.40, "Implementation," for large light-water reactor (LLWR) designs. NUREG-1021 was designed to be used for developing initial operator licensing examinations in conjunction with the applicable NUREG-series knowledge and ability catalogs. Similarly, NUREG-1478 implements operator licensing standards and regulations for RTRs. The examination methods and procedures described in these documents are based largely on the jobs and tasks that are performed by personnel at operating LLWRs and RTRs and the concepts of operations and staffing models for these facilities.

Commercial nuclear plants licensed under Part 53 will likely use different technologies and employ different operating and staffing models than LLWRs. Therefore, the NRC staff anticipates that the jobs and tasks performed by operating personnel at these plants will differ from those for operating personnel at LLWRs and RTRs. As a result, the knowledge, skills, and abilities (KSAs) that operating personnel at commercial nuclear plants licensed under Part 53 will need to perform their duties and the examination methods needed to evaluate those KSAs are also expected to differ from those in NUREG-1021 and NUREG-1478. For example, the staff anticipates that examination programs developed for commercial nuclear plants licensed under Part 53 will be of smaller scope and format than those in NUREG-1021 or NUREG-1478, given the expected reduced reliance on operating personnel to maintain safe plant operations at these facilities.

As a result, the staff determined that the existing examination guidelines could not be applied at commercial nuclear plants licensed under Part 53 without exemptions and substantial deviations from the guidance in NUREG-1021 and NUREG-1478. To ensure regulatory clarity and efficiency, the staff determined that new guidance is needed that accounts for the wide range of technologies, operating concepts, and staffing models expected to be employed at non-LLWR plants licensed under Part 53 or 10 CFR Part 50 or 10 CFR Part 52. The staff had two goals in developing the alternative guidelines:

(1) Enable applicants and licensees to leverage the results of human factors engineering design activities, such as job and task analysis, to identify the KSAs operating personnel need to have at their facility (or facilities) and to use that as the basis for developing examination standards for licensing competent operating personnel.

(2) Establish reliable guidelines for examination program development at these plants that are based on the best available knowledge from research and expertise on the measurement of knowledge and abilities.

As discussed in this ISG, applicants and licensees will be able to propose an examination program for licensing personnel at their facility that is based on or tailored to the role of operating personnel at their own facility, in lieu of using examination methods that were developed for LLWRs. This ISG offers guidance for the NRC staff to review the adequacy of examination programs for verifying the competency of operating personnel licensed under them. The NRC staff also recognizes that licensees and applicants may scope their proposed examination programs to reflect the guidance in this ISG.

APPLICABILITY

This ISG is applicable (1) to all operator licensing examination programs, and changes to those programs, submitted as part of applications for OLs and COLs under Part 53 and (2) to support the staff's review of exemptions for operator licensing examination programs for OL and COL applications under 10 CFR Part 50 or 10 CFR Part 52.

GUIDANCE

1.0 Development of a Knowledge, Skills, and Abilities List

The NRC staff should ensure that applicants for or holders of an OL or COL under Part 53 develop a design- or site-specific KSA list, or both, using a systems approach to training (SAT) process as described in 10 CFR 53.725(c) and 10 CFR 55.4, "Definitions." The staff should ensure that the applicant or licensee includes all aspects of job performance—including but not limited to tasks important to safe plant operation and the fundamentals of plant operations—in the job and task analysis of the licensed operator training program required for an SAT-based training program. The NRC gives guidance on KSA list development in DRO-ISG-2023-04, "Facility Training Programs," and NUREG-0711, Revision 3, "Human Factors Engineering Program Review Model," issued November 2012.

The NRC staff should ensure that the applicant or licensee screens the KSA list to identify those tasks and associated KSAs important to safe plant operation (as defined in section 1.3 of DRO-ISG-2023-04) and tasks related to the foundational theory of plant operations.

Applicants and licensees may perform a criticality analysis (discussed in the next section) to determine which KSAs require testing in the licensing examinations. Other types of analyses may also be used. Alternatively, deterministic criteria may be used to determine which KSAs require testing in licensing examinations. If the applicant or licensee uses cutoff values (e.g., subject-matter expert (SME) ratings) for the inclusion or exclusion of KSAs, the basis for that methodology should be documented, along with a list of excluded KSAs. Figure 1.0 depicts the expected steps in KSA list development for examination programs, including those performed as part of the separate training program.

Criticality Analysis

A criticality analysis determines the essential tasks related to doing a job that require assessment for operator licensing. Tasks that are essential are more important to be assessed for operator licensing. This analysis is a useful input to the examination development step of determining the content specification for each examination type (discussed later in this document) that determines which KSAs will be included in a specific examination type. In the criticality analysis, SMEs consider factors such as how frequently each task is done, the difficulty of each task, the importance of the tasks, and the risk involved. An acceptable approach is to obtain SME ratings on the tasks using Likert rating scales. A Likert rating scale is a numeric scale (typically five or seven points) that allows the individual to gauge to what extent they agree or disagree with a statement. For a criticality analysis, raters may judge importance (e.g., low versus high importance), frequency of use (e.g., low to high frequency), and so forth. The NRC staff should ensure that, once the data are collected, the applicant or licensee has compiled the data in a criticality matrix that displays the list of tasks and their ratings for each factor.

Acceptance Criterion:

The NRC staff should ensure that, if a criticality analysis is performed, the applicant or licensee compiles the data in a criticality matrix displaying the list of tasks and their ratings for each factor [Criterion 1.1].

Expected KSA List Coverage

The NRC staff should ensure that the KSA list that results from the task analysis in accordance with DRO-ISG-2023-04 at minimum reflects the knowledge, plant design, operator actions, administrative tasks, facility procedures, and emergency plan responsibilities that are relevant to the licensed operator function at the applicant's or licensee's facility. Items (1)–(3) give further details about these various areas:

(1) The KSA list should reflect knowledge of both the plant design and operator actions needed to achieve plant safety. The following examples are for illustrative purposes and are not meant to represent a complete or applicable list for each reactor design or operator KSA requirement. Instead, these examples detail some of the KSAs that would be expected to result from the task analysis:

- safety-related, non-safety-related, and safety-significant (if applicable) structures, systems, and components (SSCs) for the nuclear plant that the operator interacts with (e.g., controls, monitors), to include those SSCs required for radioactive material handling, radiation monitoring, and post-accident monitoring

- characteristics associated with the SSCs (e.g., connections and interdependencies, support systems, control and monitoring, behavior during normal operations and transient conditions, monitoring methods, technical specifications)

- site-specific procedures and policies (e.g., shift turnover, normal operations, administrative tasks, radiological control, plant startup) that pertain to operators

- abnormal and emergency plant events, the associated emergency and abnormal operating procedures for responding to those events, and implementation of the site emergency plan

- KSAs that address the SSCs necessary to meet the defense-in-depth (DID) requirements of the plant as they relate to licensing-basis events, increases in the cumulative plant risk, and beyond-design-basis events (this includes non-safety-related, safety-significant SSCs that provide DID to the safety-related SSCs of the plant and their associated characteristics, as applicable; if necessary, new KSAs should be developed to address plant safety DID, like the first two items on this list)

- inherent, passive, or automatic safety-significant SSCs, their characteristics, and the associated operator actions, procedures, and SSCs that provide DID to the inherent, passive, or automatic safety-significant SSCs

(2) The KSA list should reflect fundamental, theoretical knowledge associated with the specific design and operations of the reactor. Operators should have a fundamental understanding of the technologies, materials, and processes of the reactor design,

including an understanding of applicable reactor theory, thermodynamics, and chemical theory topics. Operator understanding of the physical and chemical phenomena is important to understand system behavior and to diagnose and analyze plant events.

(3) The KSA list should reflect important administrative responsibilities associated with the SRO and GLRO roles:

– conditions and limitations in the facility license

– facility operating limitations in the technical specifications and their bases

– facility licensee procedures required to obtain authority for design and operating changes in the facility

– radiation hazards that may arise during normal and abnormal situations, including maintenance activities and various contamination conditions

– assessment of facility conditions and selection of appropriate procedures during normal, abnormal, and emergency situations

– procedures and limitations involved in determining various internal and external effects on reactivity

– fuel handling facilities and procedures

Acceptance Criteria:

The NRC staff should ensure that the KSA list adequately addresses the following, as applicable:

- safety-significant SSCs that the operator interacts with and the associated characteristics of those SSCs [Criterion 1.2]
- safety-significant SSCs that provide plant safety DID and the associated characteristics of those SSCs [Criterion 1.3]
- inherent, passive, or automatic safety-significant SSCs; their characteristics; and the associated operator actions, procedures, and SSCs that provide DID to the inherent, passive, or automatic safety-significant SSCs [Criterion 1.4]
- site-specific normal operating procedures that pertain to operators [Criterion 1.5]
- abnormal and emergency plant events, associated procedures, and implementation of the site emergency plan [Criterion 1.6]
- theoretical knowledge of reactor theory, thermodynamics, and chemical theory, as applicable, associated with the technologies, materials, and processes of the reactor design [Criterion 1.7]
- knowledge of administrative topics, including notifications, radiological controls, maintenance controls, technical specifications, equipment control, conduct of operations, and the emergency plan [Criterion 1.8]

Necessary Documentation

The NRC staff should ensure that the development process for the examination program KSA list is documented. The documentation should detail the steps of the process; other processes or documents that provide input, terms or acronyms; and the resulting list of KSAs applicable to the site operator licensing examinations. The documentation should also clearly describe the methods for determining testable and nontestable KSAs. Those KSAs screened out as untestable and the associated bases should be captured in the documentation.

The initial KSA list will be reviewed and retained by the NRC staff and will be maintained for the operational lifetime of the facility by the licensee. The NRC staff should ensure that the documentation related to the KSA list gives information on the process used to maintain the KSA list, to include responsibilities, frequencies, input, and output paths. The staff should ensure that the applicant or licensee has a process to update the KSA list as needed due to site changes (e.g., equipment, procedures, training).

The KSAs should be grouped in a logical fashion in the KSA list to facilitate the formation of examinations through the selection of KSA types (i.e., to more easily sample KSAs that represent the breadth and scope of the KSA list and to meet the examination developer's goals for question significance and rigor).

Acceptance Criterion:

The NRC staff should ensure that (1) the KSA list details the development process, the determination of KSA testability (e.g., for nontestable KSAs, lack of criticality or untestable), and KSA list maintenance, and (2) the KSAs are organized in a logical format to facilitate use and review [Criterion 1.9].

