Aiming for selection perfection

The goal of all modern pilot selection systems is to identify the best applicants in a cost-effective and legally defensible manner. Despite the simplicity of this goal, many managers involved in pilot hiring misunderstand how a selection system works, writes Dr. Diane Damos.

Two common misunderstandings relate to selection systems. The first pertains to screening systems, which are often confused with selection systems. A screening system should identify applicants who do not meet a company’s minimum qualifications and eliminate them from further consideration. For example, a company may specify that an applicant needs 1,000 total hours of flight time to be considered. Pilots may apply, however, who only have 700 hours. The screening system is designed to identify these individuals and eliminate them from the applicant pool.

Air carriers typically screen applicants on education, citizenship, criminal record, and flight experience. Selection does not begin until after screening is completed. The second major misunderstanding is that selection systems like the one described below, are only for major carriers. In fact, these system will work for any carrier that hires more than around 30 pilots/year. Even if your company hires fewer than 30 pilots/year, most of the principles and features still apply.

Modern selection systems consist of five major elements: a criterion, selection instruments, a decision aid, the selection model and administrative procedures, and a feedback loop.

Note the use of the term ‘selection instrument’ rather than ‘test’. ‘Test’ refers to a paper-and-pencil examination or a computerized version of an examination. ‘Selection instruments’ is a general term that refers to any assessment method, including interviews, simulator evaluations, paper and pencil examinations, or computerized examinations.

Criterion
Managers often believe that selection systems consist entirely of selection instruments, or that the selection instruments are the most important element of the system. This is untrue; the most important element of the selection system is the criterion - the behavior(s) that the system is designed to predict. Most selection systems, particularly those of smaller carriers, consist of a simulator evaluation and an interview. Based on the results of the interview and the simulator evaluation, the chief pilot (or the flight operations manager) guesses (predicts) if the applicant will be a ‘good pilot’ for the company. Thus, the chief pilot guesses if the applicant will have a set of job behaviors that for him define a ‘good pilot’. This guessing process has three problems. First, the chief pilot may not be thinking about the same set of behaviors from day to day. Second, the amount of each behavior that an applicant must demonstrate to be hired may vary from day to day. Third, the chief pilot’s set of required job behaviors might not be the same as those of the flight operations manager. The criterion does not have to assess just one type of job performance; the criterion may consist of several measures of different aspects of job performance. For example, the management of ‘Acme Airlines’ may decide that they want to predict the score on the check ride administered at the end of training and that this score is all that is important to them. In contrast, the management of ‘Mountain High Airlines’ feel that two job behaviors are important. One of these is punctuality; they want to hire pilots who will stay on schedule. Mountain High Airlines keeps extensive records of on-time performance for each of its pilots. For Mountain High, therefore, punctuality is easy to measure. The second behavior, however, concerns learning ability: they want pilots who will have little difficulty with training and can be promoted in a timely fashion. Because Mountain High management cannot measure learning ability directly, they select two other measures that they believe are related to learning ability. One concerns the resources the new-hire pilot needs to complete initial flight training. In this case ‘resources’ may be thought of as the dollar cost of the simulator time, the aircraft time, and the ground school time the applicant needs to pass the check ride. The second behavior is the average score on the first four check rides the applicant takes after he/she begins revenue flying for the company. Thus, Mountain High uses a criterion composed of three measures: on-time performance, the resources needed to pass the end-of-training check ride, and the average score on the first four line check rides. Is it possible to select any behavior to use as a criterion? Yes, but only if the behavior is quantifiable. Clearly, the score on a check ride is a quantifiable measure. Many air carriers want to use a criterion like ‘gets along well with other pilots’. Is something like ‘gets along well with the other pilots’ quantifiable? Yes, if the company can define what ‘gets along well’ really means. Any measures that are used as criterion must have certain properties. For example, they must be sensitive to differences in levels of performance. Consider scoring a check ride as ‘pass’ or ‘fail’. Pass/fail provides only a gross
discrimination (two categories) between levels of performance. Thus, pass/fail is not generally a good criterion measure because it is not sensitive to performance differences. Scores that range from 0 to 100 are much better because they give many more categories of behavior (101 in fact).

Instruments

The second element of the selection system is instruments. These typically consist of paper and pencil tests, computerized tests, interviews, and an evaluation of the applicant’s flying skills, which may be conducted in either a simulator or an aircraft.

How many selection instruments should an air carrier have? That depends on how well the instruments predict the criterion and their cost. It also depends on how closely related two instruments are to each other.

