THE INTERNATIONAL JOURNAL OF AVIATION PSYCHOLOGY, *11*(4), 359–379

Visual Flight Rules Flight Into Instrument Meteorological Conditions: An Empirical Investigation of the Possible Causes

Juliana Goh and Douglas A. Wiegmann
Aviation Human Factors Division
University of Illinois, Urbana-Champaign

Pilots' decisions to continue or divert from a visual flight rules (VFR) flight into instrument meteorological conditions (IMC) were investigated using a dynamic simulation of a hypothetical cross-country flight. Differences in situation assessment, risk perception, and motivation between pilots who chose to continue or divert from a VFR flight into an IMC situation were examined. Results indicate that the simulation was successful in identifying pilots who would choose to either continue or divert and that differences existed between these two groups of pilots. Accuracy of visibility estimates, appraisal of one's own skill and judgment, and frequency of risk-taking behavior were most important in predicting whether a pilot would continue or divert the flight. Findings suggest that overconfidence in personal ability and inaccurate diagnoses of visibility conditions precipitate VFR flight into IMC. More research is needed, however, to identify effective methods for remedying these problems.

An analysis of the National Transportation Safety Board's (NTSB) aviation accident database indicates that between 1990 and 1997, 2.5% of the more than 14,000 general aviation accidents were classified as involving visual flight rules (VFR) flight into instrument meteorological conditions (IMC). However, these VFR flight into IMC accidents accounted for approximately 11% of the fatalities in that 8-year period. Indeed, whereas the fatality rate of other general aviation accidents that did not involve VFR flight into IMC was 18%, 75% of all VFR flight into IMC accidents, in the same time period, were fatal (Goh & Wiegmann, 2001). The aviation

Requests for reprints should be sent to Douglas A. Wiegmann, University of Illinois, Aviation Human Factors Division, 1 Airport Road, Savoy, IL 61874. E-mail: dwiegman@uiuc.edu

accident records of other countries (e.g., the United Kingdom and New Zealand) also show similar trends, indicating that VFR flight into IMC is a major hazard in general aviation (O'Hare & Owen, 2000).

Explanations for why VFR pilots risk flying into deteriorating weather have focused on issues related to decision making. Some of these explanations have begun to be subjected to empirical investigation, whereas others have remained only speculative. The focus on decision-making issues stems from the recognition that good decision making is necessary to maintain safety in aviation (O'Hare, 1992). This recognition is a result of findings that fatal aviation accidents are more often associated with decision errors than minor accidents, which tend to be associated with procedural execution errors (Jensen & Benel, 1977; O'Hare, Batt, Wiggins, & Morrison, 1994; Wiegmann & Shappell, 1997). The purpose of this article, therefore, is to further examine the multiple factors that may affect pilots' decisions to continue VFR flight into IMC.

POSSIBLE FACTORS CONTRIBUTING TO VFR FLIGHT INTO IMC

As mentioned, several factors have been proposed to contribute to VFR flight into IMC. These include situation assessment, risk perception, motivation, and decision framing.

Situation Assessment

One reason why pilots may decide to continue a VFR flight into adverse weather is that they make errors when assessing the situation. That is, pilots are seen to engage in VFR flight into IMC because they do not accurately assess the hazard (i.e., the deteriorating weather conditions). The importance of accurate situation assessment has been alluded to in studies that have found that accurate diagnoses of problem situations are essential in good decision making. For example, in an analysis of flight and flight-related mishaps in the U.S. Navy and Marine Corps, Wiegmann and Shappell (1997) found that approximately 22% of the accidents that were due to human error resulted from diagnostic errors and that these diagnostic error-related accidents were more serious than those that were related to aircraft-handling errors. Research on naturalistic decision making has also found that an expert's performance is generally better than that of a novice's because experts are able to diagnose problem situations quickly and more effectively than novices (e.g., Klein, 1997).

Errors in situation assessment could be a result of several factors. First, "fair weather" pilots who lack experience interpreting changing weather conditions may be particularly susceptible to VFR flight into IMC situations. Second, the gradual transition from minimum VFR conditions to marginal VFR conditions to instrument flight rules (IFR) conditions could make discrimination of weather conditions difficult. Finally, poor hazard awareness (i.e., considering weather as an unlikely cause of an accident) could also precipitate poor situation assessment. There is also evidence that tiredness, fatigue, and increased workload can increase the likelihood of situation assessment errors (e.g., Wright, 1974). The extent to which these factors contribute to VFR flight into IMC, however, has yet to be empirically determined.