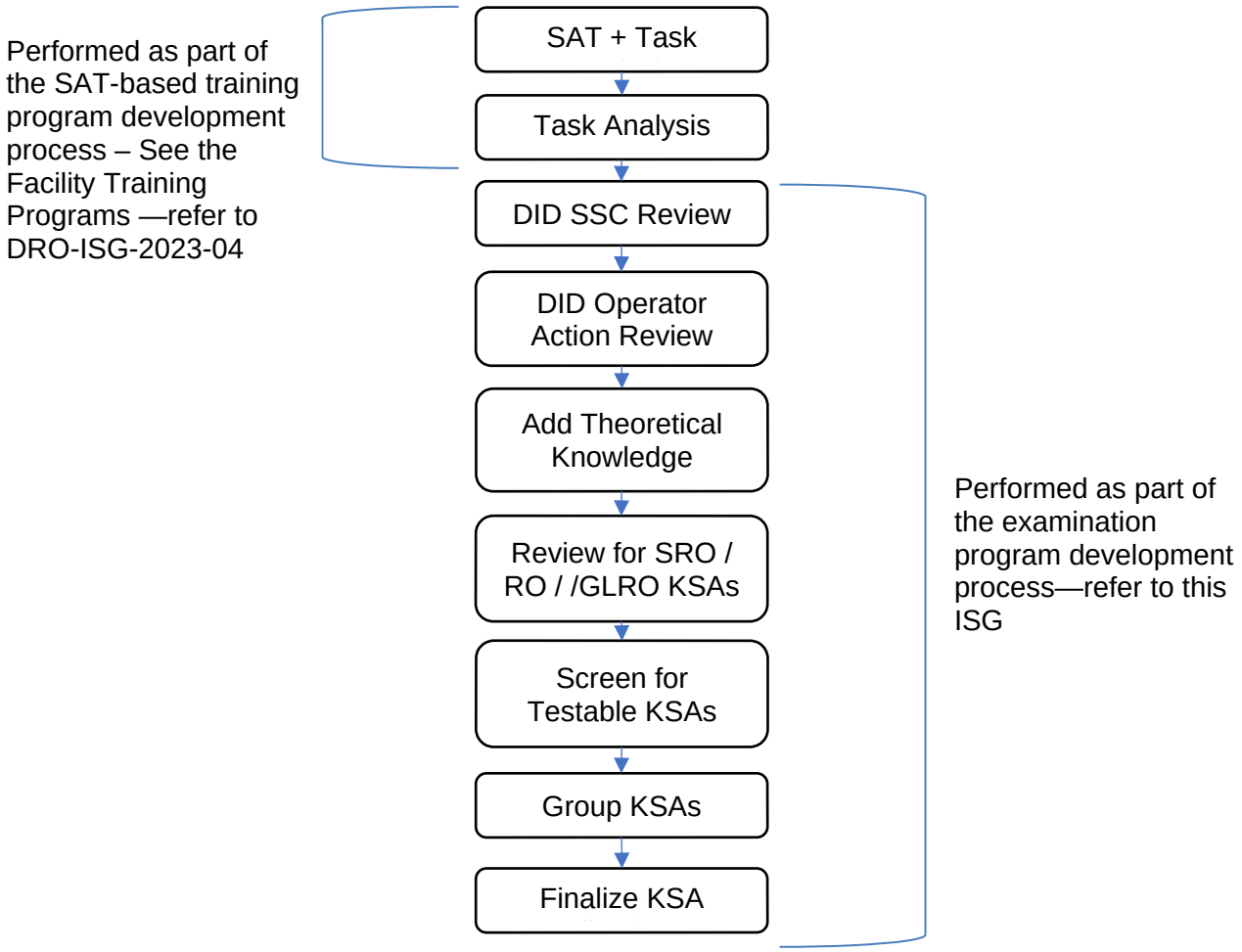

Figure 1.0 Overview of KSA Development Process

2.0 Development of Operator Licensing Test

The NRC staff should ensure that a comprehensive measurement strategy that covers all the aspects of performance that the job task analysis determined to be essential is used to determine operator competency. The first step is to develop an overall operator licensing strategy or "Test Plan." The Test Plan will typically include multiple assessment measures[1] as necessary to cover all the tasks and KSAs for a job. Once the Test Plan is established, the individual tests or examinations can be developed.

Test Plan Rationale

The Test Plan rationale is a document that describes the comprehensive strategy for determining operator qualification for the job. This includes a description of how the tasks and their respective KSAs described in the job task analysis will be measured. The NRC staff should ensure that documentation includes a mapping of tasks and KSAs to measurement tools (e.g., knowledge examination, simulation examination). The staff should ensure that Part 53 applicants or licensees have determined the appropriate assessment measures for their facility. The methods outlined in NUREG-1021 are not required to be used.

Acceptance Criteria:

The NRC staff should ensure the following:

- The applicant or licensee provides documentation of the assessment measure(s) to be used for each task and how the measures are aligned conceptually with the KSAs being measured. For example, an applicant or licensee may provide a matrix that describes each task (and the KSAs making up those tasks) and pairs each task with a measurement strategy. Sufficient detail needs to be provided to allow the NRC staff to determine adequacy of coverage [Criterion 2.1].

- The applicant or licensee provides an overall measurement strategy that may include a variety of knowledge and performance examinations (e.g., written, computer-based, simulator-based, walkthrough, or other types of examination) [Criterion 2.2].

- An instructional designer[2], industrial/organizational psychologist[3], or a human factors specialist[4] has determined the relevance of the measurement approach for the particular KSAs [Criterion 2.3].

- The applicant or licensee provides documentation on KSAs not covered by any assessment measures [Criterion 2.4].

---

[1] Assessment measures are formal, consistent, and dynamic measuring tools used to assess the ability of individuals to perform in the field.

[2] An instructional designer is someone who designs work-related training and development programs to improve individual skills or organizational performance as well as analyzing organizational training needs and training effectiveness.

[3] An industrial/organizational psychologist is someone who applies principles of psychology to areas such as policy planning; employee testing, selection, training, and development; and organizational development and analysis.

[4] A human factors specialist is someone with the training and experience in how to design and assess whether objects, facilities, and environments are optimized for the human's well-being and overall system performance by applying theory, principles, and data regarding the relationship between humans and respective technology.

Measurement Strategy and Tools: General Information

The NRC staff should ensure that the Test Plan documents that the measurement strategy and tools align conceptually with the inherent characteristics of the KSA being measured. While KSAs differ depending on the tasks, general categories include knowledge and understanding, cognitive skills (analysis, synthesis, evaluation, problem solving), and practical skills (observable behaviors). At the same time, many examination strategies and methods exist. The most common are written, computer-based, or performance-based (e.g., walkthrough or simulation). It is essential for the chosen methods to have the correct conceptual underpinnings necessary to accurately capture the KSAs of interest.

Knowledge: Knowledge is learned information, often categorized into four types. Factual knowledge consists of terminology and discrete facts. Conceptual knowledge consists of categories, theories, principles, and models. Procedural knowledge includes knowledge of a technique, process, procedure, or method. Metacognitive knowledge includes knowledge of oneself and knowledge of various learning skills and techniques.

- *Testing methods: Knowledge is typically assessed with oral, written, or computer-based tests.*

Skill: A skill is an acquired or learned proficiency to perform a certain task or role. Skills can be psychomotor, behavioral, cognitive, or some combination thereof. Examples of skills include physically controlling a vehicle (e.g., driving a car, riding a bicycle, landing an aircraft), athletics (e.g., wrestling, swimming, tennis), oral and written communication, solving problems, programming computers, various construction activities, and so on.

- *Testing methods: While some skills may be assessed with oral, written, and computer-based tests, other skills can only be assessed using a behavioral measure in which some aspect of the examinee's behavior is observed and recorded (e.g., through a simulation or in a natural setting).*

Ability: An ability is a capacity to apply knowledge and skills simultaneously to complete a task or perform an observable behavior.

- *Testing methods: While some abilities may be assessed with oral, written, and computer-based tests, other abilities can only be assessed using a behavioral measure in which some aspect of the examinee's behavior is observed and recorded (e.g., through a simulation or in a natural setting).*

Format Specifications: Format specifications provide the format of items (e.g., tasks or questions) and responses (e.g., responding to multiple choice questions, performing or demonstrating skills for a complex task) and the type of scoring procedures. Format specifications describe the rationale for the chosen format that includes consideration of the format's validity (i.e., the conceptual underpinnings).

Complex item formats include performance assessments and simulations. Specifications for complex items describe the domain from which the items (e.g., tasks or questions) are sampled, the component of the domain to be assessed, and critical features of the items for replication in alternate forms.

Test takers demonstrate their abilities or skills to perform tasks in settings that closely resemble situations encountered on the job (e.g., assemble a part, write a report, start a pump) through performance assessments. Since a performance test includes only a sample of tasks, the NRC staff should ensure that the specifications clearly define the critical dimensions or competencies to be measured (e.g., skills and knowledge, cognitive processes,[5] context[6] for performing the tasks) and how the tasks align with those competencies. When tasks are designed to elicit complex cognitive processes, specifications need to include detailed analyses of the tasks and scoring criteria to provide necessary evidence of validity.

Simulations are like performance assessments and may substitute when actual task performance is not feasible or desirable (e.g., because it would be too costly or dangerous). Specifications for simulations describe the domain of activities to be covered by the tasks, critical dimensions of performance reflected in each task, and format considerations, such as the number or duration of the tasks.

Multimethod Examinations: The central assumptions of using multiple methods are (1) that each method has its strengths and weaknesses and (2) that a combination of diverse examinations is necessary to benefit from the relative strengths of each individual examination. Research shows that using combinations of examinations can lead to high degrees of accuracy in predicting job performance.

In general, to optimize examinations for determining whether an individual is qualified, the information needed to make that determination should first be identified and then the method that is most appropriate for obtaining that information should be selected. A particular method should be used for what is best suited. For example, a written examination is best suited for assessing fundamental knowledge, comprehension, and application of knowledge. A performance-based examination using the actual work environment or a simulation is best suited for assessing an individual's skill at operating a piece of equipment. The staff anticipates that examinee performance will differ for each method if the methods actually measure different aspects of each individual's qualifications.

Examination Content Specification

Along with the overarching Test Plan giving the comprehensive measurement approach, the NRC staff should ensure that the applicant or licensee provides a detailed content specification or content framework for each type of test or examination method (e.g., written, computer-based, simulation) in the examination program. The content specifications should describe the specific content, skills, and diagnostic features of the domain or construct that the examination type covers (e.g., "knowledge of valves," "knowledge of emergency response"). The content specifications should be guided by the job task analysis. The content specification process may use varying methods to determine KSA inclusion on an examination (e.g., ranking KSAs, applying values to KSAs, grouping KSAs according to subject matter or task criticality, safety significance). The Test Plan might measure the most important work behaviors or KSAs, or a few that are prerequisite to others, or a smaller set of KSAs used to predict critical outcomes (e.g., accidents). Applicants and licensees should include a sampling method or plan that

[5] Cognitive processes are those associated with human cognition. Cognition can be defined as those mental processes involved in attending, making sense, reasoning, remembering, and making decisions.

[6] Context, for the purposes of this guidance, can be defined as the circumstances that form the setting for an event and in terms of which it can be fully understood and assessed.

discusses how to select specific KSAs for use on an administration of a specific examination, also referred to in this document as an "examination instance."

For each test or examination instance, the documentation should describe the specific KSAs included on the test, the number of KSAs, the number of questions per KSA, and the results of the sampling method used to select those KSAs. If more than one question or performance assessment item per KSA is allowed, the documentation should describe the basis for this allowance.

Documentation

The NRC staff should ensure that the applicant or licensee develops documentation with all necessary test specifications (e.g., the amount of time allowed for the testing, directions for the test takers, procedures to be used for administration, materials for test takers to use, scoring and reporting procedures, description of necessary hardware and software) and the basis for these test specifications.

Acceptance Criteria:

The NRC staff should ensure the following:

- The applicant or licensee provides documentation of the procedures used to determine KSA inclusion and exclusion on the examination [Criterion 2.5].
- For each test or examination method, the applicant or licensee provides content specifications that clearly describe the jobs and tasks covered in the test methods and the specific content, including the following:
  - the sampling methods and plan for determining inclusion on a specific examination instance (for example, the test might measure the most important work behaviors or KSAs, KSAs that are prerequisite to others, or a smaller set of KSAs used to predict critical outcomes such as accidents) [Criterion 2.6]
  - a decomposition of the tasks into the knowledge, skills, cognitive processes, abilities, attitudes, or behaviors that are included in the test method, and those that are excluded from the test method [Criterion 2.7]
  - documentation on test length (specifying the number of questions allotted to different content areas based on the job analysis) [Criterion 2.8]
  - specification of the type of personnel (examinee population) for which the test is intended (i.e., RO, SRO, or GLRO) [Criterion 2.9]
  - specification of intended uses of the test method (i.e., a clear statement of the purposes and applications for which the test is intended) [Criterion 2.10]
  - format specifications that provide the testing strategy and format of items (e.g., written, computer-based, performance-based) and describe the rationale for the formatting that includes consideration of the validity of the format and not simply ease of use [Criterion 2.11]

- evidence and supporting documentation that the test strategy aligns appropriately depending on the distinction of knowledge versus skills [Criterion 2.12]

- The applicant or licensee provides documentation with all necessary test specifications and the bases for these test specifications [Criterion 2.13].