Decision aid

The third element is the decision aid, which is designed to predict the scores on the criterion from the scores on the selection instruments. Decision aids typically produce a number that is the predicted score on the criterion for a given applicant. In some cases they produce a probability that the applicant’s score on the criterion will exceed a given value. The mathematics of the process need not concern us and in any case, the result of the decision aid is given to the person who makes the hiring decision to guide him/her in making a hire/reject decision.

Selection model

The fourth element is the administrative procedures/selection model. The choice of selection model determines when applicants will be eliminated from the pool. The simplest selection model is called a ‘single-stage’ model. A company using this model would administer all of the selection instruments to all of the candidates and the decision aid would use all of the information to generate a prediction.

The second model is called a ‘multi-stage’ model and consists of at least two steps or stages. A company using this model would administer some of the selection instruments in stage 1 to all of its applicants. The results of the instruments would be submitted to a decision aid and, based on its prediction, a substantial proportion of the applicants would be eliminated from further consideration. The remaining candidates would then complete another set of instruments in stage 2. Again, the results would be submitted to a decision aid, which would make a prediction for each candidate. This process is repeated for each stage of the
model until the final stage, where a hire/reject decision is made based on the results of the decision aid.
The third model has several different names but will be referred to here as a progressive model. The philosophy behind this model is considerably different from those behind the single and multi-stage models.

Progressive

A progressive model is essentially a multi-stage model that continues the selection process during training (see figure 1). This model uses few instruments in stage 1, a simple decision aid (Decision Aid 1), and relatively few applicants are eliminated. The remaining applicants enter training. During the first part of training, data, such as quiz scores, are collected on the trainees. These scores and scores from the selection instruments are fed into a decision aid (Decision Aid 2), which predicts how well each trainee will perform in the next stage of training. Based on these predictions, some of the trainees are eliminated. Of course, any trainees that fail the first stage of training are also eliminated.
The remaining trainees begin the second stage of training. Again, scores are collected on their performance during this stage of training. At the end of the second stage of training, any trainees who failed this stage are dropped from the program. The scores of the remainder, with the training scores from the first two stages of training and the selection instruments, are submitted to another decision aid (Decision Aid 3). Based on the results of the decision aid, some of the trainees are eliminated and the rest enter the third stage of training. This process is repeated until the trainees complete training and begin revenue flying. The single-stage model has rarely been used for pilot selection although it is commonly used to select applicants for other types of jobs. The cost of administering the selection instruments usually precludes this model; typically, simulator (or flight) evaluations and interviews by senior captains or management are very costly.

Multi-stage

The multi-stage method is the most commonly used model for pilot selection. To make this model cost efficient, the most inexpensive selection instruments are given first as the stage 1 tests. A large percentage of the applicants is eliminated from consideration at this point. Applicants continue to be eliminated after each stage but the percentage drops with each successive stage. Since transportation costs are not significant for many air carriers, the stages are often separated by many days. Applicants are sent home after the conclusion of each stage while the predictions are generated by a decision aid. Both the single-stage and the multi-stage models are best used when the number of applicants is large compared to the number of positions to be filled.
and training is expensive. In these situations missing a good person is not considered to be a large error. However, admitting someone to training who subsequently fails is considered to be a major mistake. The progressive model is best when training, at least initial training, is relatively inexpensive and the applicant pool is small compared to the number of positions that must be filled. In this situation, the company cannot afford to miss hiring good people, and missing a good person is considered to be a major error. In contrast allowing someone to enter training who subsequently fails is a minor mistake, assuming that the person leaves training early in the course.

Feedback loop

The fifth element of the selection system is the feedback loop. Its use unfortunately is more honored in the breach than in the observance. The primary purpose of the feedback loop is to determine if the selection system is actually identifying good applicants. Thus, it gathers the actual criterion data for the new hires and compares them to the predicted criterion scores. If the obtained criterion scores are close to what the decision aid predicted, then the system is working properly. If there are significant differences (determined mathematically) between the predicted and the obtained criterion scores, then the system needs to be evaluated and changes made to either the decision aid, the selection instruments, or the administrative procedures. One of the less recognized advantages of the feedback loop is that it improves corporate memory. Additionally, since the comparison process between the predicted and the obtained criterion scores results in a number, a company knows exactly how well its selection system is working.

Dr. Diane Damos has been involved in pilot selection and training for over 25 years. She received her doctorate in 1977 from the University of Illinois in aviation psychology. She was a professor of human factors for 19 years and most recently was on the faculty of the Institute of Safety and Systems Management at the University of Southern California. In 1995 Dr. Damos founded Damos Research Associates, Inc., a company specializing in pilot selection.