Risk Perception

Decision making under uncertainty also involves the perception of risk. Risk is defined as the likelihood of suffering a loss due to a hazard. In the case of VFR flight into IMC, pilots may assess the situation accurately (i.e., detect the deteriorating weather), but they may not realize the risks involved in continuing with the flight. Sanders and McCormick (1993) suggested that risk perception can be influenced by personal experience and ability, even when the hazard is diagnosed accurately. In other words, a pilot's overconfidence in his or her own ability may encourage continuation with the flight because of an unrealistic optimism about being able to avoid harm through personal control of the flight. Unfortunately, this overconfidence may be a by-product of pilot training, during which pilots are taught to be confident of their ability to control the plane. The U.K. Civil Aviation Authority (1988) cited "excessive optimism," "reluctance to admit limited capability," and "lack of appreciation of real dangers" as factors associated with pilot errors that lead to weather-related accidents. Also, O'Hare (1990) found, from administering the Aeronautical Risk Judgment Questionnaire, that general aviation pilots tended to exhibit low levels of risk awareness and higher than average self-appraisals of skill and judgment in flying. These findings suggest that risk perception should be addressed in any investigation of VFR flight into IMC.

Motivational Factors

Another possibility for why pilots engage in VFR flight into IMC has to do with motivation. In other words, pilots may diagnose and perceive the risks accurately, but other motivational factors bias their decisions. These motivational factors may be due to "Get-home-itis" or other personal or social pressures. These pressures are considered to bias the pilot's decision to continue with the flight even though an assessment of the situation suggests otherwise. O'Hare and Smitheram (1995) asked pilots to indicate how important eight personal and social factors were in contributing to their decision to continue or divert from the flight. Their results indicate that factors relating to self-motivation were more important than those relating to others.

Decision Framing

According to Prospect Theory (Kahneman & Tversky, 1982), a person's choice between a risky and safe course of action depends on whether he or she frames an option as a gain or a loss. In the case of VFR flight into IMC, if pilots frame the decision to divert as a loss (e.g., wasted time, money, and effort), the pilot might tend to be risk-seeking and choose to continue with the flight. On the other hand, if the pilot frames diverting as a gain (e.g., it is safer), he or she should be risk-averse and choose to divert from the flight. O'Hare and Smitheram (1995) investigated the effects of framing on a pilot's decision to continue or divert from a VFR flight into IMC situation. These investigators found that pilots were less likely to continue with the flight into adverse weather when they were encouraged, via experimental manipulations of the alternatives, to frame VFR into IMC in terms of losses and diverting in terms of gains. However, this did not occur when pilots used their own natural frames.

LIMITATIONS OF PREVIOUS RESEARCH

Many of the factors that could potentially influence a pilot's decision to continue VFR flight into IMC have either not been investigated empirically or have only been examined in isolation. For example, the effects of risk perception and framing have been examined in separate studies, whereas errors in situation assessment have not been investigated within the context of VFR flight into IMC. Furthermore, participants in many of these studies have had to respond only to static information about hypothetical cross-country flights. These static simulations do not represent the true workload conditions involved in having to fly a plane and, at the same time, getting and processing relevant information so that a decision can be made. It is under such high workload conditions that errors in decision making often arise, and a valid method of investigation would require this characteristic to be part of the simulation.

One of the major reasons for the current lack of consistency in both explaining and researching VFR flight into IMC is that many of the investigators currently involved in these research efforts generally come from different academic backgrounds or philosophies and therefore approach the issue from different perspectives. As such, different investigators have used different methodologies to study decision making, and hence, comparisons of different explanations about pilot decisional factors have not been attempted. What is needed, therefore, is a unifying theory or framework that ties these different approaches together so that the multiple factors affecting pilots' decisions to continue VFR flight into IMC can be systematically investigated and compared.

AN ORGANIZING FRAMEWORK

One possible framework for conceptualizing the factors that potentially affect pilot decision making is Jensen's (1995) Judgment Model. The model, shown in Figure 1, describes multiple stages involved in the judgment process that might take place in aeronautical decision-making situations, such as the decision to continue to VFR flight into IMC. Each stage of the model is described briefly, as well as how the model accounts for the multiple factors that have been proposed to affect pilots' decisions to continue VFR fight into adverse weather.