Test Development

The following guidance, in combination with the guidance in NUREG-1021, Revision 12, issued September 2021, appendix A, "Overview of Generic Examination Concepts," is intended to help the NRC staff determine whether applicants and licensees followed adequate test and examination development processes.

For any new tests or examinations, the NRC staff should ensure that the applicant or licensee follows an effective test development process and describes this process. Effective test development typically includes generating items (for written tests) or tasks (for performance tests), reviewing items using statistical analysis or qualitative review of items (or both), assembling the test, establishing validity of the overall test, and developing documentation and test manuals.

The NRC staff should ensure that test development processes are effective and do the following:

- Generate more questions or tasks than are needed for one version of the test.

- Review the pool of items for content quality, clarity, and any construct-irrelevant issues:

  - Ideally, new test items or tasks are given to people who represent (to the extent possible) the target test population. Depending on the sample size, statistical analysis can assess some psychometric properties, such as item difficulty and the ability of an item to distinguish between different levels of test taker knowledge or skill (refer to the validity sections of this document for detailed information on establishing the validity of the entire test).

  - Qualitative review of items may occur. This may include using a structured interview or think-aloud protocol to demonstrate that the desired cognitive processes are elicited. For extended response items or performance tasks, test developers should provide example responses that represent the various scoring levels.

- After review, the test developer assembles items into one or more test forms (i.e., versions) in accordance with the content and length specifications. This includes documenting that the items selected meet specifications for content and psychometrics. Content reviews may entail replacing items that are too similar or that have other negative characteristics (such as providing answers to other items). Multiple test forms can be generated.

- Establish the validity of the overall test (refer to the validity section of this document).

- Verify compliance with 10 CFR 53.780(b), 53.780(c), 10 CFR 53.815(b)(3)(iv), or 53.815(b)(3)(v), as applicable. The NRC staff considers the following list of general

categories to be the minimum set of KSAs needed to safely perform operator duties. Ensure the elimination of any of these categories is adequately justified.

- reactor theory and thermodynamic principles
- plant systems and components
- reactivity management and manipulations
- radiation control and safety
- emergency, abnormal, and normal operations
- administrative requirements and conditions of the facility license
- technical specifications

The NRC staff should ensure that Test Plan documentation includes the following:

- Description of what important work behaviors, activities, and worker KSAs are included in the domain, how the content of the work domain is linked to the test development process, and why certain parts of the domain were or were not included in the testing method. This information should be based on accurate and thorough information about the work, including analysis of work behaviors and activities, responsibilities of the job incumbents, and the KSAs prerequisite to effective performance on the job.

- Evidence that demonstrates that the test adequately samples and is linked to the important work behaviors, activities, and worker KSAs. The documentation of the processes used to develop the test is the primary evidence that the scores will be sufficient to predict on-the-job performance. The NRC staff should understand that the sufficiency of the match between the test and the work domain is a matter of professional judgment based on the evidence provided. However, the documentation of the process should include the criteria and the rationale for the inclusion and exclusion criteria.

- A rationale for content sampling (i.e., selecting the particular content to include) that is based on the professional judgment of the testing professional and an analysis of work that details important work behaviors and activities, important components of the work context, and the KSAs needed to perform the work, if sampling is required. True random sampling of the content of the work domain is usually not feasible or appropriate. Consequently, a systematic, quasi-random sampling procedure should be implemented to ensure that appropriate content sampling and examination security are maintained. The rationale underlying the content sampling should be documented in a Test Plan specifying which KSAs are to be measured by which assessment methods. If sampling is not performed because the content domain is small enough to be tested on every examination, the applicant or licensee should state that and explain how examination security concerns related to predictability are addressed.

Acceptance Criteria:

The NRC staff should ensure the following:

- The applicant or licensee describes the test development process, to include generating items; reviewing items using statistical analysis or qualitative review of items; evidence of overall test validity; and the conditions, directions, procedures, and materials for taking the test [Criterion 2.14].

- The applicant's or licensee's description of the test development process specifies how it will ensure that, at a minimum, topics from the following list of general categories of KSAs will be sampled during examinations. Elimination of any of these categories is adequately justified.
  - reactor theory and thermodynamic principles
  - plant systems and components
  - reactivity management and manipulations
  - radiation control and safety
  - emergency, abnormal, and normal operations
  - administrative requirements and conditions of the facility license
  - technical specifications [Criterion 2.15]
- The applicant's or licensee's Test Plan documentation includes the following:
  - what important work behaviors, activities, and worker KSAs are included in the domain, how the content of the work domain is linked to the testing procedure, and why certain parts of the domain were or were not included in the testing method [Criterion 2.16]
  - evidence that demonstrates that the test adequately samples and is linked to the important work behaviors, activities, and worker KSAs, including the criteria and the rationale for the inclusion and exclusion criteria [Criterion 2.17]
  - a rationale for a systematic, quasi-random content sampling procedure based on the professional judgment of the testing professional and an analysis of work that details important work behaviors and activities, important components of the work context, and KSAs needed to perform the work; if sampling is not performed due to small content domain size, an explanation of how examination predictability is addressed is included [Criterion 2.18]

3.0 Examination Validation

The NRC staff should ensure that the applicant or licensee has a plan to validate the new examination's ability to assess operator KSAs and task performance capability.

Rationale for Validity Plan

The NRC staff should ensure that the applicant or licensee describes the type of evidence collected to support test validity. The degree and type of evidence for the test will depend on SMEs. SME judgment should be based on the following factors:

- how the test results will be interpreted
- how the test results will be used
- the type and degree of harm that might result if the test is not valid

- the probability that incorrect inference can be corrected before harm occurs
- availability of research on similar examinations taken by similar personnel for similar reasons
- sufficiency of qualified individuals in validation samples
- practicality in data collection
- availability of appropriate criteria for conducting criterion-based validation studies

Acceptance Criterion:

The NRC staff should ensure that the applicant or licensee demonstrates that the validity plan will be implemented before the first administration of a test developed under the examination program. As part of the examination program, applicants and licensees should monitor and collect data following the validity plan over training cycles. If monitoring identifies a significant deviation from estimated risk associated with the type or degree of harm that may result or from the estimated probability that incorrect inference can be corrected before that harm occurs, the NRC staff should ensure that applicants and licensees have processes to document the deviation, refine the validity plan, and reassess the examination [Criterion 3.1].

Validity of the Test

As part of the validity plan, the NRC staff should ensure that the applicant or licensee provides evidence that the test will work as intended. The validity of a test can be demonstrated through one or more kinds of validity for each test, including content validity, predictive validity, and concurrent validity. The validity plan should require content validity. Ideally, predictive and concurrent validity would be demonstrated as well; however, the practical constraints stemming from the small examinee population size inherent to the nuclear domain may limit the ability of applicants and licensees to provide this information (refer to the section on criterion validity for more information).

Demonstrating the validity of a test of human knowledge and capabilities is a rigorous and difficult process. In particular, if the KSA/construct of interest is not directly observable, then multiple and different types of validity is demonstrated to establish that the actual KSA/construct of interest is being measured. For example, attention is not directly observable but the behaviors associated with attention are. Thus, a test developer could create a test that measures the directly observable behaviors associated with attention under strictly controlled conditions. However, for the test to be valid, the test developer would need to demonstrate that the test "behaves" in the way that it should. It should have content validity (i.e., cover the entire realm of relevant information). It should predict future outcomes. This is predictive validity. For our example, an attention test should predict future academic scores because attention is important for learning. Concurrent validity is similar to predictive validity but the test scores and the criterion scores are assessed in the same session. Test scores are the predictor scores, and the criterion scores are the outcome scores. For instance, the researcher could measure the person's attention and reading comprehension scores in the same session. Through statistical analysis, we can assess whether the attention scores predict the reading comprehension scores. Other types of validity exist. These are the most relevant for this document. Below, more information is provided on these validities.

Content Validity

A test with content validity has a strong link between the content of the test and key work behaviors and activities. The aim is for the test to measure relevant KSAs required for the job while not including irrelevant knowledge and skills.

A content validation study provides evidence that the testing method samples the worker KSAs and/or work behaviors and activities that are necessary for the job. SMEs are an integral part of the test development team.

The NRC staff should verify that the validity plan documents the SMEs involved in test development, including the following:

- The SMEs have in-depth knowledge of the work, the responsibilities of the job incumbents, and/or the KSAs necessary for effective job performance. Documentation of the qualification of each SME should include a description sufficient for the reader to infer that the SME was knowledgeable about factors such as shift, location, type of equipment used, and software and hardware.

- SMEs have defined the work domain and identified the key work behaviors, activities, and worker KSAs. SME judgments are important for both defining the content of the test and the methods used for testing.

- Documentation includes details on how SME judgments were used to determine content validity. That is, if SME judgments were ratings on the match between test content and work requirements, documentation includes the criteria and rating procedures used for each aspect. Evidence of rating standardization should also be included.

Criterion Validity

There are two general approaches to estimate the correlation between test scores and criterion scores (e.g., a measure of job performance): predictive and concurrent validation strategies. Predictive validation strategies are the most accurate but can present practical and ethical concerns, as explained below. Concurrent validation strategies are more practical procedures to assess validity.

Predictive Validation Strategies

From a scientific standpoint, predictive validation is the simplest and most accurate method to determine validity. However, this approach is not practical or realistic for real-world application, as it requires the population in the validity study to be similar to the population of individuals who would apply for the job.

The general process to complete a predictive validation study is as follows:

(1) Step 1: Obtain test scores from a group of applicants and hire all the applicants.

(2) Step 2: At some point in time, obtain performance measures of those hired and correlate these performance measures with the prior test scores.

In other words, a predictive validation study requires random hiring, which sets up an ethically risky position of selecting some people who are very likely to fail (and potentially create costly and/or dangerous situations in the process of failing). However, this is the strategy necessary for a true predictive validity approach. Consequently, a true predictive validity approach is not suitable for demonstrating test validity because public health and safety are paramount.

Concurrent Validation Strategies

Concurrent validation involves (1) obtaining both test scores and criterion scores in an intact, preselected population, such as existing employees, students accepted into college, pilots who have completed a certain level of training, and (2) correlating the two sets of scores (i.e., correlating the test scores with the criterion scores). The fundamental difference between this approach and the predictive validation approach is the lack of a random sample. That is, the population of workers at a plant, students accepted into college, pilots with several certificates, etc. is narrower and with much less variability than the population who initially applied for but were not selected or did not choose to continue through the hiring process. From a statistical standpoint, this generates a restriction of range in the scores, which, in turn, impacts the correlation between the test scores and the criterion measures. Thus, the correlation between test scores and criterion scores in a preselected (not random) sample will likely not be the same as for the broader population in general. While test theory suggests that concurrent validation would underestimate this correlation between test scores and criterion scores in the broader population, studies suggest concurrent validation studies are often sufficiently similar to predictive validation studies.

If a concurrent validity study is used in a setting that is very selective (e.g., when only those with high test scores are included), coefficients will generally underestimate the validity of the test, and the underestimation may be severe. Considering the practical limitations placed on the applicant or licensee in demonstrating criterion validity at the time of the first administration of these examinations, the NRC staff acknowledges for the purposes of review that the validity plan may treat criterion validity as an ongoing goal as more data are collected on the examinations.

