

**FY2004 Task 5, Report 2:**

**METHODOLOGY FOR VALIDATION OF TASKS AND COMPLETION  
STANDARDS IN FAA PRACTICAL TEST STANDARDS**

Submitted by:

Peter M. Schumacher  
University of North Dakota

Beth Blickensderfer, Ph.D.  
Michele Summers, M.S.  
Embry-Riddle Aeronautical University

**May 13, 2005**

**This effort is the second deliverable under the FITS Program FY2004 Task 5: *Development of a methodology to justify the inclusion or removal of maneuvers from flight training curriculums.***

# METHODOLOGY FOR VALIDATION OF TASKS AND COMPLETION STANDARDS IN FAA PRACTICAL TEST STANDARDS

## INTRODUCTION

The FITS Task 5 team recently completed a review of common practices used to validate the inclusion of specific tasks in training curriculum and job related evaluations. (See FY2004 TASK 5, REPORT 1: INVESTIGATION OF CURRENT PRACTICES RELATING TO THE PTS.) The purpose of the current report is to propose a methodology for justification of the tasks and completion standards in the FAA Practical Test Standards (PTS). This report first presents an overview of the methodology. This is followed by a detailed description of a proposed demonstration (i.e., test) of this methodology.

### ***PROPOSED METHODOLOGY: Overview***

Based on Report 1 from this tasking, we recommend using a typical *job analysis* approach (to assess validity of the tasks listed in the PTS) with an *evaluation of test validity* (to validate the test completion standards included in the PTS). Job analysis is a common tool within applied psychology and education used as a first step to evaluate human performance and instructional design. As outlined in the previously mentioned report, a job analysis identifies the key tasks required to perform a job, and then lists the skills and knowledge that are necessary to accomplish each task. Once a job analysis is completed, the validity of exams used for that job can be established. The two types of validity relevant to the current tasking are content validity and criterion validity. First, *Content Validity* is the degree to which a particular task reflects the knowledge, skills, or abilities needed to accomplish the real job or assignment. Second, *Criterion Validity* is the degree to which the standards for satisfactory test completion relate to the performance measures that are expected during actual job completion.

The PTS provide a readily available framework for identification of the required tasks and necessary skills to pilot an aircraft. The PTS identify broad areas of operation (e.g., Airport Operations) and describe the individual tasks that make up each area (e.g., Traffic Patterns). The PTS also list the skills that must be successfully demonstrated to complete each task. These skills are referred to as completion standards.

The proposed methodology implements a rating system to evaluate the validity of each task and associated completion standards. To accomplish this, a representative sample of qualified Subject Matter Experts (SMEs) would assess each task, subtask, and skills on a variety of dimensions. Depending on which PTS is evaluated, the sample might include certified flight instructors, AOPA-sanctioned Master CFI's, and FAA Examiners. Based on the responses of the SMEs, changes could be made to the PTS.

If the FAA adopts this methodology, during each five-year cycle for PTS updates this evaluation process would occur again. That is, a representative sample of SMEs would be asked to complete a PTS Update Survey once every 5 years (or more if necessary).

A detailed description of the SME evaluation process, including items to be evaluated and rating dimensions, is described in the next section of this report.

## ***Evaluation Process***

The first step in the validation process is *to identify the job that is being evaluated* by the PTS. For this discussion, we have narrowed our scope to the Private Pilot and Commercial Pilot Practical Test. The first task is to define the job of a Private Pilot; which in the most basic sense is *to safely and efficiently fly an aircraft from one airport to another*. All further assessment of specific tasks and completion standards must be related to this job description.

The next step is to *list the tasks required to complete this job*. The PTS has already identified these tasks. To prepare the survey, we list the tasks and then describe the basic skill incumbent in each (See attached Private/Commercial Pilot PTS Validation Survey). As you can see in our example, the defined skills are quite broad and simply provide a general overview of the task. The Completion Standards further define the necessary skills and therefore are not listed on the survey. (Please note: the attached spreadsheet survey is only a draft to demonstrate tasks and completion standards.)

Each task on the PTS is then evaluated for *content validity*—is it an activity that a private pilot would accomplish during a real life flight scenario? Questions are asked regarding how frequently the task is required, with a range of responses from 1 (Seldom) to 5 (Always); and how important is the task 1 (Unnecessary) to 5 (Critical). Seldom accomplished and non-critical tasks are generally considered not content valid.

Next, the *completion standards* are listed. These are evaluated in terms of degree to which the criteria are realistic and reflective of the performance that would be expected in real life? If the measures (in this case, the standards) are too critical OR if the measures do not correspond to how performance is assessed in real life OR if the measures require a knowledge or skill that is artificially inserted for the test, then the criterion may not be valid. Once again, a range of responses is provided: 1 (Unrealistic) to 5 (Realistic). (Note: In an exact test of criterion validity, data would be collected on the test criteria and correlated with actual performance data. The method we suggest here—SME ratings—requires fewer resources and would give insight into the PTS, although it is not an exact test of criterion validity.)

The next step in the analysis is to *identify evaluation tasks that are redundant*. If the required knowledge and skills incumbent in a particular task are evaluated in other test items, then consideration should be given to eliminating this redundancy—especially if training and evaluation costs are prohibitive. Responses range from 1 (Excessively Redundant) to 5 (Not Redundant).

After a ranking has been applied to all tasks and criterion, the final step is to calculate the average responses from all of the SME ratings. High mean responses identify sub-tasks and skills that are necessary for effective accomplishment of the job/mission. Conversely, low mean responses identify tasks that have low, if any, relevancy to the job. The final step then is to

modify the PTS. In the interest of available time and resources for pilot testing, those sub-tasks, skills, and criteria with low scores should be considered for elimination from the test battery.

### ***PROPOSED TEST***

To demonstrate the job analysis approach to PTS validation, we propose the following test:

- 1) A sample Private Pilot Validation Survey will be developed. The survey will include a cross section of tasks from the Private Pilot Practical Test Standards in accordance with the attached example.
- 2) The survey will be distributed to 100 Subject Matter Experts. Due to research deadlines, this group will primarily include UND and ERAU flight instructors. As available, other members of the FITS team may be included.
- 3) Participants will rate individual tasks and completion standards in accordance with the methodology previously described.
- 4) Survey data will be analyzed and collated by the Task 5 FITS research team.
- 5) Results will be evaluated by the FITS research team and disseminated to appropriate personnel within the FAA and industry for review.