Jensen's Judgment Model

Problem vigil. This stage describes the constant state of vigilance the pilot is in so that he or she can detect changes in the environment. Detection of these changes takes place through the various human senses (e.g., vestibular system, kinesthetic system) and can be affected by the pilot's attention, workload, and fatigue.

Recognition. This second stage in the judgment process is when the pilot realizes that changes in the environment could affect the safety of the flight. The pilot's perceptions and expectations of the situation can affect his or her ability to rec-

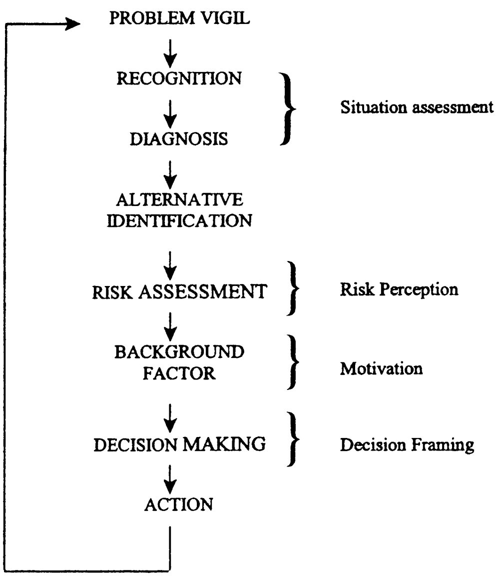

FIGURE 1 Detailed judgment model. From *Pilot Judgment and Crew Resource Management* (p. 37), by R. S. Jensen, 1995, Brookfield, VT: Ashgate Publishing Company. Copyright 1995 by Ashgate Publishing Company. Adapted with permission.

ognize that a problem has emerged. These difficulties are well documented in the literature on bottom-up and top-down processing (Wickens & Hollands, 2000).

Diagnosis. During diagnosis, the pilot attempts to understand the nature of the problem. This is affected by the assessment he or she makes of the situation and his or her experience and knowledge in dealing with similar situations. It requires the pilot to consider changes in the environment and how these changes can have an effect on the current state. Knowledge of hazards and the risks associated with these hazards may help the pilot make a correct diagnosis, which is important for a correct decision to be made.

Alternative identification. This stage involves the identification of various courses of action that can be taken to resolve the problem. The number of alternatives that one can generate may ultimately affect the quality of a decision.

Risk assessment. The pilot tries to determine the risks associated with the alternatives that have been generated at this stage of the judgment process. This is very much affected by the next step in the judgment model, the background factor.

Background factor. The background factor can affect all stages of the judgment process, but it has particular importance just as a decision is about to be made. According to Jensen (1995), "it represents the motivational forces that keep us from following purely rational decision functions" (p. 46).

Decision making. At this stage of the process, the pilot makes a decision or chooses which course of action to take.

Action. This stage involves application of the pilot's decision. These actions may involve interacting with air traffic control (ATC), passengers, and moving flight controls.

When applying Jensen's (1995) framework to VFR flight into IMC, it becomes clear that the many explanations that have been proposed are neither independent nor conflicting. Indeed, Jensen's framework makes it clear that each of the different perspectives is related, differing only in the stage of the judgment process that it is experienced. Explanations related to situation assessment and risk perception fit in the earlier stages of Jensen's model. It is during these stages that pilots attempt to recognize the hazard in the environment and try to make a diagnosis. Motivation factors and decision framing come into effect in the later stages of the decision-making process, as shown in Figure 1. At these later stages, a correct decision may not be made even if the assessment and diagnosis of the situation are accurate, as the motivational pressures and framing of the situation have the effect of biasing the pilot's decision.

Purpose of This Study

Given that Jensen's (1995) Judgment Model is correct, or at least provides an adequate explanation of the relations among variables affecting pilots' decisions to continue VFR flight into IMC, it would seem logical to examine these factors together. Furthermore, Jensen's model clearly indicates that pilot judgment is a dynamic, iterative process that poses temporal demands and limitations on the individual decision maker. Therefore, it necessitates the use of a more "real-world" simulation of VFR flight to test and understand the underlying nature of pilots' judgment processes.