Interpreting Validity Coefficients

Correlations between test scores and criterion measures could range in absolute value between 0.0 and 1.0. The higher the criterion-related validity, the more accurate the estimate based on the predictor measure can be. In practice, a good, carefully chosen test is unlikely to show a correlation coefficient with an important criterion greater than 0.5, and validity coefficients greater than 0.3 are not common in applied settings. The levels of criterion validity achieved by most tests rarely exceed 0.6 or 0.7. The validity coefficient in and of itself does not provide a complete measure of the effects of the tests on decisions.

Evaluating a Predictive or Concurrent Validation Study

This section outlines the information needed to evaluate criterion validity. However, as noted for the application approval process, predictive validity is very unlikely to constitute an acceptable approach for the reasons discussed previously.

A predictive or concurrent validation study includes a written description of the study method that describes the following:

- a set of participants who are representative of the desired population (i.e., operators)
- well-defined, reasonable standards of quantitative assessment
- procedure that describes how the data are collected
- statistical test results

Acceptance Criterion:

The NRC staff should ensure that the applicant or licensee has a program that ensures that the tests will work as intended and that the tests will demonstrate one or more kinds of validity (i.e., content validity, predictive validity, and/or concurrent validity) [Criterion 3.2].

4.0 Scoring Specifications

Determining the scoring specification is a key aspect of examination development.

The scoring specification should be criterion referenced. Criterion-referenced scoring focuses on the absolute score relative to a defined level of competence or "criterion." The criterion-referenced method is appropriate for operator licensing testing. Note that norm-referenced scoring, which ranks a score for an individual within a distribution of scores or compares it to the average performance of test takers in a reference population (e.g., based on age, job classification), is not appropriate for operator licensing.

The scoring specification should describe how each test item is scored and how the item scores are combined to produce an overall score.

For procedures where ratings are based on scorer observation, scoring specifications should describe scorer qualifications, scorer training, identification and resolution of rating discrepancies, and steps to eliminate bias in judgments.

Cutoff Scores

When cutoff scores are used, the staff should ensure that the applicant or licensee clearly defined the passing score and included a rationale. Criteria for reasonable cutoff scores should include the following:

- The score reflects the minimum qualification necessary to perform the job in a safe and efficient manner.
- The score represents criteria consistent with a job analysis.
- The score includes evidence that a thoughtful approach was used to determine the cutoff score and involved both SME judgments about the cutoff score and evidence of the validity of the test overall.
- The score uses a content validity approach (often referred to as "criterion-referenced cutoff scores," in which the criterion is SME judgment):
  - The three approaches for evaluating the test items are as follows:
    - SMEs examined the test items and rated the likelihood that a barely qualified or barely competent person could answer that question correctly. The cutoff score is the average of the ratings.
    - SMEs may have rated the relative importance and relevance of each item.
    - SMEs may have rated the capability of barely qualified individuals to identify the correctness of the possible responses (i.e., options in a multiple-choice test).
  - Note that a lack of interrater agreement is a disadvantage of all three methods. In organizations where scores are based on the subjective judgment of an evaluator, evaluators should be limited to a single or constant group of evaluators

for all candidates to reduce error among evaluators and to increase reliability and validity within the scoring process.

- The score does not use a norm-referenced approach. Norm-referenced methods to establish cutoff scores are based on an entire distribution of test scores (e.g., mean + one standard deviation). The norm-referenced approach is not acceptable for situations requiring establishment of minimum competency, subject-matter mastery, licensure, and other situations.

Acceptance Criterion:

The NRC staff should ensure that the applicant or licensee clearly defines appropriate, criterion-referenced cutoff scores and their bases. If a cutoff score greater than or equal to 80 percent is used, then no additional basis needs to be provided, as NUREG-1021 has determined 80 percent to be an acceptable cutoff score for power reactors. If a cutoff score less than 80 percent is used, then the basis for that cutoff score is provided [Criterion 4.1].

5.0 Reliability of the Test

Reliability is a necessary, but insufficient, requirement for validity. If a test is reliable, there will be a low amount of measurement error in the test. If the trainees repeated the test, their results would be relatively the same. This particular type of reliability is called “test-retest reliability.”

Assessing individuals at two different time points even with use of alternate forms of the test is not always feasible or desirable. Thus, reliability is usually assessed by examining the internal consistency of the items. When examining internal consistency, the researcher assesses how the items of the test relate to one another. Split-half reliability correlates one half of the test to the other half. It is still possible that test items of poorer quality could be grouped together while the better items are grouped together through random chance. To minimize this item-related bias, typically the Cronbach’s alpha—average of all possible split-half reliabilities—is reported. Statistical software can quickly generate all possible groupings and generate correlations between two groups and calculate the average of all these correlations. Essentially, each split-half reliability is analogous to a dart on a dart board; if all the darts strike close to the center, then the test is internally consistent or reliable.

The NRC staff should ensure that applicants and licensees include the following information on how test reliability is determined, as appropriate:

- information on the reliability of a test (i.e., the test manual provides data adequate to permit judgment on whether scores are sufficiently dependable for the recommended use)

- sufficient evidence to support a thorough judgment of the reliability of a test (e.g., parallel form reliability, internal consistency and interitem relations, test-retest reliability, interrater reliability)

- findings that are adequate to justify use of a test for operator licensing-related purposes

- information confirming that the participants who supply the data used to compute reliability coefficients are appropriately suited to support operator licensing examination

development-related work and sufficiently described in the documentation to permit judgment of whether the data derived from them should apply

- confirmation that appropriate procedures for computing the reliability coefficients are used
- reliability data presented in the conventional statistical form of product-moment correlation coefficients and standard errors of measurement, with standard errors given for different levels of performance
- for tests with alternate versions, data allowing for comparisons between the versions
- evidence reported for internal consistency (or interitem correlations)
- confirmation that the test manual provides evidence of the stability of test performance over time (i.e., test-retest reliability)

Acceptance Criterion:

The NRC staff should ensure that the applicant or licensee provides information on the reliability of tests and supports that information with a reasonable description of the methodology used to determine its adequacy [Criterion 5.1].

6.0 Test Manual

A comprehensive test manual should accompany each examination program that is submitted for programmatic approval under 10 CFR 53.780(b), 53.780(c), 10 CFR 53.815(b)(3)(iv), or 53.815(b)(3)(v), as applicable.

Acceptance Criteria:

The NRC staff should ensure the following:

- The applicant or licensee provides a comprehensive test manual that covers each type of test. The manual includes information on the purpose of the test, how the test was developed, and how to administer and score the test [Criterion 6.1].
- The test manual includes the following:
  - a clear statement of the purposes and applications for which the test is intended [Criterion 6.2]
  - clearly defined constructs the test intends to measure, the groups for which the test is intended, and the specific application of the test [Criterion 6.3]
  - justification of the relevance of the test content for the constructs to be measured [Criterion 6.4]
  - a summary of research findings for the test [Criterion 6.5]
  - dates of any revisions to the manual and any additional statistical tests for the revised versions [Criterion 6.6]

- information on revisions to the test instruments associated with technology changes [Criterion 6.7]
- any limitations of the test [Criterion 6.8]
- a description of the expertise required to administer and interpret the test [Criterion 6.9]
- the time needed to administer the test [Criterion 6.10]
- the time allowed for testing [Criterion 6.11]
- a description of all—
  - procedures to be used for administration (and allowed variations) [Criterion 6.12]
  - materials for test takers to use [Criterion 6.13]
  - scoring and reporting procedures, and a description of the necessary scoring hardware and software [Criterion 6.14]
- clear and complete instructions to administer the test and interpret the results properly [Criterion 6.15]
- all necessary forms and any necessary aids to interpreting the test results [Criterion 6.16]
- case descriptions showing how test scores can be interpreted [Criterion 6.17]
- a clearly defined domain and discussion of the procedures used for sampling from that domain [Criterion 6.18]
- discussion of any potential threats to the valid use of the test scores [Criterion 6.19]

• For a specific test instance, documentation will include the following:

- name of the test, authors of the test (by name and position), publisher of the test and the date published, and the existence of alternate forms [Criterion 6.20]
- dates of any revisions to the test and any additional statistical tests for the revised versions [Criterion 6.21]
- information on revisions to the test associated with technology changes [Criterion 6.22]
- a report on the evidence of the validity of the specified use of the test [Criterion 6.23]
  - The documentation should avoid referring to correlations between items and total test score as evidence of validity [Criterion 6.24].

- a description of criterion variables, adequacy of criterion variables, and aspects of criterion performance that are not adequately reflected in the criterion measure(s) when criterion validity is reported [Criterion 6.25]
- statements expressing relationships that are presented in quantitative terms so that the reader can tell how much confidence to attach to them (e.g., "the correlation with variable X is 0.55" rather than "the test correlates substantially with variable X") [Criterion 6.26]

- Additionally, for computer-based testing, the test manual also includes the following:
  - information on any necessary computer software installation [Criterion 6.27]
  - information on operation of the software [Criterion 6.28]
  - provision for technical support [Criterion 6.29]

7.0 Additional Characteristics of High-Quality Test Materials

Some characteristics of high-quality test materials apply specifically to written and to computer-based tests. These are described in the sections that follow.

High-Quality Written Test Materials

NUREG-1021, Form 4.2-2, "Question Development Checklist," can be used as a job aid during the process of developing and reviewing written examination questions. NUREG-1021, Appendix B, "Examples of Written Examination Questions," gives examples that illustrate psychometric errors to avoid.

Additional characteristics of high-quality written test materials include the following:

- correctly formulated items (refer to NUREG-1021, Appendix B)
- standardized test items
- the items, test booklet, answering scales, and answer form follow user-centered[7] design principles in a way that errors can be avoided when completing the test
- clear and complete instructions for the test
- a scoring system that is consistent and error resistant, including the following:
  - an objective scoring system
  - if the test is being scored by raters or observers, a clear and complete system for assessment and observation

High-Quality Computer-Based Test Materials

[7] "User-centered" refers to a design that is based on looking at the way that people will use something and considering what they will do with that design.

High-quality computer-based test materials include the following:

- description of scoring (e.g., computerized or an objective scoring system)
- documentation describing whether the test items are standardized or adaptive
- if adaptive scoring, documentation describing the following:
  - the decision rules
  - how the test meets the provisions for uniformity in licensing
- error-resistant user interface
- clear and complete instructions for the administrator and examinee
- items correctly formulated (refer to NUREG-1021)
- sufficiently secure test

8.0 Other Examination Program Considerations

The following sections of NUREG-1021 contain additional items that are applicable to any examination program:

- section 1.3, “Examination Security”
- section 2.1.G, “Guidelines for Freezing Plant Procedures”
- applicable portions of section 2.2, “Applications, Medical Requirements, and Waiver and Excusal of Exam and Test Requirements,” for specifically licensed operators
- section 2.3, “Reviewing and Approving Operator Licensing Initial Examinations,” for specifically licensed operators
- applicable portions of section 5.1, “Issuing Operator Licenses and Post-Examination Activities,” for specifically licensed operators
- section 5.2, “Application Denials and Requests for Informal NRC Staff Review,” for specifically licensed operators
- section 5.3, “Maintaining, Changing, and Renewing Operator Licenses,” for specifically licensed operators
- appendix A, “Overview of Generic Examination Concepts”

Acceptance Criteria:

The NRC staff should ensure that the applicant or licensee documents how it will meet examination security requirements under 10 CFR 53.780(d) or 10 CFR 53.815(d), as applicable. If the facility will have specifically licensed operators, it will also have procedures for preparing and submitting applications, including any waivers and excusals, and procedures for ensuring medical requirements are met. Facilities with specifically licensed operators will also have procedures for maintaining, updating, and renewing operator licenses [Criterion 8.1].

