WISE CHOICES

In *CAT 2*, 2001, aviation psychologist, Dr. Diane Damos described the differences between a screening system and a selection system. In this article, she discusses how a selection system is developed for use in civil aviation. The examples given are for the selection of experienced pilots. Ab initio systems designed to produce pilots for a specific carrier follow a similar process.

All professionally developed selection systems are designed to meet three goals. First, they must identify the best candidates for the company. Second, they must perform this identification cost-effectively. Third, all of the components and processes of the system must be legally defensible.

The importance of this third goal varies widely from country to country. In the United States and South Africa for instance, meeting the third goal is of enormous importance. For this reason, all of the processes described below meet the highest levels of 'best practice'.

Frequent use is made of the term 'selection instrument' rather than 'test'. 'Test' refers to paper and pencil examinations or computerized versions of these examinations. 'Selection instrument' is a more general term that refers to any assessment method, including simulator evaluations and interviews.

Most of the principles for test development and battery construction were developed in the 1930s and 1940s. Although the guidelines concerning the use of selection tests have become more exacting, the process of developing a cost-effective battery remains unchanged and can be described in seven major steps.

1. Conduct **a job analysis**

This typically begins with a task analysis to identify the knowledge, skills, abilities, and traits required by the job. This is time consuming but many air carriers already have completed task analyses for their AQP program. This type of a task analysis, a training analysis, differs from a selection analysis in its level of detail. Figure 1 shows part of a generic training task analysis for a two-person jet aircraft. Training analyses are too detailed to use

Left
A task analysis identifies the knowledge, skills, abilities and traits required by the job. *Image Credit: Ted Fahn/SAS*

for selection system development and must be modified to reflect large categories of activities. Usually, this modification can be accomplished by having a group of five to nine subject matter experts (check airmen, senior captains, simulator instructors, etc.) cluster the elements into higher-order activities.

For example, they might include the elements from tasks 4.2.2.2 to 4.2.2.7.3 and elements from other sections of the training task analysis into a cluster called 'communication'. **See Figure 1.**

A second group of experts identifies the knowledge, skills, and abilities required by the various clusters of task elements. This process is usually facilitated by an outside consultant. At this point, the second group is allowed to describe the required knowledge, skills, and abilities in its own words.

Tasks analyses provide useful information on the abilities required to be a successful pilot at a given carrier. However, analyses should not be expected to produce substantive information on the personality traits that are most important to an air carrier. Descriptions of useful personality information may be obtained using one or several different methods. For a small operator, the minimum acceptable process involves individual interviews of first-level and second-level managers in the flight operations and the training departments. These interviews identify past problems the company has experienced with its pilots, i.e., insufficient job skills, lack of dependability and poor stress management. They also identify relevant personality strengths associated with successful pilots.

For medium-sized carriers, facilitators also may hold small-group discussions with current pilots. In these discussions, the problem areas identified by management are broken down into their component parts and operationally defined.

For large carriers, the management interviews are the first step. The second step may involve small-group discussions, such as those described above, and/or the use of a survey. If management decides to use a survey, the results of the interviews are used to develop the survey questions, which are administered to a large group of line-qualified captains. Small focus groups then translate the survey results into specifications of desirable and undesirable personality traits.

2. Literature **review**

Literature reviews are performed for two reasons. Sometimes identifying personality traits is difficult. In such cases, a literature review may reveal traits that have been identified at other carriers. More commonly, however, the literature review is conducted to identify potential selection instruments. Many air carriers lack sufficient resources to conduct the review and interpret the results. Consequently, a consulting firm may be hired to conduct the review and evaluate the results.

