ARELLO® EXAMINATION ACCREDITATION PROGRAM

Purpose and Process of Licensing Examination Development

The following information is provided only as an educational resource for the purpose of assisting members with an understanding of examination development methodology and process. The information is a compilation derived from research of published works such as the *Standards for Educational and Psychological Testing* and internet based literature; and should not be disseminated as an ARELLO® work product.

1) Purpose of Examinations (Public Protection)

a) A primary purpose of regulatory licensing is to protect the public by ensuring that entry-level real estate licensees (both salesperson and broker level), possess the knowledge and ability to competently and safely perform the activities and tasks essential to brokerage services.

b) Exams used in licensing are intended to provide the public and government agencies with a dependable mechanism for identifying those who have met particular standards and are designed to determine whether the essential knowledge and skills of a profession have been mastered by the candidate. The focus of performance standards is on levels of knowledge and performance necessary for safe and appropriate practice.

c) For purposes of licensing, the activities and tasks essential to brokerage services distinguish between those tasks that are related to successful practice in the business of real estate and those task competencies required to protect the public's interest. Activities and tasks that may be important to success in the profession, however are not directly related to the purpose of licensure (public protection), should not be included in pre-licensing exam development except as authorized or approved by the licensing agency.

2) Critical Issues in Exam Development (Validity, Reliability, Fairness, Legal Defensibility)

a) Validity

i) The validity of an exam relies primarily on "evidence" that the subject area content reflects the job-related knowledge required to perform the activities of the profession in conformance with standards of practice, laws and prescriptive rules. Such "evidence" of validity supports that the exam score is reflective of the exam's ability to discriminate between those who are qualified and those who are not.

ii) To be valid and defensible, exams must test job related competencies. The subject area content and areas of knowledge being tested must be those that are necessary to perform tasks which are essential for an entry-level practitioner to be minimally competent. Exam content validity is concerned with an exam's ability to include or represent all of the subject area content of what is essential to perform the job.

iii) An accepted method of determining that the subject area content of an exam is valid is the use of job/task surveys and subject matter expert analysis and determination of what is essential to perform the job.

b) Reliability

i) Reliability is one of the most important elements of exam quality. It has to do with the consistency, or reproducibility, of an examinee's performance on the exam.

(1) For example, if you were to administer an exam with high reliability to an examinee on two separate occasions, you would expect that the results of the examinee's performance would be consistent. An exam with poor reliability, on the other hand, might result in very different scores for the examinee across the two exam administrations.
ii) If an exam yields inconsistent scores, it is likely that the exam may be legally challenged as an ineffective measure of preparedness to practice.

c) Fairness

i) The fairness of an exam refers to its absence of bias. The exam should be appropriate for all qualified examinees irrespective of race, religion, gender, or age.

ii) The exam should not disadvantage any examinee, or group of examinees, on any basis other than the examinee's lack of the knowledge and skills. Item writers should be instructed in methods of assuring fairness as they undertake the task of writing items.

iii) In addition, the items should also be reviewed for potential fairness problems during the item review phase. Any items that are identified as displaying potential bias or lack of fairness should then be revised or dropped from further consideration.

d) Legal Defensibility

i) For an exam program to have legal defensibility there must be “evidence” as to the exam's suitability of purpose that would stand up to a legal challenge.

(1) Such evidence would consist of proof that sound, generally accepted and recommended guidelines were followed throughout the design, development, and maintenance of the exam program.

(2) Professional guidelines for examinations are offered by such entities as: the American Psychological Association (APA), American Educational Research Association (AERA), the National Council on Measurement in Education (NCME), U.S. Equal Employment Opportunity Commission (EEOC) and the Association of Real Estate License Law Officials (ARELLO®).

ii) The single most important element of validity evidence related to legal defensibility is that the exam content subject areas and questions (items) are based on an appropriate job task survey and analysis. In other words the exam content subject areas and exam questions must be proven to have a high degree of "job relatedness".

3) Exam Content Development

a) Job Task Survey (Questionnaire)

i) The content of the survey questionnaire is developed through utilization of “Subject Matter Experts” committees to identify the activities involved in real estate brokerage and the range of tasks performed in relationship to these activities.

(1) Subject Matter Experts (SMEs) are people who are generally acknowledged as having expert knowledge in a particular content area or areas that are being measured on a particular test. All testing programs make use of SMEs, whether they call them that or by another name. SME Committees for real estate licensing exams generally include those persons with experience in:

   (a) teaching and/or developing pre-licensing education courses that are relevant to the exam content areas.

   (b) authorship of principles and practices, law or reference books that are relevant to the exam content area.

   (c) the practice of real estate brokerage as a licensee

   (d) the regulation of practice relevant to the profession of the exam being administered.

