FAA Pilot Knowledge Tests: Learning or Rote Memorization?

Stephen M. Casner  
NASA Ames Research Center, Moffett Field, California

Karen M. Jones, Antonio Puentes  
San José State University Foundation, San Jose, California

Homi Irani,  
Embry Riddle Aeronautical University, Oakland, California
Acknowledgements

We thank all of the flight schools in the San Francisco Bay Area for their help with this research. Mikkel Grandjean-Thomsen and John Shopland provided valuable insights about knowledge testing in other countries.
Summary

The FAA pilot knowledge test is a multiple-choice assessment tool designed to measure the extent to which applicants for FAA pilot certificates and ratings have mastered a corpus of required aeronautical knowledge. All questions that appear on the test are drawn from a database of questions that is made available to the public. The FAA and others are concerned that releasing test questions may encourage students to focus their study on memorizing test questions. To investigate this concern, we created our own database of questions that differed from FAA questions in four different ways. Our first three question types were derived by modifying existing FAA questions: (1) rewording questions and answers; (2) shuffling answers; and (3) substituting different figures for problems that used figures. Our last question type posed a question about required knowledge for which no FAA question currently exists. Forty-eight student pilots completed one of two paper-and-pencil knowledge tests that contained a mix of these experimental questions. The results indicate significantly lower scores for some question types when compared to unaltered FAA questions to which participants had prior access.

Introduction

A requirement for most every FAA pilot certificate or rating is the knowledge test that each applicant must pass to demonstrate mastery of the aeronautical knowledge required to safely exercise the privileges of the certificate or rating being sought. Required by the U. S. Code of Federal Regulations, knowledge tests present applicants with a series of multiple-choice questions designed to assess applicants’ knowledge of aeronautical topics such as aerodynamics, weather, regulations, navigation, performance and flight planning. The questions that appear on every knowledge test are drawn from an item bank, a fixed database of questions created by the FAA for each pilot certificate or rating. For example, the item bank for the Private Pilot Airplane knowledge test currently contains 915 questions. Each Private Pilot Airplane knowledge test presented to an applicant consists of 60 questions selected from the 915-question item bank.

Since the mid-1980s, the FAA has made the item banks available to the public. This allows every pilot applicant to access, in advance, all of the questions that can potentially appear on every knowledge test. This situation has led to great controversy for two reasons. First, pilots now have the opportunity to limit their study to just those questions that appear in the item bank. Since the item bank does not contain questions pertaining to all knowledge required to fly safely, it seems possible for pilot applicants to skip over some knowledge and still achieve a high score on the knowledge test. Second, pilots have the opportunity to simply memorize the questions and answers, rather than develop an understanding of the knowledge that the questions aim to test.

The problems associated with neglecting to learn material that is required but that is not formally tested is evident and needs no further discussion. The negative outcomes associated with rote memorization strategies have been documented by more than a century of psychological research. In the case of the FAA knowledge tests, the crudest memorization strategies allow students to avoid looking at questions at all. Flash cards that contain an FAA question number on the front, and the correct answer letter designation on the back allow students to engage in the most primitive paired-associate learning
[Calkins, 1896; Mandler, 1970]. Other strategies link the letter designation for the correct answer to the content of the question. For example, the FAA Private Pilot knowledge test contains a flight-planning question using an airport named Addison. The answer to the question was traditionally “A.” Other psychologists have demonstrated the effectiveness of using simple mental images to successfully associate pairs of key words or phrases [Kothurkar, 1963; Paivio, 1965; Bower, 1972]. For example, 8 minutes is the answer to an FAA flight-planning question involving a flight to an airport called Redbird. The answer is easily recalled after visualizing a red bird flying in a figure-eight pattern. Strategies like these are particularly useful for test questions that otherwise require time-consuming and effortful calculations (e.g., weight and balance, flight planning, aircraft performance, etc.)