The NRC staff should also ensure that the applicant or licensee documents how examinees will be briefed on the examination process before examination administration, similar to the guidelines in section 1.2, "Guidelines for Taking NRC Examinations," of NUREG-1021 but tailored to the specific examination process developed for the facility [Criterion 8.2].

Preparing for Operator Licensing Initial Examinations

The NRC staff should ensure that the applicant's or licensee's examination program reflects the following, as applicable:

- The program will test a representative sample of the KSAs needed to safely perform RO and SRO or GLRO duties, to include both the examination methods and criteria to be used to assess passing performance, as required by 10 CFR 53.780(b)(1) and 10 CFR 53.780(c)(2)(i) or 10 CFR 53.815(b)(3)(iv) and 10 CFR 53.815(b)(3)(v), as applicable.

- The program will be approved by the NRC before its use in developing or administering initial examinations and requalification examinations for either ROs and SROs or GLROs as described under 10 CFR 53.730(g) or 10 CFR 53.815(b)(vi), as applicable.

- The program will be maintained according to the requirements of 10 CFR 53.1565, "Evaluating changes to programs included in licensing basis information."

- Prepared examinations for RO and SRO applicants will be made available to the NRC for review and approval in advance of their administration, as required by 10 CFR 53.780(b)(2). This should include the expected timeline for submittal of these prepared examinations based on the complexity and scope of the examinations.

- Sufficient advance notification will be given to the NRC to enable an NRC representative to be present during examination administration, as required by 10 CFR 53.780(b)(3), 10 CFR 53.780(c)(2)(ii)(B), and 10 CFR 53.815(b)(3)(v), or to administer the examination, as required by 10 CFR 53.780(b)(3). Because the NRC needs sufficient advance notice to ensure availability of staff resources to review and approve the examination as well as coordinate and observe examination administration, an acceptable timeframe for notification is 18 months in advance of the scheduled examination administration. Other timeframes may be acceptable depending on the complexity of the design and scope of the examination.

Acceptance Criteria:

The NRC staff should ensure that the applicant or licensee documents how it will ensure that the facility's examination program does the following:

- tests a representative sampling of the KSAs to safely perform licensed duties [Criterion 8.3]
- is approved by the NRC before use [Criterion 8.4]
- will be maintained [Criterion 8.5]
- has provisions for how the prepared examinations for RO and SRO applicants will be made available to the NRC for review and approval before administration, if applicable [Criterion 8.6]
- has provisions for giving sufficient advance notification to the NRC to enable an NRC representative to be present during examination administration and, as applicable, to also administer the examination [Criterion 8.7]

9.0 Simulation Facilities

Part 53 includes requirements about the use of simulation facilities for licensed operator training, experience requirements, and examinations. Such requirements could apply to both onsite and offsite simulation facilities and licensee-operated or third-party-operated simulation facilities, depending upon the individual circumstances of a given applicant or licensee. The following supplemental information can help inform the development of an applicant's or licensee's simulation facility for use in training, meeting experience requirements, or for the conduct of examinations consistent with the requirements in 10 CFR 53.780(e) or 10 CFR 53.815(e), as applicable.

The term "simulation fidelity" refers to the level of realism of the physical, psychological, and functional aspects. The following definitions apply to this discussion of human-in-the-loop[8] simulation in non-LLWR reactor technologies:

- Physical fidelity is the degree to which the simulator looks, feels, and is designed to replicate the actual environment.
- Functional fidelity is how well the simulator functions and provides stimuli to reflect the actual environment.
- Task fidelity is the replication of tasks and maneuvers executed by the user.
- Psychological-cognitive fidelity is how well the simulator replicates psychological and cognitive factors (e.g., communication, decision-making, and situational awareness).

Digital technologies, such as desktop computers, augmented reality, and virtual reality, use varying levels of simulation realism to represent task scenarios that reflect the actual job situation.

[8] Human-in-the-loop simulations involve humans taking part in interactive simulations in a manner that allows for human actions to influence the outcome of the simulation.

The NRC staff should ensure that, when using new simulation technologies, the applicant or licensee documents and demonstrates how the simulation provides a level of fidelity sufficient to assess the intended knowledge and skills as required by 10 CFR 53.780(e) or 10 CFR 53.815(e), as applicable.

To best replicate scenarios for non-LLWR facilities, simulation facilities should have the same cognitive requirements as the real environment. If a simulation facility is less cognitively demanding than the real environment, it may not be as useful a tool in training or examination.

Acceptance Criteria:

The NRC staff should ensure the following:

- If a simulation facility is used in examinations, the applicant or licensee documents how the simulation facility provides a level of fidelity and cognitive requirements sufficient to assess the intended knowledge and skills of operators [Criterion 9.1].
- If a simulation facility is used in examinations, the applicant or licensee documents policies and procedures to demonstrate compliance with simulator fidelity requirements under 10 CFR 53.780(e) or 10 CFR 53.815(e), as applicable [Criterion 9.2].

A crucial aspect of simulation-based assessment is the structure or the linking together of task-based content, scenario events, performance measures, and feedback. An integrated approach is essential to demonstrate the validity of a simulation-based examination. For each proposed simulation examination, the applicant or licensee will describe the purposes of the examination, the definition of the construct or domain measured, the intended examinee population, the measurement tools, and the interpretations for intended uses, consistent with methods for content determination and validation described in this document. This includes identifying the jobs and tasks, scenario events, metrics, and feedback collection, as described in the following sections.

Identifying the Jobs and Tasks

- Provide a clear description of the jobs and tasks covered in the examination and the specific content. Document in an examination plan the rationale for the tasks covered.
- Provide a task decomposition in terms of the knowledge, skills, cognitive processes, attitudes, or behaviors that are included in the examination and that are excluded from the examination (i.e., identify the scope of the examination).
- Specify the type of personnel (i.e., examinee population) for which the examination is intended.
- Specify the intended uses of the examination (i.e., a clear statement of the purposes and applications for which the examination is intended).
- Provide evidence or supporting documentation that the simulation strategy is appropriate based on the distinction of knowledge versus skills.

Acceptance Criteria:

The NRC staff should ensure that, if a simulation facility is used in examinations, the applicant or licensee provides content specifications that clearly describe the jobs and tasks covered in the simulation and the specific content, as follows:

- Tasks are decomposed into both those knowledge, skills, cognitive processes, attitudes, or behaviors that are included in the examination and those that are excluded from the examination [Criterion 9.3].
- The type of personnel (i.e., examinee population) for which the examination is intended is specified [Criterion 9.4].
- The intended uses of the examination (i.e., a clear statement of the purposes and applications for which the examination is intended) are specified [Criterion 9.5].
- Evidence and/or supporting documentation are provided showing that the simulation strategy is appropriate based on the distinction made between knowledge and skills [Criterion 9.6].

Identifying the Scenario Events

For any scenario that is part of a simulation facility examination, the staff should ensure that the applicant or licensee provided a list of scenario events that link to the overarching tasks. NUREG-1021, Examination Standard (ES)-3.3, "General Testing Guidelines for Dynamic Simulator Scenarios," and ES-3.4, "Developing Scenarios," contain examples of possible types of events. The applicant or licensee should include documentation describing how the scenario events give the examinee an opportunity to demonstrate the KSAs necessary to perform the task. Additionally, the applicant or licensee should include documentation related to acceptable qualitative and quantitative scenario attributes that, if met, would ensure that the overall scenario and individual scenario events are appropriately realistic and at a satisfactory difficulty level. Subsequent sections of this ISG offer considerations about scenario events and the examination criteria determination.

Acceptance Criterion:

The NRC staff should ensure that, if a simulation facility is used in examinations, the applicant or licensee provides policies and procedures to link scenario events to the operator tasks and to document that information for scenario-based examinations [Criterion 9.7].

Identifying the Metrics

For each scenario event, the applicant or licensee will provide event-based metrics to assess the degree to which the examinee achieved the task objectives. The precise event-based performance measures will assess how effectively the examinee is demonstrating the desired knowledge and skills. The metrics will be designed to facilitate ease of use by the examination administrator to avoid errors in grading.

Acceptance Criterion:

The NRC staff should ensure that the applicant or licensee provides policies and procedures to develop event-based metrics for scenario-based examinations [Criterion 9.8].

Feedback

The applicant or licensee should describe the process the examiner will use to integrate the performance results and provide feedback to the examinee on their performance in terms of scenario events that link to the job tasks.

10.0 Administering Operating Tests

The NRC staff should ensure that applicants and licensees have examination procedures similar to those in section 3.5, "Administering Operating Tests," of NUREG-1021. Qualified facility staff (e.g., certified instructors, licensed personnel) may administer the examinations and should follow procedures similar to those used by NRC examiners. For operator licensure, the applicant or licensee should propose a schedule that will protect examination security while providing an efficient license examination process. The applicant or licensee should inform the NRC of examination schedules. The applicant or licensee should support the administration of the examination by providing the following as needed:

- personnel to operate simulation facility equipment
- surrogate operators
- monitors
- approved examination materials
- examination administrators

The NRC staff should ensure that the applicant or licensee has measures to ensure that examiners behave in accordance with NRC examination and integrity codes of conduct, which should be documented as part of the facility examination program to ensure examination integrity, as required by 10 CFR 53.780(d) or 10 CFR 53.815(d), as applicable.

Applicants and licensees must retain examination administration records in accordance with the relevant requirements of 10 CFR 53.780, "Training, examination, and proficiency program," or 10 CFR 53.815, "Generally licensed reactor operator training, examination, and proficiency programs," as applicable, and the NRC staff should ensure that this is reflected in the operator examination plan.

Acceptance Criterion:

The NRC staff should ensure that the applicant or licensee documents how to administer the operating test portion of the examination while maintaining examination security and examiner codes of conduct to ensure that test integrity is maintained [Criterion 10.1].

11.0 Examination Program Change Management Process

The NRC staff should ensure that the examination program (1) documents the change management process and (2) specifies both which changes to approved examination programs can be made without prior NRC approval and which changes require prior NRC approval. Changes to the approved initial and requalification examination programs for ROs and SROs, and to the approved requalification examination program for GLROs, must be consistent with

the regulatory requirements of 10 CFR 53.1565. In general, the applicant or licensee should specify that changes require prior NRC approval if they do the following:

- Require an exemption from a regulation
- Require a change to technical specifications
- Negatively impact examination security or integrity
- Negatively impact consistent examination administration
- Negatively impact consistent examination evaluation

In general, the applicant or licensee should specify that the following changes do not require prior NRC approval:

- minor editorial changes (such as to correct typos or to provide clarification)
- administrative changes associated with test proctoring
- changes or deletions to validity determination (with the exception of content validity)
- refinements to the validity plan based on test program results
- updating SME lists
- documenting a new validation or reliability study using the same process as that previously approved, provided that it does not result in substantive changes to the examination program
- test manual updates in the areas of research findings, reporting and scoring procedures (such as hardware or software), forms, the validity report, software for computer-based training, or technical support information
- changes associated with section 7.0 of this guidance
- changes to guidance associated with freezing procedures
- changes related to simulation facilities except for changes that may affect their use for training, examining, or meeting experience requirements for ROs and SROs
- changes to simulation facility policies and procedures (provided that fidelity is maintained)
- updating evidence or documentation that the simulation strategy aligns with the distinction between knowledge and skills, provided that no other test changes are recommended
- changes to the feedback mechanism addressed in section 9.0 of this guidance

In general, changes other than those described above should be identified as requiring NRC approval before implementation to ensure that there is no impact on the consistent, valid administration of the examination program.