   (e) other relevant occupation such as experience as an attorney-educator that is relevant to the exam content area.

ii) The survey questionnaire is typically completed by large sample of real estate practitioners and is used to validate brokerage activities and the related tasks. The survey seeks to determine:

   (1) what task competency is actually needed for an entry-level practitioner and ranked in order of importance for entry level
(2) the frequency of which the tasks are performed (e.g. daily, monthly, occasionally, seldom)

(3) the importance of the task as it relates to effective performance in protection of the public’s interest

(4) and the relative importance in the actual practice of a profession.

b) Job Task Analysis

An analysis of the survey results is performed by Subject Matter Experts and the examination provider development staff. The objectives of the Job Task Analysis are to:

i) Distinguishes between those task competency is needed at entry level as opposed to what can be learned or enhanced on the job.

ii) Distinguishes between those task competencies that are critical at entry level and what can be learned or enhanced on the job.

iii) Distinguishes between those tasks that are related to successful practice in the business of real estate and those task competencies required to protect the public’s interest. Not always easily distinguishable.

iv) Identify what knowledge, skill and ability (KSA), is required to perform the tasks related to the activities involved in real estate brokerage. This is referred to as “the knowledge statement”. Then SMEs are asked to participate in a linkage activity to establish an explicit relation between each KSA and one or more tasks from the job analysis. After further review and comment, the document is finalized and weights are assigned to indicate the number of exam questions to be allocated to each KSA category.

For instance, “knowledge of determining market value” is the knowledge statement and it is linked to what the job analysis identifies as required to perform the task identified by the job survey category area as determining a listing price in the activity area identified by the job survey of completing a listing agreement contract form.

Identify the various subject areas of activity in professional practice (e.g. sales contracts, closings, finance, listing, etc.), these are called domains, and categorize these domains by subject content area and sub-categories.

Categorize the knowledge statements, tasks and activities by content subject areas.

(a) For instance, knowledge of determining market value, the task of determining a listing price and the activity of completing a listing agreement are categorize in the major content area of Listing Agreements

Assign weights to the categories to indicate the emphasis to be allocated to each category. The rationale for the weight for a content category should be a function of (a) the number of tasks linked to the topics in that category and (b) the magnitude of the importance.

c) Creating the Exam Specifications

After the overall content of the exam has been established through adherence to the specifications relative to the job/task survey and analysis (as identified above), the next step in developing an exam is to create detailed specifications for statistical values pertaining to: difficulty levels, weighting of importance of the content areas, determining the equivalence of alternate forms of the same test, exam reliability (consistency of scores candidates would receive on alternate forms of the same exam), cognitive levels, passing scores, discrimination, editorial quality, item writing and performance evaluation; Exam specifications are like a blueprint that establishes processes, procedures, scientific methodologies and the criterion (or standards) for the development of the exam.

4) Developing the Exam Questions (Item Writing)

a) Subject Matter Experts are convened to develop the exam questions. They are provided guidance by the exam development staff to write questions in conformance with basic item writing principles (see ARELLO® Basic Item Writing Principles) such as:
i) The exam question must have only one clearly correct or clearly best answer (called the “key”). The stem and answer must contain or result in an accurate statement(s).

ii) The subject matter being tested by an item must be relevant and appropriate.

iii) Any scenarios used should be realistic and appropriate.

iv) Each item must be based on an important point, idea, or skill that reflects an appropriate level of competency in knowledge or skill related to tasks shown to be needed for entry level.

v) Subject Matter Experts are given additional direction that ensures each exam question is written so that it can be demonstrated to be appropriately linked to job-related activities (as identified through the job analysis).

5) Exam Question (Item) Performance Analysis

The item analysis is an important phase in the development of an exam program. In this phase statistical methods are used to identify any exam items that are not working well. If an item is too easy, too difficult, failing to show a difference between skilled and unskilled examinees, or even scored incorrectly, an item analysis will reveal it. The two most common statistics reported in an item analysis are the item difficulty, which is a measure of the proportion of examinees who responded to an item correctly, and the item discrimination, which is a measure of how well the item discriminates between examinees who are knowledgeable in the content area and those who are not. An additional analysis that is often reported is the distractor analysis. The distractor analysis provides a measure of how well each of the incorrect options contributes to the quality of a multiple choice item. Once the item analysis information is available, an item review is often conducted.

6) Scoring and Evaluation of the Exam

a) The validity of exam results (pass/fail) depends on the specification for the “standard for passing” being able to make a distinction between those with adequate and those with inadequate knowledge.

b) A Committee of Subject Matter Experts (SME’s) is utilized by examination developers to evaluate each question on an exam and determine the percentage of likelihood that a minimally competent licensed practicing professional will be able to answer the exam item/question correctly. These proportions or percentages are referred to as “p-values”.

i) Each item on an exam is assessed in terms of how likely a minimally competent candidate (those who would barely meet mastery standards) is to answer that item correctly.

1) The SME’s are each asked to (independently) think of a group of minimally competent candidates who would border on the mastery/non-mastery cut-off. The most typical instruction is for the SME’s to think of a pool of 100 candidates who would “just barely” meet the performance criteria. The SME’s, working independently, then estimate what proportion or percentage of that sample of 100 minimally acceptable candidates would answer each item in the exam correctly.

ii) The proportions (or percentages) of each SME are then summed and usually denoted as the Minimum Passing Level (MPL). The MPL represents an individual SME’s score for the exam (called a “cut score”). The mean of these cut scores is the final cut score for the exam. When SME’s are asked to assess the probability of candidates correctly answering an item, they are in essence determining the difficulty of the item.