Even in the cases in which students read through the questions and memorize answers, the distinction between rote memorization and understanding has been demonstrated by other psychological studies [Tversky, 1973; Craik and Lockhart, 1972; Kieras and Bovair, 1984]. Students who achieve deeper levels of understanding are more successfully able to solve similar problems, recall solution steps, and use their knowledge in novel ways. Later research demonstrated the same phenomena in the domain of flight training. This research led to the same conclusions: when complex skills and concepts are to be learned, rote memorization is a poor substitute for understanding [Telfer, 1993; Moore and Telfer, 1990; Telfer, 1991]. Consequently, these ideas are emphasized in the FAA’s Aviation Instructor’s Handbook [FAA, 2001].

Indirect Evidence of Question Memorization

The FAA has been concerned with the validity of the knowledge tests for some time. Their principal concerns are, if applicants are focusing their study on questions and answers, that student understanding may be negatively impacted, and that the test scores being awarded to pilot applicants may not accurately reflect their knowledge.

Flouris (2001) made an initial attempt to relate differences in study methods to differences in knowledge test scores. Flouris compared the scores of students who completed a formal ground school at Auburn University, and students who studied on their own. Flouris stressed that the Auburn ground schools stress learning for understanding, and try to dissuade students from question memorization. Flouris found no significant difference between the two groups. Since the true study practices of both groups of pilots remain unknown, it is impossible to know what kind of preparation led to what test scores, or the real level of mastery of either group at the time of the tests. Flouris’ study cast further suspicion on the extent to which the FAA tests accurately measure knowledge, and motivated future research.

A compelling piece of evidence in support of the question-memorization hypothesis was gathered by the FAA. This evidence is the amount of time that test takers spend completing knowledge tests. Today, all FAA knowledge tests are administered by computer. The computer systems record the amount of time that test takers require to complete each exam, and the amount of time required to complete each question. Figure 1 shows timing data gathered by the FAA in 2002 for the knowledge tests required for the private and commercial
<table>
<thead>
<tr>
<th>Test</th>
<th>Ave. Completion Time (minutes)</th>
<th>Minimum Time (minutes)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Private Pilot – Airplane</td>
<td>73</td>
<td>2</td>
</tr>
<tr>
<td>Instrument Rating – Airplane</td>
<td>79</td>
<td>2</td>
</tr>
<tr>
<td>Commercial Pilot - Airplane</td>
<td>91</td>
<td>11</td>
</tr>
</tbody>
</table>

Figure 1: Average and minimum completion times for the Private, Commercial, and Instrument Airplane knowledge tests [Data provided by the FAA (FAA, 2003)].

The data show that many applicants complete the test in far less time than would be required for the average human to even read the questions and answers on the test.

Figure 2 shows timing data for a few individual test questions. These questions require applicants to work through complex calculations needed for flight planning under specified conditions. The data show that questions that would typically require several minutes to work through are being answered in merely a few seconds.

The FAA data clearly suggests that memorization is at work. For at least some test takers, the data clearly demonstrate that the test is not measuring the applicants’ ability to exercise the required knowledge and skills.

The evidence for question-memorization gathered by the FAA is compelling. However, there remains one credible refutation of the question-memorization hypothesis: it may be possible that pilot applicants satisfactorily learn the material, and then memorize the questions in the item bank in order to expedite the taking of the test. In other words, in can be argued that the FAA’s timing data demonstrates the presence of question-memorization, but it does not demonstrate the absence of understanding.

To directly address the concerns about the validity of the test, we conducted an experiment using private pilot students recruited from local flight schools. The experiment aimed to more directly measure what pilot applicants know at the time of their FAA knowledge test.
After departing GJT and arriving at Durango Co., La Plata Co. Airport, you are unable to land because of weather.

How long can you hold over DRO before departing for return flight to the alternate, Grand Junction Co., Walker Field Airport?

Total usable fuel on board, 68 gallons.
Wind and velocity at 16,000, 2308-16
Average fuel consumption 15 GPH.