The purpose of this study, therefore, was to investigate the extent to which cognitive and motivational factors such as situation assessment, risk perception, and motivation affect pilots' decisions to continue or divert from adverse weather, using a dynamic simulation of a VFR flight into IMC situation. The potential effects of decision framing were not examined in this study, however, as the main purpose was to obtain a profile of the pilots who tend to either continue or divert from adverse weather conditions, without having to manipulate the pilots' decisional processes. The results of this study, therefore, provide insight into both the individual and combined contribution of the multiple factors that naturally affect pilots' decision to continue VFR flight into adverse weather.

METHOD

Participants

Thirty-two non-instrument rated pilots (28 male, 4 female) from the University of Illinois's pilot-training program participated in the study. The participants' total VFR flight experience ranged from 30.0 to 259.4 hr (*Mdn* = 60.0 hr), whereas total solo flight hours ranged from 3.5 to 150.3 hr (*Mdn* = 11.5 hr). Only 14 had actual IFR flight experience, and this IFR experience ranged from 0.3 to 10.0 hr. All pilots had flown at least 1 solo cross-country flight (*Mdn* = 3) and between 2 and 13 dual cross-country flights (*Mdn* = 3) at the time of the study. Their ages ranged from 18 to 47 years (*Mdn* = 19 years). Participants were compensated $6 an hour for their participation.

Apparatus and Materials

Pre-experimental questionnaire. The pre-experimental questionnaire was based on the Aeronautical Risk Judgment Questionnaire (O'Hare, 1990). Participants were required to answer questions regarding the following:

1. Background information: Information on age, gender, total flight hours (dual and solo), total VFR hours, total IFR hours (simulated and actual), total hours of cross-country flight, and total number of cross-country flights (dual and solo) were recorded.

2. Self-judgment: Using a 7-point Likert scale, participants were required to rate their own skill and judgment compared to other pilots, ranging from 1 (*much worse*) to 7 (*much better*), and their propensity to take risks compared to other pilots, ranging from 1 (*very unwilling*) to 7 (*very willing*). The frequency of risk-taking was also rated on a scale from 1 (*very infrequently*) to 7 (*very frequently*).

3. Hazard awareness: This was assessed through two types of questions. First, participants were required to estimate the percentages of accidents that are attributable to six broad causal factors (e.g., power plant or engine failure, pilot error, weather). Second, they ranked seven specific factors (e.g., flying into adverse weather, spatial disorientation, fatigue) in terms of how likely they are to cause accidents, ranging from 1 (*most likely*) to 7 (*least likely*). Participants were asked to estimate these in two ways: (a) the likelihood that the factors would cause accidents in general and (b) the likelihood the factors would cause an accident they may be involved in.

4 Risk awareness: Participants were asked to assess the risks of being involved in an accident as a passenger in a commercial airliner, riding a motorcycle, and flying a general aviation flight. The options ranged from 1 in 1,000,000 to 1 in 10.

X-plane flight simulation program. A Pentium III 450MHz computer with 320Mb RAM was used to run the X-Plane Version 5.01 (Laminar Research Inc., 1998) flight simulation program. A yoke and rudder pedals were used by the participants to interact with the program. The simulation was presented on a 22-in. SVGA Dell monitor. The X-plane flight simulation program allowed real time and programmable weather. This included selection of cloud ceiling, visibility, wind direction, and wind speed. A database of airports from around the world was also available for inclusion into the program. Terrain and real-world structures (e.g., power line towers, smoke stacks, buildings) could also be created and manipulated. The program also allowed collection of flight parameters such as altitude, distance traveled, pitch and roll, latitude and longitude, and airspeed.

Postexperimental questionnaire. On completion of the flight simulation program, participants were required to answer questions on a postexperimental questionnaire regarding the following:

1. Situation awareness: Participants were asked questions on the amount of time that had elapsed during the flight, when the program was terminated, and the distance they were from their origination and destination. They were also required to state what they perceived the weather conditions (visibility, cloud ceiling, wind

direction, wind speed, temperature, and precipitation) to be when the program was terminated.

2. Self-judgment: Participants were asked to judge their own skill and judgment compared to other pilots of similar experience, ranging from 1 (*much worse*) to 7 (*much better*), and their willingness to take risks compared to other pilots of similar experience, ranging from 1 (*very willing*) to 7 (*very unwilling*), in the situation they had just encountered. They also rated how frequently they took risks, ranging from 1 (*very infrequently*) to 7 (*very frequently*).