As an example of the proper implementation of an examination program's change management process, although adding an item to a KSA list may appear to be a conservative change that should not require prior NRC approval, a closer analysis may indicate that such an addition could impact the sample and Test Plan as it relates to the number of questions being asked. If such a KSA were added without evaluating these impacts, it could result in a skewed sample plan and, perhaps more importantly, KSAs being tested at a reduced frequency, thereby diluting the testing pool with a KSA of lesser importance. If a significant number of such KSAs were to be added, it might necessitate a revision to the number of items on the test to ensure that appropriate sampling is maintained.

Acceptance Criterion:

The NRC staff should ensure that the applicant or licensee documents the change management process for the examination program and specifies which changes to approved programs can be made without and which require prior NRC approval. This process is consistent with the requirements of 10 CFR 53.1565 and should also conform to the guidance of this section [Criterion 11.1].

12.0 Static Computer-Based Testing

Static computer-based testing is a type of examination that does not have interactive features but only allows multiple choice, selection on the screen, or text entry. This would not include computer-adaptive tests or any form of simulator. Although static computer-based testing may be permissible for examination, this document does not directly address this testing method. To use such a testing approach, the applicant or licensee would likely need to create similar guidance for examination development that can provide an equivalent testing measure to the guidance in this document.

13.0 Additional Guidance for Requalification Programs

Applicants and licensees are required under Part 53 to have requalification programs for the ROs and SROs or GLROs at their facilities. While requalification programs are generally subject to the same guidance as for initial examination programs in this document, the reviewer should consider additional elements of requalification programs.

For all requalification programs, the program must have a provision to ensure that any requalification examination failures are properly remediated and retested, and that a retake examination is passed before allowing the licensed operator to return to licensed duties per 10 CFR 53.780(c)(2)(ii)(C) or 10 CFR 53.815(b)(3)(v)(A). The retake examination would need to encompass all the portions of the examination that the licensed operator failed but would not need to include portions that the licensed operator passed. For example, if the licensed operator failed a 35-question written examination and passed the job performance measure (JPM) and simulator sections, the retake examination would consist of a 35-question written examination but would not need to include JPMs or simulator scenarios. Alternatively, if the licensed operator failed three out of five JPMs, the retake examination would need to consist of five JPMs, not three.

The NRC staff should ensure that, for applicants and licensees with ROs and SROs (i.e., facilities subject to the requirements of 10 CFR 53.760 to 10 CFR 53.795), the applicant or licensee establishes programs with the following programmatic aspects:

- Applicants and licensees are required by 10 CFR 53.780(c)(2)(i) and 10 CFR 53.780(c)(2)(ii)(C) to administer a requalification examination to all ROs and SROs during a period that is not to exceed 24 months between examinations. This time frame is limited by the duration of the requalification training cycle required under 10 CFR 53.780(c)(1)(i). Thus, the program must include sufficient provisions to facilitate this required examination process on a recurring basis.

- The program must provide for making prepared requalification examinations available for NRC review, as required by 10 CFR 53.780(c)(2)(ii)(A).

- The program must ensure that the NRC is notified of upcoming requalification examinations and has the opportunity to be present during administration, as required by 10 CFR 53.780(c)(2)(ii)(B). Because the NRC needs sufficient advance notice to ensure availability of staff resources to observe examination administration and coordinate that observation with the facility, an acceptable timeframe for notification is 12 months in advance of the scheduled examination administration. Other timeframes may be acceptable depending on the complexity of the design and scope of the examination.

- The program must provide for ensuring that a summary of examination results for each RO and SRO is promptly forwarded to the NRC after the completion of the requalification examination, as required by 10 CFR 53.780(c)(2)(ii)(D). Because the NRC needs to ensure that licensed operators are maintaining the requisite knowledge to safely operate the plant, this information must be provided to the NRC promptly to support that determination. If the program provides the results within 30 days, the NRC staff will accept that as being prompt. Other timelines may be proposed and will be evaluated on a case-by-case basis to determine if the timeline will support the NRC's need to ensure the continued competence of the currently licensed operators.

The NRC staff should ensure that, for applicants and licensees with GLROs (i.e., facilities subject to the requirements of 10 CFR 53.800 to 10 CFR 53.830), the applicant or licensee establishes programs meeting the following requirements of 10 CFR 53.815(b)(3)(v):

- A requalification examination must be administered to each GLRO within a recurring periodicity that is defined within the program.

- If the requalification examination periodicity is less than or equal to 24 months, the NRC staff does not need to perform further review. If the applicant or licensee proposes a periodicity exceeding 24 months, the NRC staff should ensure that the applicant or licensee provides bases for that periodicity, using the following insights:

  – the SAT process

  – operator performance trends

  – industry operating experience

  – changes in the aggregate experience level and turnover of GLRO staffing

  – significant changes to the design or operation of the facility

- The program must ensure that the NRC is notified of upcoming requalification examinations and has the opportunity to be present during administration. Because the

NRC needs sufficient advance notice to ensure availability of staff resources to observe examination administration and coordinate that observation with the facility, an acceptable timeframe for notification is six months in advance of the scheduled examination administration. Other timeframes may be acceptable depending on the complexity of the design and scope of the examination.

Acceptance Criteria:

The NRC staff should ensure the following:

- The requalification examination program has a provision to ensure that any requalification examination failures are properly remediated and retested, and that a retake examination is passed before allowing the licensed operator to return to licensed duties [Criterion 13.1].
- The requalification examination program designates an appropriate periodicity for the administration of requalification examinations. For ROs and SROs, this periodicity may not exceed 24 months. For GLROs, this periodicity may exceed 24 months with adequate provisions for informing the periodicity [Criterion 13.2].
- The requalification examination program has sufficient provisions to satisfy the additional requirements of 10 CFR 53.780(c)(2) or 10 CFR 53.815(b)(3)(v), as appropriate, that are discussed in this section [Criterion 13.3].

14.0 Proficiency Programs for Specifically Licensed Operators and Senior Operators

Specifically licensed operators and senior operators are required under Part 53 to be subject to a Commission-approved proficiency program. The applicant or licensee must include a program that is adequate to meet the following requirements of 10 CFR 53.780(g):

- Ensure that ROs and SROs will actively perform the functions of an RO or SRO, as appropriate.
- Maintain proficiency regarding shift functions.
- Maintain familiarity with plant status.
- Include those steps that will be taken to reestablish proficiency when it cannot be maintained.

Proficiency programs should clearly define the specific duties, locations, and amounts of time associated with each of the items in the preceding list. The specifics of each should be informed by considering the following:

- the facility-specific concept of operation
- the facility-specific staffing plan
- human performance considerations
- the full scope of duties assigned to the individual while on shift (e.g., collateral functions for maintenance, radiation protection)

Changes that reduce the scope or requirements of an approved proficiency program must be approved by the NRC before implementation in accordance with 10 CFR 53.1565.

While GLROs are required by 10 CFR 53.805(a)(4) and 10 CFR 53.815(g) to be subject to a proficiency program that is administered by the facility licensee, GLRO proficiency programs do not require prior approval by the NRC and, therefore, are not within the scope of this ISG.

Acceptance Criteria:

The NRC staff should ensure the following:

- The applicant or licensee provides a proficiency program for ROs and SROs that addresses the requirements of 10 CFR 53.780(g) and has appropriate justification for the adequacy of its provisions. Proficiency programs that are equivalent to, or more conservative than, the requirements of 10 CFR 55.53(e) and (f) may be considered acceptable without additional justification [Criterion 14.1].

- The applicant's or licensee's programmatic provisions for change control specifically disallow making changes to the approved proficiency program that reduce its scope or requirements without obtaining prior NRC approval as required by 10 CFR 53.1565 [Criterion 14.2].

15.0 Waivers for Generally Licensed Reactor Operators

Pursuant to 10 CFR 53.815(f), the requirements for a test or examination for GLROs may be waived in accordance with the facility's approved examination program. Therefore, the applicant or licensee should include appropriate criteria that may be used to waive the requirements for a test or examination. If the applicant or licensee is using requirements that are equivalent to, or more conservative than, those in 10 CFR 55.47, "Waiver of examination and test requirements," no further NRC staff review is required. Otherwise, the applicant or licensee should also provide a basis for these criteria describing how the criteria will ensure that the individuals are able to safely and competently operate the facility as required by 10 CFR 53.815(b).

Acceptance Criterion:

The NRC staff should ensure that the applicant or licensee includes criteria for waiving the requirement for a test or examination for GLROs, including a basis for those criteria as needed [Criterion 15.1].

16.0 Acceptance Criteria Summary

To approve an application for an OL or COL under Part 53, the NRC staff must determine, among other things, that the requirements in Subpart F are met for the design and technology under review. This determination should be based on whether the information in the application is sufficient to conclude that the acceptance criteria discussed in this guidance are met, as summarized below.

Summary of Acceptance Criteria		
Section 1.0		
1.1	If a criticality analysis is performed, the applicant or licensee compiles the data in a criticality matrix displaying the list of tasks and the ratings for each factor.	
The coverage of the KSA list should adequately address…		
	1.2	…safety-significant SSCs that the operator interacts with and the associated characteristics of those SSCs.
	1.3	…safety-significant SSCs that provide plant safety DID and the associated characteristics of those SSCs.
	1.4	…inherent, passive, or automatic safety-significant SSCs; their characteristics; and the associated operator actions, procedures, and SSCs that provide DID to the inherent, passive, or automatic safety-significant SSCs.
	1.5	...site-specific normal operating procedures that pertain to operators.
	1.6	…abnormal and emergency plant events, associated procedures, and implementation of the site emergency plan.
	1.7	…theoretical knowledge of reactor theory, thermodynamics, and chemical theory, as applicable, associated with the technologies, materials, and processes of the reactor design.
	1.8	…knowledge of administrative topics, including notifications, radiological controls, maintenance controls, technical specifications, equipment control, conduct of operations, and the emergency plan.
1.9	The KSA list details the development process, the determination for KSA testability (e.g., for nontestable KSAs, lack of criticality or untestable), and KSA list maintenance, and the KSAs are organized in a logical format to facilitate use and review.	
Section 2.0		
2.1	The applicant or licensee provides documentation of assessment measures to be used for each task and how the measures are aligned conceptually with the KSAs being measured. Sufficient detail needs to be provided to allow the NRC staff to determine adequacy of coverage.	
2.2	The applicant or licensee provides an overall measurement strategy that may include a variety of knowledge and performance examinations.	
2.3	An instructional designer, industrial/organizational psychologist, or a human factors specialist has determined the relevance of the measurement approach for the particular KSAs.	
2.4	The applicant or licensee provides documentation on KSAs not covered by any assessment measures.	
2.5	The applicant or licensee provides documentation of the procedures used to determine KSA inclusion and exclusion on the examination.	
For each test or examination method, the applicant or licensee provides content specifications that clearly describe the jobs and tasks covered in the test methods and the specific content, including…		
	2.6	…the sampling methods and plan for determining inclusion on a specific examination instance.
	2.7	…a decomposition of the tasks into the knowledge, skills, cognitive processes, abilities, attitudes, or behaviors that are included in the test method, and those that are excluded from the test method.
	2.8	…documentation on test length.