Figure 1

4.2.2	Perform cockpit crew communications*
4.2.2.1	Complete cruise checklist if appropriate
4.2.2.2	Callout fault/warning messages as appropriate
4.2.2.3	Make standard callouts as appropriate
4.2.2.4	Perform augmented crew turnover briefing as applicable
4.2.2.4.1	Communicate A/C current position and assigned altitude
4.2.2.4.2	Communicate current clearance versus flight plan and any route/altitude requests
4.2.2.4.3	Communicate waypoints loaded versus flight plan
4.2.2.4.4	Communicate navigation system(s) accuracy
4.2.2.4.5	Communicate frequencies and controlling agency
4.2.2.4.6	Communicate fuel status/optimal altitude
4.2.2.4.7	Communicate enroute, destination, and alternate weather
4.2.2.4.8	Communicate any incidents or irregularities
4.2.2.4.9	Communicate last company position report
4.2.2.5	Perform communications with cabin crew
4.2.2.5.1	Discuss turbulence situation and passenger seating requirements as appropriate
4.2.2.5.2	Discuss special passenger requirements and time available for cabin service
4.2.2.5.3	Discuss aircraft servicing requirements and cabin discrepancies
4.2.2.6	Perform communications with company
4.2.2.6.1	Communicate rerouting, delay, schedule change, or gate information with dispatch
4.2.2.6.2	Acquire destination airport status/weather from dispatch
4.2.2.6.3	Discuss aircraft systems discrepancy, systems troubleshooting, continuation criteria, and/or deferral actions with maintenance
4.2.2.7	Perform communications with ATC
4.2.2.7.1	Respond to traffic callouts, vectors, or clearance changes
4.2.2.7.2	Read back ATC clearances
4.2.2.7.3	Respond to weather and/or delay information

** from Longridge, October, 1995, personal communication*

3. Select performance measures **to use as criteria**

All selection systems are prediction systems and as such, they are designed to predict something. This 'something' is referred to as the 'criterion'. Today air carriers have many choices of criterion measures, from assessments of a candidate's performance during training, probation, and line flying. Air carriers may, and probably should, use more than one criterion in their selection system.

4. Identify types of instruments **that may be used**

This step is usually performed by a consultant. The consultant maps the skills, abilities, and personality traits that were identified in step 1 to accepted psychological constructs. For example, if one of the necessary personality traits identified by management is, 'arrives for work on time', the consultant may match this to the psychological construct, 'conscientiousness'. The consultant, with the subject matter experts, also will categorise the knowledge requirements into broad categories, such as 'high altitude meteorology', 'regulations' or 'aircraft performance characteristics'.

5. Identify the most promising **selection instruments**

The consultant uses the results of the literature review and other sources to identify the most promising commercially available selection instruments. Identifying the best selection instruments is usually straightforward. The major considerations typically are cost per candidate, time to administer the instrument, and reliability.

Frequently, the task analysis, interviews, small group discussions, and surveys identify such a large number of knowledge areas, skills, abilities, and traits that assessing all of them is impractical. In such cases, management must determine how much money and time can be devoted to assessing each applicant. The system designer finds a selection instrument for the most important item on the list and adds instruments to the proposed selection system until the cost per applicant or time per applicant limit is reached.

Occasionally no instrument is commercially available. In such cases, management has three choices. The first option is to wait. The second option is to develop an instrument, although this process will take resources and time. A third option is to include assessment of the missing elements as part of the interview process.

6. Administer the selection instruments **to obtain** predictive validities

Each instrument must be tested to determine how well it predicts performance on the criterion before management can use it to make a hiring decision. This testing can be done in one of two ways. If a cri-

terion is selected that is already in use, such as scores on a yearly check ride, then the 'concurrent validity' approach may be used. To use this approach, two groups of current pilots are selected: one group who scored well on the criterion and one who scored relatively poorly. The selection instruments then are given to each group.