3. Decisional factors: Participants were asked to rate the importance of 12 factors in terms of importance in their decision to divert or continue with the flight. Similar factors were used by O'Hare and Smitheram (1995). The 12 factors were grouped into the following four categories: (a) tangible gains and losses for self, (b) tangible gains and losses for others, (c) self-approval and disapproval, and (d) social approval and disapproval.

Procedure

Participants signed a consent form after reading a brief description of the study. They were then asked to complete the pre-experimental questionnaire. Participants then read the scenario, which explained that they were going to fly two routes, the first of which was a practice flight from Champaign to Terre Haute, to allow participants to familiarize themselves with the controls. In the second, experimental flight, they were to imagine they were on a solo cross-country flight from the Champaign airport to the Illinois Valley Regional airport (approximately 85 nautical miles), as part of their training required to earn a private pilot's license. In both the practice and experimental flights, participants were provided with a map and flight plan, which detailed the routes, landmarks, and alternate airports along these routes. They were also given the same Terminal Aerodrom Forecast and Winds Aloft information. Participants were given as much time as they needed to study the map and flight plan for each route.

Participants were introduced to the simulator set-up and told they were flying a Cessna 172. All flight instruments and controls were clearly labeled, including the distance measuring equipment and clock for determining time and distance flown. All participants flew approximately 15 min of the practice route. They were then given the map, flight plan, and weather information for the experimental route that had an estimated time in route of approximately 1 hr. The participants were told to treat the flight as they would any other flight that they may make in the real world. They were also told to be aware of any possible failures, mechanical or otherwise, that might occur during the flight. It was also indicated that failures might not necessarily occur. However, they were instructed that should they decide to divert for any reason, they should toggle a predetermined switch and then proceed to any airport on the sectional map they deemed appropriate.

Weather conditions at take-off were just above VFR conditions (5 miles visibility, 5,000-ft mean sea level [MSL] cloud ceiling, overcast) until approximately 45 min into the flight when conditions deteriorated to below VFR minimums (2 miles visibility, 1,500-ft MSL cloud ceiling). Participants were given a 5-min window from the time the weather deteriorated to below VFR minimums to make their decision whether to discontinue the flight. If the participant continued flying at the end of the 5-min window, he or she was considered to have made the decision to continue with the flight. In either case, the simulation was terminated at this point in the flight and the screen was blanked. Following completion of the simulation, participants completed the postexperimental questionnaire. They were then thanked and compensated for their time and participation and dismissed.

RESULTS

The analysis of the data consisted of five phases, which are reflected in the organization of the Results section. The first phase examined the percentage of pilots who chose to either continue or divert the flight after encountering the adverse weather. The second phase focused on examining possible differences between those pilots who chose one course of action over another (i.e., to continue vs. divert the flight) based on their responses on the pre- and postexperimental questionnaires. The questions were related to experience, self-judgment, hazard awareness, and risk awareness. The third phase focused on situation awareness, and the fourth phase focused on responses to postexperimental questions related to personal and social factors that may have affected the participants' decisions to continue or discontinue with the flight. The fifth and final phase examined the relative contribution of several factors to a pilot's decision of whether to continue with the flight, using a multivariate discriminant analysis.

Decision to Continue or Divert

Of the 32 pilots who participated in this study, 22 (68.75%) chose to fly into the deteriorating weather, whereas only 10 (31.25%) chose to divert the flight. Assuming an equal likelihood (50–50 chance) that pilots would chose to either continue or divert the flight, the observed proportions exceeded chance expectations, $\chi^2(1, N = 32) = 4.5, p < .05$, suggesting that pilots in this study, on average, were more likely to continue the flight rather than divert.

Background Information

In general, participants varied greatly in terms of the number of total VFR flight hours, which ranged from a high of approximately 259 hr to a low of 30 hr (*Mdn* =

78.2). They also varied greatly in age, with the median age being 20.8 years (range = 18 to 47 years). Slightly less than half of the participants were certificated pilots ($n = 12$). The remaining participants possessed student pilot certificates. However, median tests performed on all background variables did not reveal any statistically significant differences between pilots who decided to continue with the flight (i.e., the continue group) and those who chose to divert (i.e., the divert group).

Self-Judgment

Participants provided self-ratings of their skill and judgment, as well as their willingness and frequency of taking risks. These questions were asked on both the pre- and postexperimental questionnaires. The pre-experimental questionnaire asked participants to rate these items compared to the "average" pilot in general. The postexperimental questionnaire items required participants to rate themselves compared to an average pilot with regard to the specific flight scenario they had just encountered in the experiment. High scores indicate a higher self-rating on each item, with a score of 4 indicating a self-rating of *average*.