Summary of Acceptance Criteria		
	2.9	…specification of the type of personnel for which the test is intended.
	2.10	…specification of intended uses of the test method.
	2.11	…format specifications that provide the testing strategy and format of items and describe the rationale for the formatting that includes consideration of the validity of the format and not simply ease of use.
	2.12	…evidence and/or supporting documentation that the test strategy aligns appropriately depending on the distinction of knowledge versus skills.
2.13	The applicant or licensee provides documentation with all necessary test specifications and the bases for these test specifications.	
2.14	The applicant or licensee describes the test development process, to include generating items; reviewing items using statistical analysis and/or qualitative review of items; evidence of overall test validity; and the conditions, directions, procedures, and materials for taking the test.	
2.15	The applicant's or licensee's description of the test development process specifies how it will ensure that, at a minimum, topics from the following list of general categories of KSAs will be sampled during the course of examinations. Elimination of any of these categories is adequately justified: • reactor theory and thermodynamic principles • plant systems and components • reactivity management and manipulations • radiation control and safety • emergency, abnormal, and normal operations • administrative requirements and conditions of the facility license • technical specifications.	
The applicant's or licensee's Test Plan documentation includes the following…		
	2.16	…what important work behaviors, activities, and worker KSAs are included in the domain, how the content of the work domain is linked to the testing procedure, and why certain parts of the domain were or were not included in the testing method.
	2.17	…evidence that demonstrates that the test adequately samples and is linked to the important work behaviors, activities, and worker KSAs, including the criteria and the rationale for the inclusion and exclusion criteria.
	2.18	…a rationale for a systematic, quasi-random content sampling procedure based on the professional judgment of the testing professional and an analysis of work that details important work behaviors and activities, important components of the work context, and KSAs needed to perform the work; if sampling is not performed due to small content domain size, an explanation of how examination predictability is addressed is included.
Section 3.0		
3.1	The applicant or licensee provides that the validity plan will be implemented before the first administration of a test developed under the examination program. As part of the examination program, applicants and licensees should monitor and collect data following the validity plan over training cycles. If monitoring identifies a significant deviation from estimated risk associated with the type or degree of harm that may result or from the estimated probability that incorrect inference can be corrected before that harm occurs, applicants and licensees have processes to document the deviation, refine the validity plan,	

Summary of Acceptance Criteria		
	and reassess the examination.	
3.2	The applicant or licensee has a program that ensures that the tests will work as intended and that the tests will demonstrate one or more kinds of validity.	
Section 4.0		
4.1	The applicant or licensee clearly defines appropriate cutoff scores and their bases. If a cutoff score greater than or equal to 80 percent is used, then no additional basis needs to be provided, as NUREG-1021 has determined 80 percent to be an acceptable cutoff score for power reactors. If a cutoff score less than 80 percent is used, then the basis for that cutoff score is provided.	
Section 5.0		
5.1	The applicant or licensee provides information on the reliability of tests and supports that information with a reasonable description of the methodology used to determine its adequacy.	
Section 6.0		
6.1	The applicant or licensee provides a comprehensive test manual that covers each type of test. The manual includes information on the purpose of the test, how the test was developed, and how to administer and score the test.	
Test manuals include the following…		
	6.2	…a clear statement of the purposes and applications for which the test is intended.
	6.3	…clearly defined constructs the test intends to measure, the groups for which the test is intended, and the specific application of the test.
	6.4	…justification of the relevance of the test content for the constructs to be measured.
	6.5	…a summary of research findings for the test.
	6.6	…dates of any revisions to the manual and any additional statistical tests for the revised versions.
	6.7	…information on revisions to the test instruments associated with technology changes.
	6.8	…any limitations of the test.
	6.9	…a description of the expertise required to administer and interpret the test.
	6.10	…the time needed to administer the test.
	6.11	…the time allowed for testing.
	6.12	…a description of all procedures to be used for administration (and allowed variations).
	6.13	…a description of all materials for test takers to use.
	6.14	…a description of all scoring and reporting procedures, and a description of the necessary hardware and software.
	6.15	…clear and complete instructions to administer the test and interpret the results properly.
	6.16	…all necessary forms and any necessary aids to interpreting the test results.
	6.17	…case descriptions showing how test scores can be interpreted.
	6.18	…a clearly defined domain and discussion of the procedures used for sampling from that domain.
	6.19	…discussion of any potential threats to the valid use of the test scores.
For a specific test instance, documentation will include the following…		

Summary of Acceptance Criteria		
	6.20	…name of the test, authors of the test (by name and position), publisher of the test and the date published, and the existence of alternate forms.
	6.21	…dates of any revisions to the test and any additional statistical tests for the revised versions.
	6.22	…information on revisions to the test associated with technology changes.
	6.23	…a report on the evidence of the validity of the specified use of the test.
	6.24	…avoid referring to correlations between items and total test score as evidence of validity.
	6.25	…a description of criterion variables, adequacy of criterion variables, and aspects of criterion performance that are not adequately reflected in the criterion measures when criterion validity is reported.
	6.26	…statements expressing relationships that are presented in quantitative terms so that the reader can tell how much confidence to attach to them.
For computer-based testing, the test manual also includes the following…		
	6.27	…information on any necessary computer software installation.
	6.28	…information on operation of the software.
	6.29	…provision for technical support.
Section 8.0		
8.1	The applicant or licensee documents how it will meet examination security requirements. If the facility will have specifically licensed operators, it will also have procedures for preparing and submitting applications, including any waivers and excusals, and procedures for ensuring medical requirements are met. Facilities with specifically licensed operators will also have procedures for maintaining, updating, and renewing operator licenses.	
8.2	The applicant or licensee documents how examinees will be briefed on the examination process before examination administration, similar to the guidelines in section 1.2 of NUREG-1021 but tailored to the specific examination process developed for the facility.	
The applicant or licensee documents how it will ensure that the facility's examination program…		
	8.3	…tests a representative sampling of the KSAs to safely perform licensed duties.
	8.4	…is approved by the NRC before use.
	8.5	…will be maintained.
	8.6	…has provisions for how the prepared examinations for RO and SRO applicants will be made available to the NRC for review and approval before administration, if applicable.
	8.7	…has provisions for giving sufficient advance notification to the NRC to enable an NRC representative to be present during examination administration and, as applicable, to also administer the examination.
Section 9.0		
9.1	If a simulation facility is used in examinations, the applicant or licensee documents how the simulation facility provides a level of fidelity and cognitive requirements sufficient to assess the intended knowledge and skills of operators.	
9.2	If a simulation facility is used in examinations, the applicant or licensee documents policies and procedures to demonstrate compliance with simulator fidelity requirements.	

Summary of Acceptance Criteria		
If a simulation facility is used in examinations, the applicant or licensee provides content specifications that clearly describe the jobs and tasks covered in the simulation and the specific content, as follows…		
	9.3	…tasks are decomposed into both those knowledge, skills, cognitive processes, attitudes, or behaviors that are included in the examination and those that are excluded from the examination.
	9.4	…the type of personnel for which the examination is intended is specified.
	9.5	…the intended uses of the examination are specified.
	9.6	…evidence and supporting documentation are provided showing that the simulation strategy is appropriate based on the distinction made between knowledge and skills.
9.7	If a simulation facility is used in examinations, the applicant or licensee provides policies and procedures to link scenario events to the operator tasks and to document that information for scenario-based examinations.	
9.8	The applicant or licensee provides policies and procedures to develop event-based metrics for scenario-based examinations.	
Section 10.0		
10.1	The applicant or licensee documents how to administer the operating test portion of the examination while maintaining examination security and examiner codes of conduct to ensure that test integrity is maintained.	
Section 11.0		
11.1	The applicant or licensee documents the change management process for the examination program and specifies which changes to approved programs can be made without and which require prior NRC approval. This process is consistent with the requirements of 10 CFR 53.1565 and should also conform to the guidance of section 11.0.	
Section 13.0		
13.1	The requalification examination program has a provision to ensure that any requalification examination failures are properly remediated and retested, and that a retake examination is passed before allowing the licensed operator to return to licensed duties.	
13.2	The requalification examination program designates an appropriate periodicity for the administration of requalification examinations. For ROs and SROs, this periodicity may not exceed 24 months. For GLROs, this periodicity may exceed 24 months with adequate provisions for informing the periodicity.	
13.3	The requalification examination program has sufficient provisions to satisfy the additional requirements of 10 CFR 53.780(c)(2) or 10 CFR 53.815(b)(3)(v), as appropriate, that are discussed in section 13.0.	
Section 14.0		
14.1	The applicant or licensee provides a proficiency program for ROs and SROs that addresses the requirements of 10 CFR 53.780(g) and has appropriate justification for the adequacy of its provisions. Proficiency programs that are equivalent to, or more conservative than, the requirements of 10 CFR 55.53(e) and (f) may be considered acceptable without additional justification.	
14.2	The applicant’s or licensee’s programmatic provisions for change control	

Summary of Acceptance Criteria	
	specifically disallow making changes to the approved proficiency program that reduce its scope or requirements without obtaining prior NRC approval.
Section 15.0	
15.1	The applicant or licensee includes criteria for waiving the requirement for a test or examination for GLROs, including a basis for those criteria as needed.

With the satisfaction of these acceptance criteria, the NRC staff can conclude that the examination program complies with the requirements of Subpart F of 10 CFR Part 53, as applicable. Thus, there is reasonable assurance that the examination program meets the standards for examination in a manner that is commensurate with the design-specific safety role of the associated ROs and SROs or GLROs.

IMPLEMENTATION

The NRC staff will use the information in this ISG in reviewing operator licensing examination programs submitted as part of applications for OLs and COLs under 10 CFR Part 53. Additionally, this guidance may be used to inform staff consideration of requests for exemptions from the requirements in 10 CFR Parts 50, 52, and 55 for applicants and licensees under 10 CFR Part 50 and 10 CFR Part 52, or for proposals for using alternative methods to those described in NUREG-1021 or NUREG-1478 by applicants and licensees under 10 CFR Part 50 or 10 CFR Part 52. The NRC has incorporated feedback obtained during the public comment period for the Part 53 proposed rule and associated guidance into this ISG.

BACKFITTING AND ISSUE FINALITY DISCUSSION

DRO-ISG-2023-01 does not constitute backfitting as defined under 10 CFR 53.1590, “Backfitting,” and as described in Management Directive (MD) 8.4, “Management of Backfitting, Forward Fitting, Issue Finality, and Information Requests.” It does not constitute forward fitting as that term is defined and described in MD 8.4 or affect the issue finality of any approval issued under Part 53. The guidance does not apply to any current licensees or applicants or existing or requested approvals under Part 53. Therefore, its issuance cannot be a backfit or forward fit or affect issue finality. Further, applicants and licensees are not required to comply with the positions of this ISG.

CONGRESSIONAL REVIEW ACT

This ISG is a rule as defined in the Congressional Review Act (5 U.S.C. 801-808). However, the Office of Management and Budget has not found it to be a major rule as defined in the Congressional Review Act.

PAPERWORK REDUCTION ACT

This ISG provides voluntary guidance for implementing the mandatory information collections in 10 CFR Part 53 that are subject to the Paperwork Reduction Act of 1995 (44 U.S.C. 3501 et. seq.). These information collections were approved by the Office of Management and Budget (OMB), under control number 3150-0274. Send comments regarding this information collection to the FOIA, Library, and Information Collections Branch (T6-A10M), U.S. Nuclear Regulatory Commission, Washington, DC 20555 0001, or by e-mail to Infocollects.Resource@nrc.gov, and

to the OMB Office of Information and Regulatory Affairs (3150-0274), Attn: Desk Officer for the Nuclear Regulatory Commission, 725 17th Street, NW Washington, DC 20503.

PUBLIC PROTECTION NOTIFICATION

The NRC may not conduct or sponsor, and a person is not required to respond to, a collection of information unless the document requesting or requiring the collection displays a currently valid OMB control number.

FINAL RESOLUTION

The NRC staff will transition the information and guidance in this ISG into the regulatory guide or NUREG series, as appropriate. Following the transition of all pertinent information and guidance in this document into the regulatory guide or NUREG series, or other appropriate guidance, this ISG will be closed.