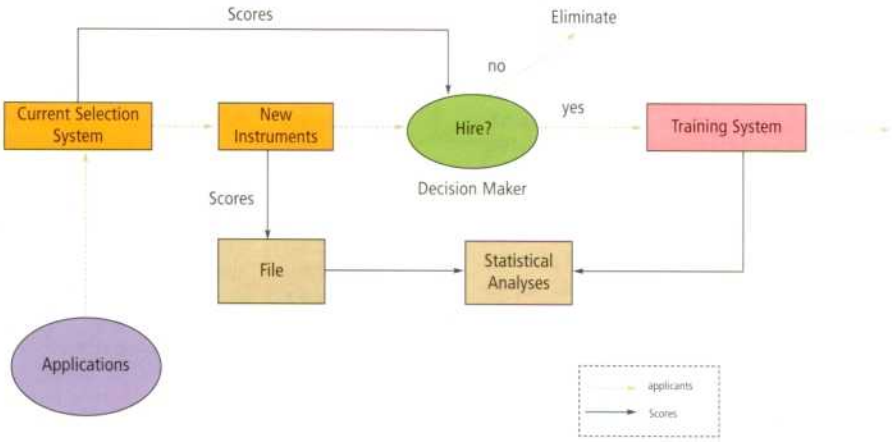
In countries like the United States or South Africa, both the total scores from each instrument and individual test items will be checked for bias (the correlation between scores on the selection instrument and scores on the criterion differ for different groups of people) and adverse impact (a lower-than-expected proportion of a protected group passes the instrument). Any selection instrument demonstrating bias or adverse impact will be dropped from the proposed selection system. Then the concurrent validity (the correlation between scores on an instrument and scores on the criterion) will be calculated for each instrument.

If the concurrent validity of a selection instrument is low, it may be dropped from the proposed selection system. If many instruments must be dropped, the testing process may have to be repeated with new instruments and two new groups of pilots.

Finally, the scores on the remaining instruments with the scores from the existing selection instruments are mathematically weighted to provide the best possible prediction.

The second approach to obtaining validity data uses 'predictive validity'. This is the traditional method used and assumes that some type of selection system is already in place. An example of testing using the predictive validity approach is shown below in **Figure 2**.

Figure 2



Applicants complete the current selection system as well as the new instruments. The results of the new instruments are not given to the decision makers; they make their hiring decisions using the same information they had previously. As soon as the new hires produce the criterion data, the predictive validity (the correlation between the scores on an instrument and the criterion scores) is calculated. Any instruments that demonstrate low predictive validity are dropped from the battery. Again, the scores on the remaining instruments with the scores from the existing selection instruments are mathematically weighted to provide the best possible prediction. The next group of applicants is assessed using both the new instruments and the old instruments from the old selection system. Management then may use all of the information in the hiring decision.

Assume that a carrier is using the predictive validity approach to test new selection instruments and that a personality test assessing a trait of major importance to the air carrier is dropped from the battery because of its low predictive validity. How can the operator assess the trait of interest? First, the operator must find another selection instrument or develop one. The process described above then is repeated: a group of applicants receives the new personality test (in addition to all of the other selection instruments), but its data are not used to make any hiring decisions. Mathematical analyses are then performed on the scores to determine if they increased the predictive validity of the selection battery. If so, the test is added to the battery for the next group of applicants. If not, another personality test must be located or the test revised.

7. Monitor **the system**

Once the battery contains all of the desired components, it is constantly monitored to ensure that its predictive validity does not decrease over time. The predictive validity of the selection system may decrease because of changes in the characteristics of the applicants. If this occurs, someone with the necessary technical background must be contacted to recalculate the weighting of the selection instruments or find replacements for instruments that have become obsolete or compromised.

Widely used

The process described is the most widely used one for developing selection systems. Some details may vary depending on the size of the air carrier and the laws of the country where the hiring decisions are made. At first glance this process may seem formidable. If training task analyses are available, management should allow from three to 12 months to develop a selection system using the process described here. The amount of time required depends in part on the size of the carrier; larger carriers usually require more development time. The process should be thoroughly and carefully documented to improve the carrier's corporate memory and to decrease the company's legal exposure.