A) 1 hour 33 minutes.
B) 1 hour 37 minutes.
C) 1 hour 42 minutes.

Mean Time: 569 seconds
Minimum Time: 6 seconds

An airplane descends to an airport under the following conditions:

Cruising altitude 6,500 ft
Airport elevation 700 ft
Descends to 800 ft AGL
Rate of descent 500 ft/min
Average true airspeed 110 kts
True course 335
Average wind velocity 060 at 15 kts
Variation 3 W
Deviation +2
Average fuel consumption 8.5 gal/hr

Determine the approximate time, compass heading, distance, and fuel consumed during the descent.

A) 10 minutes, 348, 18 NM, 1.4 gallons.
B) 10 minutes, 355, 17 NM, 2.4 gallons.
C) 12 minutes, 346, 18 NM, 1.6 gallons.

Mean Time: 195 seconds
Minimum Time: 5 seconds

Figure 2: Average and minimum completion times for flight planning questions [Data provided by the FAA (FAA, 2003)].
Method

Our approach to measuring pilot understanding of the required aeronautical knowledge was to administer an experimental knowledge test to private pilot applicants who had just prepared for and completed the FAA knowledge test. Our experimental knowledge test contained questions that varied from FAA questions in systematically increasing ways. Some experimental questions made cosmetic changes to existing FAA questions. Other experimental questions asked pilot applicants to use their aeronautical knowledge in different ways. The goal of our experimental test was to discover if pilot applicants’ knowledge of the required aeronautical knowledge was based on understanding, or on a more superficial review of the questions appearing in the FAA item bank. It was hypothesized that if applicants understood the required material, then their performance on the experimental questions would be comparable to their performance on FAA questions.

We distinguished two types of questions that appear in the FAA item banks. Skills questions require the applicant to exercise a procedure they had learned such as interpreting a chart, or working through a flight planning or airplane performance calculation problem. Knowledge questions require the applicant to recall a fact, or to reason about a concept they had learned.

To create our experimental item banks, we began with knowledge and skills questions that we randomly sampled from the FAA item banks. We left some of these FAA questions in their unaltered form to use as a control. We modified the remaining FAA questions to generate our treatment questions. Our selection and modification of the FAA knowledge and skills questions resulted in six different experimental questions types described below.

1. Unaltered Skills Questions

Figure 3 shows an example of an unaltered FAA skills question. Our purpose in including questions of this type was to establish a baseline measure for how well test takers performed on FAA skills questions, without having to ask them to reveal their score on the FAA test.

The wind and temperature at 12,000 feet MSL as reported by a pilot are

A. 080° at 21 knots and -7 °C.
B. 090° at 21 knots and -9 °C.
C. 090° at 21 MPH and -9 °F.

Figure 3: Unaltered skills question.
The wind and temperature at 6,000 feet MSL as reported by a pilot are

A. 280˚ at 26 knots and -8 °C.
B. 280˚ at 26 knots and 8 °C, measured.
C. 220˚ at 9 MPH and -8 °F.

Figure 4: Different figure skills question.

How should an aircraft preflight inspection be accomplished for the first flight of the day?

A. Thorough and systematic means recommended by the manufacturer.
B. Quick walk around with a check of gas and oil.
C. Any sequence as determined by the pilot-in-command.

Figure 5: Unaltered knowledge question.

2. Different Data Skills Questions
Figure 4 illustrates our only modification of the FAA skills questions. This modification substitutes different data (a different pilot report (PIREP) in this case) for the original data appearing in the question.

3. Unaltered Knowledge Questions
Figure 5 shows an example of an unaltered FAA knowledge question. These questions were intended as a control to establish a baseline score for each applicant on FAA knowledge questions.

4. Shuffled Knowledge Questions
Figure 6 illustrates the simplest of the modifications we made to FAA knowledge questions. This modification shuffles the order in which the answer choices appear in the question. This modification was designed to detect reliance on the crudest of memorization strategies that associate questions with memorized letters and answers (e.g., “A” for Addison).
How should an aircraft preflight inspection be accomplished for the first flight of the day?