A 2 × 3 × 2 mixed analysis of variance (ANOVA) was used to analyze participants' responses to these questions. Decision Group (continue vs. divert) was a between-groups factor. Both Question Type (skill level, willingness to take risks, and risk-taking frequency) and Questionnaire (pre vs. post) were within-participants factors. Results of this analysis revealed a significant main effect for Question Type, $F(2, 30) = 23.99, p < .001$. In general, pilots rated themselves as being better than the average pilot in terms of skill and judgment ($M = 4.52$), $t(31) = 5.61, p < .001$. However, they indicated they were generally less willing to take risks ($M = 3.7$), $t(31) = 1.9, p = .06$, and that they took risks less frequently than the average pilot, ($M = 3.12$), $t(31) = 5.49, p < .001$. The Question Type × Questionnaire interaction, was also significant, $F(2, 30) = 4.88, p < .05$, with the pilots rating themselves on the postexperimental questionnaire as having less skill in situations similar to the one they had encountered in the experiment than on the pre-experimental questionnaire, $t(31) = 3.97, p < .001$. No differences were found for willingness to take risks and frequency of risk-taking behavior between the pre- and postexperimental questionnaires, $p > .05$.

A significant Decision Group × Question Type interaction was also observed, $F(2, 30) = 5.18, p < .01$. This interaction is depicted in Figure 2. An examination of this figure reveals that pilots in the continue group had significantly higher self-ratings of skill and judgment ($M = 4.64$, $SD = .5$) than individuals in the divert group ($M = 4.25$, $SD = .49$), $t(30) = 2.05, p < .05$. Also, although those in the continue group tended to have higher ratings of their willingness to take risks ($M = 3.85$, $SD = .87$ vs. $M = 3.35$ $SD = .94$), this difference was not significant, $p > .05$. Finally, participants in the continue group tended to rate themselves as slightly less

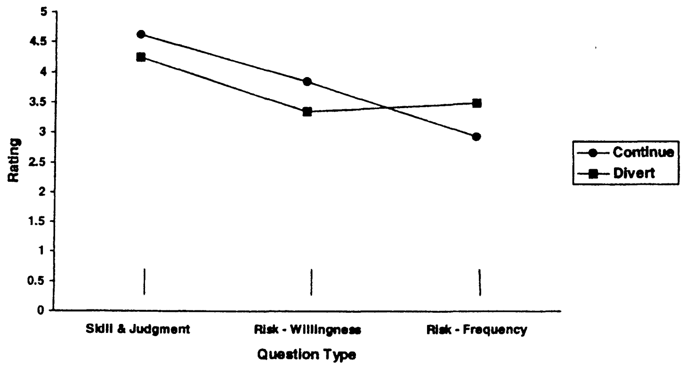

FIGURE 2 Ratings on skill and judgment, willingness to take risks, and frequency of risk taking as a function of Decision Group.

frequent risk takers ($M = 2.94$, $SD = .89$) than participants in the divert group ($M = 3.5$, $SD = .88$), but this difference was not significant, $p > .05$.

Hazard Awareness

Global causal factors. Participants estimated the likelihood that given an accident had occurred it would have been the result of a specific "causal factor." They made these estimates for both accidents in general and a hypothetical accident in which they might be involved personally in the future. For clarity purposes, participants were allowed to make these estimates using a percentage scale; however, for purposes of analysis and interpretation, these values were converted to likelihood estimates by dividing each value by 100 (i.e., percent estimate/100). Graphical depictions of the "weather" and "pilot error" likelihood estimates, which are of particular interest to this study, are presented in Figures 3 and 4, respectively. The "general" estimates refer to the pilots' perception of all accidents, whereas the "personal" estimates refer to an accident in which the pilots themselves might be involved.

Weather. A 2 (decision group) × 2 (type of accident: personal vs. general) mixed ANOVA was used to analyze participants' estimates pertaining to the weather factor. This analysis revealed a significant main effect for Group, $F(1, 30) = 4.436$, $p < .05$. An examination of Figure 3 reveals that pilots in the divert group had higher estimates of the likelihood that weather would be a cause of an accident ($M = .50$, $SD = .23$) than pilots in the continue group ($M = .34$, $SD = .18$). The difference between groups was slightly larger for the general variable than the personal variable; however, no significant interaction was observed.