ACRONYMS

CFR	*Code of Federal Regulations*
COL	combined license
DID	defense in depth
ES	examination standard
GLRO	generally licensed reactor operator
ISG	interim staff guidance
JPM	job performance measure
KSA	knowledge, skills, and abilities
LLWR	large light-water reactor
MD	management directive
non-LLWR	non-large light-water reactor
NRC	U.S. Nuclear Regulatory Commission
OL	operating license
RO	reactor operator
RTR	research and test reactor
SAT	systems approach to training
SME	subject-matter expert
SRO	senior reactor operator
SSC	structure, system, and component

APPENDIX A

Currently Approved Examination Methods

The U.S. Nuclear Regulatory Commission (NRC) has determined the methods described in this appendix to be adequate testing methods for use on operator licensing examinations as documented in NUREG-1021, “Operator Licensing Examination Standards for Power Reactors,” and the use of these methods in an applicant or licensee examination program will not require additional review by the NRC.

Written Examinations

An applicant or licensee examination program that involves written examinations developed, administered, and graded in accordance with the guidance in Examination Standard (ES)-4, “Initial Written Examinations,” of NUREG-1021 and meeting the guidance of appendix A, “Overview of Generic Examination Concepts,” and appendix B, “Examples of Written Examination Questions,” to that NUREG does not require further review with respect to written examinations before NRC approval. Because the sample plans of NUREG-1021 do not necessarily accommodate the designs of non-large light-water power reactors, the applicant or licensee should provide the basis for the sampling plan to be used and explain how it meets the guidance of appendix A to NUREG-1021 to ensure examination validity. The applicant or licensee should also provide a basis for the use of the written examination test format to test the specific knowledge and ability items.

Any difference from the guidance in NUREG-1021 should be documented and a basis for that difference provided. Examples of differences include, but are not limited to, the number of questions on the written examination, the passing score, how to sample (random and systematic or other method), and differences in administration (such as using multiple locations to test simultaneously or using computers to provide proctoring services).

Operating Test Job Performance Measures

An applicant or licensee examination program that involves operating tests using job performance measures developed in accordance with the methodology in ES-3.1, “Overview of the Operating Test for Operator Licensing Initial Examinations,” and ES-3.2, “Developing Job Performance Measures,” of NUREG-1021 does not require further review with respect to the use of the job performance measure as a test instrument before NRC approval. The applicant or licensee should provide a basis for the use of the job performance measure test format to test the specific knowledge and ability items as well as establish a basis for the number of job performance measures used on the test and what constitutes passing performance. The format of the job performance measure that follows the methodology of NUREG-1021 does not require further review before NRC approval.

Job performance measures developed per the NUREG-1021 methodology should also be graded in accordance with the methodology in ES-3.6, “Grading and Documenting Operating Tests,” of NUREG-1021. Any differences should be documented and explained.

Operating Test Dynamic Simulator Scenarios

An applicant or licensee examination program that involves operating tests using dynamic simulator scenarios developed in accordance with the methodology in ES-3.1, ES-3.3, “General

Testing Guidelines for Dynamic Simulator Scenarios," and ES-3.4, "Developing Scenarios," of NUREG-1021 does not require further review with respect to the use of the dynamic simulator scenario as a test instrument before NRC approval. The applicant or licensee should provide a basis for the use of the dynamic simulator scenario format to test the specific knowledge and ability items.

Different considerations should be applied to commercial nuclear plants with designs that significantly differ from existing large light-water reactor plants in terms of the degree of automation, human-system interaction, or interaction between people. The extent to which earlier scenarios are modified to avoid predictability under the guidance of NUREG-1021 might not be necessary, depending on the scope of the knowledge, skills, and abilities (KSA) list. For example, a commercial nuclear plant design with inherent safety systems and automation might only have a couple of tasks important to safe plant operations that are screened into the operator licensing examination. In this instance, it might be appropriate to include those tasks in every task-based examination to ensure that the operator can complete those tasks satisfactorily. The operating test will need to be validated to determine whether it assesses the target KSAs using the selected operator tasks necessary for maintaining efficient and safe control of the plant under normal and abnormal situations. The applicant or licensee should describe the basis for any portion of the dynamic simulator scenario development process that differs from that described in NUREG-1021.

An applicant or licensee should grade dynamic simulator scenarios developed in accordance with the guidance of NUREG-1021 using the methodology in ES-3.6 of NUREG-1021. Any differences in grading methodology should be documented and explained.

APPENDIX B

Resolution of Public Comments

This ISG accompanies the 10 CFR Part 53 rulemaking package. Stakeholder comments, as well as the resolution of those comments, are included in the comment resolution documents associated with the 10 CFR Part 53 rule. See Docket ID NRC-2019-0062.

APPENDIX C

Bibliography

Part 1: Regulatory Bibliography

American National Standards Institute/American Nuclear Society (ANSI/ANS), "Selection, Qualification, and Training of Personnel for Nuclear Power Plants," ANSI/ANS 3.1-2014.

Nuclear Energy Innovation and Modernization Act (Public Law 115-439).

Nuclear Energy Institute (NEI), "Template for an Industry Training Program Description," NEI 06-13A, Revision 2, March 2009.

U.S. Code of Federal Regulations, "Domestic Licensing of Production and Utilization Facilities," Part 50, Chapter I, Title 10, "Energy."

U.S. Code of Federal Regulations, "Licenses, Certifications, and Approvals for Nuclear Power Plants," Part 52, Chapter I, Title 10, "Energy."

U.S. Code of Federal Regulations, "Operators' Licenses," Part 55, Chapter I, Title 10, "Energy."

U.S. *Code of Federal Regulations*, "Risk-Informed, Technology-Inclusive Regulatory Framework for Commercial Nuclear Plants," Part 53, Chapter I, Title 10, "Energy."

U.S. Nuclear Regulatory Commission, "Human Factors Engineering Program Review Model," NUREG-0711, Revision 3, November 2012.

U.S. Nuclear Regulatory Commission, "Operator Licensing Examination Standards for Power Reactors," NUREG-1021, Revision 12, September 2021.

U.S. Nuclear Regulatory Commission, "Knowledge and Abilities Catalog for Nuclear Power Plant Operators: Pressurized Water Reactors," NUREG-1122, Revision 3, September 2020.

U.S. Nuclear Regulatory Commission, "Knowledge and Abilities Catalog for Nuclear Power Plant Operators: Boiling Water Reactors," NUREG-1123, Revision 3, September 2020.

U.S. Nuclear Regulatory Commission, "Qualification and Training of Personnel for Nuclear Power Plants," Regulatory Guide 1.8, Revision 4, June 2019.

U.S. Nuclear Regulatory Commission, "Facility Training Programs," DRO-ISG-2023-04 (under development as of December 2022 and subject to change).

Part 2: Psychometric Bibliography

Adams, N.E. (2015). Bloom's taxonomy of cognitive learning objectives. *Journal of the Medical Library Association.* 103(3), 152–153. http://dx.doi.org/10.3163/1536-5050.103.3.010

Aggarwal, R., Mytton, O., Derbrew, M., Hananel, D., Heydenburg, M., et al. (2010). Training and simulation for patient safety. *Quality and Safety in Health Care*. 19(2), 33–43.

American Educational Research Association, American Psychological Association, & National Council on Measurement in Education (Eds.). (2014). "Standards for Educational and Psychological Testing." American Educational Research Association. https://www.testingstandards.net/uploads/7/6/6/4/76643089/standards_2014edition.pdf

Carey, J.M., & Rossler, K. (2021). The how when why of high fidelity simulation. *National Library of Medicine; StatPearls Publishing*. https://www.ncbi.nlm.nih.gov/books/NBK559313/

Cascio, W.F., Alexander, R. A., and Barrett, G.V. (1988). Setting cut-off scores: Legal, psychometric and professional issues and guidelines. *Personnel Psychology*. 41(1).

Chiang, I.A., Jhangiani, R.S., & Price, P.C. (2019). *Understanding Psychological Measurement. Research Methods in Psychology* (2nd ed.). https://opentextbc.ca/researchmethods/chapter/understanding-psychological-measurement/

Echtelt, R.V. (2020). What's the difference between knowledge and skill? *AG5.* https://www.ag5.com/whats-the-difference-between-knowledge-and-skills/

Evers, A., Sijtsma, K., Lucassen, W., & Meijer, R.R. (2010). The Dutch review process for evaluating the quality of psychological tests: History, procedure, and results. *International Journal of Testing*. 10(4), 295–317. DOI: 10.1080/15305058.2010.518325

Finan, E., Bismilla, Z., Whyte, H.E., LeBlanc, V., & McNamara, P.J. (2012). High-fidelity simulator technology may not be superior to traditional low-fidelity equipment for neonatal resuscitation training. *Journal of Perinatology*. 32, 287–292. https://www.nature.com/articles/jp201196

Goldstein, I.L., Zedeck, S., & Schneider, B. (1993). An exploration of the job-analysis-content validity process. In N. Schmitt & W. Borman (Eds.), *Personnel Selection in Organizations* (pp. 3–34). San Francisco, CA: Jossey-Bass.

Kardong-Edgren, S., Anderson, M., & Michaels, J. (2009). Does simulation fidelity improve student test scores? *Clinical Simulation in Nursing*. 3(1), 21–24. https://doi.org/10.1016/j.ecns.2009.05.035

Lievens, F., & Patterson, F. (2011). The validity and incremental validity of knowledge tests, low-fidelity simulations, and high-fidelity simulations for predicting job performance in advanced-level high-stakes selection. *Journal of Applied Psychology*. 96(5), 927–940. DOI: 10.1037/a0023496

Martin, S.L., & Raju, N.S. (1992). Determining cut-off scores that optimize utility: A recognition of recruiting costs. *Journal of Applied Psychology*. 77(1), 15–23.

Massoth, C., Roder, H., Ohlenurg, H., Hessler, M., Zarbock, A., Popping, D.M., & Wenk, M. (2019). High-fidelity is not superior to low-fidelity simulation but leads to overconfidence in medical students. *BMC Medical Education*. 19.

Motowidlo, S.J., Dunnette, M.D., & Carter, G.W. (1990). An alternative selection procedure: The low-fidelity simulation. *Journal of Applied Psychology*. 75(6), 640–647.

Murphy, K.R., & Davidshofer, C.O. (2005). *Psychological Testing: Principles and Applications* (6th ed.). Pearson/Prentice Hall.

Norman, G., Dore, K., & Grierson, L. (2012). The minimal relationship between simulation fidelity and transfer of learning. *Medical Education*. 46(7), 636–647. https://doi.org/10.1111/j.1365-2923.2012.04243.x

National Center for O*NET Development. (n.d.). *O*NET OnLine*. U.S. Department of Labor, Employment & Training Administration. Retrieved May 22, 2025 from https://www.onetonline.org/

Society for Industrial and Organizational Psychology. (2018). "Principles for the Validation and Use of Personnel Selection Procedures." American Psychological Association. https://www.apa.org/ed/accreditation/about/policies/personnel-selection-procedures.pdf

Salas, E., & Burke, C.S. (2002). Simulation for training is effective when… *Quality Safety and Healthcare*. 11, 119–120.

Salas, E., Paige, J.T., & Rosen, M.A. (2013). Creating new realities in healthcare: The status of simulation-based training as a patient safety improvement strategy. *BMJ Quality and Safety*. 22, 449–452. doi:10.1136/bmjqs-2013-002112

Schmidt, F.L., & Hunter, J.E. (1998). The validity and utility of selection methods in personnel psychology: Practical and theoretical implications of 85 years of research findings. *Psychological Bulletin*. 124, 262–274.

Thorndike, R.M. (2005). *Measurement and Evaluation in Psychology and Education* (7th ed.). Pearson.