A. Quick walk around with a check of gas and oil.
B. Any sequence as determined by the pilot-in-command.
C. Thorough and systematic means recommended by the manufacturer.

Figure 6: Shuffled knowledge question.

For the first flight of the day, an aircraft should be preflighted using

A. the procedure recommended by the manufacturer.
B. a walk around with a check of gas and oil.
C. a systematic procedure determined by the pilot-in-command.

Figure 7: Reworded knowledge question.

5. Reworded Knowledge Question
   Figure 7 illustrates our second and slightly more sophisticated modification of the FAA knowledge questions. This modification slightly rewords both the question and the answer. The rewordings used for these questions were limited to simple rearrangements of sentence structures and substitutions of non-critical words. In no case was a technical word or phrase (e.g., angle of attack) changed. This modification was designed to further test for memorization strategies based on simple question and answer recognition.

6. Different Knowledge Question
   Figure 8 illustrates our third and most sophisticated modification of the FAA knowledge questions. This modification extracts the target concepts tested by each question, and asks a slightly different question about the same concepts. Questions of this type were created by a group of three certified flight instructors who worked until consensus was reached on two key points: (1) that the modified question tested the same concepts as the original question; and (2) that the modified question was not more difficult than the original question. The aim of this type of question was to measure applicants’ performance when question-memORIZATION strategies were no longer possible.

   Figure 8 shows a modified question about wing flaps, along with all of the existing FAA questions about wing flaps. The question in Figure 8 illustrates our attempt to write questions that “fell between the cracks” of the existing FAA questions.
**Our Different Question About Wing Flaps**

One purpose of wing flaps is to
A - permit touchdown at higher airspeeds.
B - maintain a lower angle of attack at slower airspeeds.
C - permit safe flight at slower airspeeds during approach
and landing.

**Two Existing FAA Questions About Wing Flaps**

One of the main functions of flaps during approach and
landing is to
A - decrease the angle of descent without increasing the
airspeed.
B - permit a touchdown at a higher indicated airspeed.
C - increase the angle of descent without increasing the
airspeed.

What is one purpose of wing flaps?
A - To enable the pilot to make steeper approaches to a
landing without
   increasing the airspeed.
B - To relieve the pilot of maintaining continuous pressure
   on the controls.
C - To decrease wing area to vary the lift.

Figure 8: Different knowledge question.

**Apparatus**

Two different paper and pencil knowledge
tests were used for data collection. Each
knowledge test contained equal numbers
of randomly-sampled and ordered
questions drawn from the item bank of
question types described above. The first
knowledge test contained 50 questions,
and measured performance for the first
five question types. The second
knowledge test contained 20 questions,
and compared performance for the third
and sixth question types described above.

The number of questions appearing in our
experimental test was a compromise
between the statistical power needed, and
the expected fatigue and motivation levels
among our student pilot participants.

**Participants**

A total of 48 student pilots from local
flight schools agreed to participate in the
study. We attempted to recruit every
private pilot applicant that was scheduled
to take the FAA Private Pilot knowledge
test at every local flight school that offered
testing services. Every applicant that was both qualified and willing participated in the study. Each pilot completed exactly one of the two different experimental knowledge tests.

Participants were not told any details about the experiment in advance, and were told that they did not need to undertake any special studying or preparation for the experiment.

Student pilots who completed the 50-question experimental knowledge test received payment equal to the cost of their FAA knowledge test. Pilots who completed the 20-question experimental knowledge test received a NASA Aviation t-shirt.

**Procedure**

Our experimental knowledge tests were completed by student pilots during a scheduled appointment. Most student pilots took our experimental knowledge test directly after taking their FAA Private Pilot Airplane knowledge test, although some completed the test up to three days after completing the FAA test. Unlike the FAA knowledge test, we did not pose a time limit for the experimental test. When completing the test, participants were allowed to use the same materials permitted by the FAA for their knowledge tests (e.g., blank paper, calculators, pencils).