Pilot error. A 2 (decision group) × 2 (type of accident: personal vs. general) mixed ANOVA was also performed to analyze participants' estimates of the likelihood that pilot error would be a cause of an accident. Figure 4 shows the participants' responses.

The analysis revealed a significant interaction between these factors, $F(1, 30) = 6.345$, $p < .05$. The divert group estimated that pilot error was equally likely to be the cause of an accident, whether they were personally involved or not ($M = .277$, $SD = .21$, for general; $M = .299$, $SD = .19$, for personal), whereas those in the continue group rated pilot error as significantly more likely to be a cause of accidents

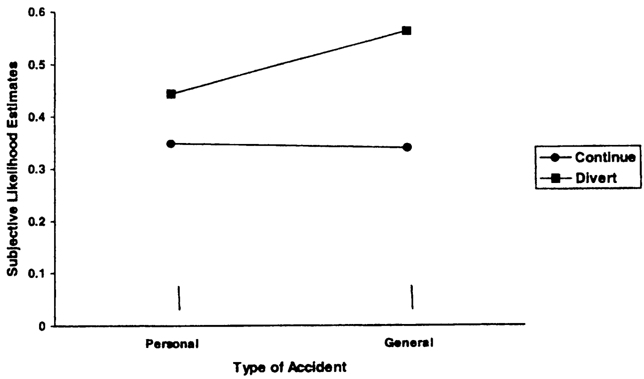

FIGURE 3 Mean likelihood estimates that an accident would have been caused by weather as a function of group and type of accident.

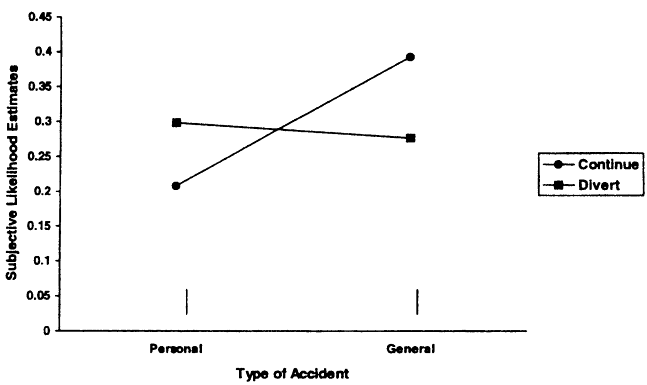

FIGURE 4 Mean likelihood estimates that an accident would have been caused by pilot error as a function of group and type of accident.

in general ($M = .393$, $SD = .24$) than if they were personally involved ($M = .208$, $SD = .19$), $t(21) = 3.763$, $p < .01$. As a result, pilots in the continue group had somewhat lower estimates than pilots in the divert group on the personal variable, but slightly higher estimates on the general variable; however, these between-group differences were not statistically significant.

Specific pilot factors. Participants rank-ordered a list of specific pilot factors in terms of the likelihood that they would be the cause of an accident, given an accident had occurred. Again, participants made these estimates for both accidents in general and a hypothetical accident in which they might be involved personally in the future.

General accident occurrences. On average, pilots ranked Flying into Adverse Weather and Exercised Poor Judgment the most likely causes of accidents ($Mdn = 2.0$). Spatial Disorientation and The Misuse of Flight Controls had a tied ranking as the second most likely ($Mdn = 4.0$), whereas Violation of Rules was ranked second least likely to cause an accident ($Mdn = 5.5$).

Personal accident occurrences. Overall, the pilots also ranked Flying into Adverse Weather as the factor that would most likely be a cause of an accident ($Mdn = 2.0$) in which they might be involved. Spatial Disorientation was the second highest most likely factor ($Mdn = 2.5$). Exercised Poor Judgment and Fatigue were tied for the third most likely cause ($Mdn = 3.0$), whereas Violation of Rules and Misuse of Flight Controls were tied for the lowest ranking by the pilots ($Mdn = 6.0$).

Statistical comparisons of these variables using a Wilcoxon Sign Rank Test revealed significant differences between pilots' rankings of general and personal variables for Misuse of Flight Controls ($p < .01$), Problems with ATC ($p < .01$), and Fatigue ($p < .01$). Mann–Whitney tests, however, did not reveal any differences in rankings between the continue and divert groups on either general or personal variables, $p > .05$.