All participants were informed that their responses would remain anonymous, and no names were recorded.

**Results and Discussion**

Figure 9 shows the scores for the two knowledge tests that compared the six question types.

![Figure 9: Average scores for two tests that compared the six different question types.](image-url)
There was no difference between average scores for question types 3, 4, and 5: the unaltered, shuffled, and reworded knowledge questions. These results seem to rule out our worst fear: that participants relied solely on the crudest of memorization strategies in which learners used superficial cues available in the questions and answers. Shuffling and rewording questions had little effect on participants’ ability to answer questions correctly.

There was a significant difference \([F(1, 23) = 15.4, p < .001]\) between average scores for question types 1 and 2: the unaltered skills questions and the skills questions for which different data had been substituted for the original data. The means and standard deviations for the two question types were 87.9% (0.10) and 73.8% (0.18), respectively. This result seems to confirm our suspicions about the generality of the problem-solving skills being learned by students. Few instructors or evaluators would disagree – a student whose chart interpretation or flight-planning skills are limited to only those airports appearing in FAA test questions is operating under an obvious knowledge deficiency. Any test instrument that rewards such a pilot with a passing score is therefore also deficient.

There was also a significant difference \([F(1, 23) = 31.2, p < .001]\) between average scores for question types 3 and 6: unaltered knowledge questions and knowledge questions that participants did not have the opportunity to see in advance. The means for the two question types were 87.5% (0.12) and 64.6% (0.12), respectively. The less-than-passing scores may indicate fairly serious knowledge deficiencies, since these questions tested aeronautical knowledge that was also tested by existing FAA questions.

**Conclusion**

Our results support the FAA’s concerns that releasing test questions in advance may: (1) negatively affect the way students learn and ultimately understand required aeronautical knowledge; and (2) reduce the validity of the knowledge test as an assessment tool. If the design and execution of our study are accepted as reasonable, our data suggest that some pilot applicants may leave the test center with two things: (1) aeronautical knowledge that is deficient; and (2) a notarized document certifying that the student’s aeronautical knowledge is not deficient.

**Limitations of Our Study**

We must admit a number of limitations of our study. First, despite our experts’ efforts to control for question content and difficulty, the questions we created for question type 6 may have differed in content from, or may have been more difficult than existing FAA questions. Second, there is the possibility that some falloff in scores was due to some subjects’ unwillingness to work through the unfamiliar questions, accepting a lower score in the interest of time. Regardless, even for these “lazy” subjects, the falloff in scores between the two question types still suggests that question and answer memorization was used for the original FAA example questions, and their underlying knowledge of the material is not revealed by the current FAA test.

**Changes Made by the FAA to the Knowledge Test**

In July 2003, the FAA made a number of changes to the knowledge tests, based on their own investigations and on the evidence and inputs presented to them by outside parties. Our work was among the research that was reported to the governing
branch in the FAA. These changes include:

1. FAA question numbers are no longer associated with questions that are released to the public.

2. Answer choices are now randomly shuffled.

3. Only a few examples of each type of skills question are now made available. Questions appearing on the tests will be of the same form as the released question, but will use different data.

The first two changes were designed to foil crude memorization strategies that make simple associations between questions numbers and letter designations, or order of appearance among answer choices. The results for our shuffled-answers condition did not suggest a presence of this kind of memorization for knowledge questions. We did not, however, test our shuffling manipulation on skills questions. It may be the case that test takers sometimes do resort to this type of strategy for more difficult and time-consuming questions.

The third change agrees with our finding that significant numbers of test takers may have focused their study on individual questions rather than mastering the underlying skills being tested. There remains at least one credible threat to skills questions despite this countermeasure. Since skills questions require significantly more time to answer, any one test contains a small number of these questions. Under the current scoring system, all questions have the same value. Test takers may opt to simply guess at the more time-consuming skills questions, accepting a slightly lower overall test score.

One remedy for this problem is to increase the value of skills questions. There are a number of problems with this approach. First, it suggests that aeronautical knowledge areas that require more time to exercise are more important than other areas, a notion that is difficult to defend. Second, skills questions are often missed because of simple mathematical errors. Penalizing test takers several points for a simple math error seems unjustified.

**What Depth of Understanding Is Enough?**

In arguing for measures that might potentially lead to “deeper” understanding of aeronautical knowledge by students, we must recognize a few natural limitations.

A first limitation we encounter is that some aeronautical knowledge elements are simple facts, and that the learning of facts inevitably leads to a process of rote memorization. For example, Class D airspace extends by default to an altitude of 2,500 feet AGL. The FAA offers no particular theoretical reason why this altitude is chosen instead of 2,400 or 2,600 feet. The student is left to learn that 2,500 feet AGL is the correct number. For aeronautical knowledge of this type, we should be sure to include questions in the FAA item bank that exhaustively cover all of the facts to ensure that students are learning them.

A second limitation we encounter is that student understanding of any given topic is naturally limited at any given level of expertise. For example, a new student pilot’s understanding of flaps may be little more than a collection of memorized facts. Such a student simply has not yet had the opportunity to think through and experience the use of flaps in different flight situations; to develop theories about how they work, and to validate or invalidate those theories during practical experience. If we admit this limitation, care should be taken to ensure that the FAA item bank includes questions that cover a set of facts about flaps that will
allow them to act accordingly, for the time being, in all required flight situations. This collection of facts must serve as an acceptable substitute while the student progresses toward a more mature understanding that only further study, practice, and experience can provide.

**Knowledge Testing in the International Aviation Community**

It is interesting to compare FAA knowledge testing practices to those used in other countries. The European Joint Aviation Authorities (JAA) requires the applicant to pass seven separate knowledge tests for the Private Pilot License (PPL). These tests cover: air law and operational procedures, human performance and limitation, flight performance and planning, aircraft general and principles of flight, navigation and radio aids, meteorology, and radiotelephony. Some tests contain as few as 15 questions. All tests are multiple-choice and the JAA does not reveal the potential test questions in advance. Questions from old tests and questions similar to JAA test questions are available for student study, as well as a suggested reading list.

Similarly, the Australian Civil Aviation Safety Authority (CASA) requires applicants for the Private Pilot License to pass Basic Aeronautical Knowledge and Private Pilot License knowledge tests. Similar to JAA tests, these tests are multiple-choice and sample questions (not drawn from the existing item bank) are made available for student study.

India requires five written tests: navigation, meteorology, regulations, technical, and radio telephony. All tests are multiple choice, and the questions are not released to applicants in advance.

Whether aeronautical knowledge topics are tested in one or several tests, the move away from publishing test questions makes the FAA more consistent with testing practices used in other countries.
References


Federal Aviation Administration (2003). Personal communication.


The FAA pilot knowledge test is a multiple-choice assessment tool designed to measure the extent to which applicants for FAA pilot certificates and ratings have mastered a corpus of required aeronautical knowledge. All questions that appear on the test are drawn from a database of questions that is made available to the public. An ongoing concern about this practice is that the availability of the test questions may encourage students to focus their study on memorizing test questions. To investigate this concern, we created our own database of questions that differed from FAA questions in four different ways. Our first three question types were derived by modifying existing FAA questions: (1) rewording questions and answers; (2) shuffling answers; and (3) substituting different figures for problems that used figures. Our last question type posed a question about required knowledge for which no FAA question currently exists. Forty-eight student pilots completed one of two paper-and-pencil knowledge tests that contained a mix of these experimental questions. The results indicate significantly lower scores for some question types when compared to unaltered FAA questions for which participants had prior access.