Risk Awareness

Overall, 19 of the 32 pilots (59.4%) were accurate in indicating the higher risks involved in general aviation verses commercial aviation. Specifically, 54.5% of the continue group correctly indicated this trend, whereas the rest of the group indicated that the risks were equal or less so for general aviation than for commercial flights. The proportions were slightly better for the divert group, with 70% making correct risk assessments. However, a chi-square analysis did not reveal any statistically significant relation between accuracy of risk awareness and decision to continue or discontinue with the flight, $p > .05$.

Situation Awareness Variables

Distance and time. Analyses were based on how much the pilots' estimates deviated from the actual distances and times. Actual distances and times were subtracted from estimated distances and times. A negative value would indicate underestimation (further or more time away from location than estimated) of distance and time from the departure and destination airports. A positive value would indicate an overestimation (nearer or less time away from location than estimated), and a value of zero would indicate perfect estimation. Table 1 presents a summary of the mean differences between actual and estimated distance and time from both the departure and destination airports as a function of decision group.

Participants generally underestimated the distance traveled from the departure airport, as well as their distance to the destination airport. A total of 21 participants underestimated the distance they had already traveled from the departure airport, $\chi^2(1, N = 31) = 3.9$, $p < .05$. A somewhat smaller and nonsignificant proportion of pilots ($n = 19$) also underestimated their distance from the destination airport. Although pilots in the divert group tended to have more accurate distance estimates than pilots in the continue group for both the departure and destination airports, a Mann–Whitney test indicated that these differences were not significant, $p > .05$.

For time estimates, almost all of the pilots ($n = 31$) underestimated the time they had already traveled from the departure airport, $\chi^2(1, N = 32) = 28.13$, $p < .001$. However, less than half ($n = 15$) underestimated the amount time they needed to travel to reach the destination airport, $p > .05$. Again, pilots in the divert group tended to have slightly more accurate time estimates than participants in the continue group for both the departure and destination airports, but a Mann–Whitney test indicated that these differences were not significant, $p > .05$.

TABLE 1
Mean Differences Between Actual and Estimated Distance and Time From Departure and Destination Airports as a Function of Decision Group

	Continue		*Divert*	
Difference Between Estimated and Actual	*M*	*SD*	*M*	*SD*
From departure airport				
Distance (nautical miles)	–11.21	13.65	–5.38	15.42
Time (minutes)	–3.51	3.47	–3.24	2.05
From destination airport				
Distance (nautical miles)	–1.29	2.90	0.88	7.41
Time (minutes)	–0.54	1.90	0.29	3.35

Weather conditions. Differences between estimated and actual visibility and cloud ceiling were calculated in the same manner as that for distance and time. Similarly, a negative value would indicate an underestimation of the conditions (thinking conditions are worse than they actually are), whereas a positive value would be an overestimation (thinking conditions are better than they actually are). The mean differences between actual and estimated visibility and cloud ceiling as a function of decision group is presented in Table 2.

Overall, a significant proportion of the participants ($n = 21$) overestimated the visibility, $\chi^2(2, N = 31) = 16.71, p < .001$, as well as cloud ceiling ($n = 27$), $\chi^2(1, N = 28) = 24.14, p < .001$. A Mann–Whitney test revealed, however, that the continue group's visibility estimates were significantly higher than divert group's estimates ($U = 33, p < .01$). Group differences in cloud ceiling estimates were not significant, $p > .05$.

Decisional Factors

The rankings of importance of 14 items in the pilots' decisions to continue or divert, grouped into two categories of Factors Pertaining to Self and Factors Pertaining to Others, are presented in Table 3 for both groups. Items were coded such that higher values indicate greater importance.

Average "self" and "other" ratings were analyzed using a 2 (decision) × 2 (factors type: self vs. others) mixed ANOVA. The results revealed only a significant main effect of Factor Type, $F(1, 30) = 130.378, p < .001$. On average, pilots indicated that personal factors in general had a bigger impact on their decisions to either continue or divert a flight than do factors related to others. The most important personal factor was fear of "not being able to handle the situation," whereas the personal factor of lowest importance was the "convenience of being able to get in your own car and going home." The most important social variable was "not killing anyone" and of least importance was of passengers being "disappointed with your abilities as a pilot if you diverted."

TABLE 2
Mean Differences Between Actual and Estimated Visibility and Cloud Ceiling as a Function of Decision Group

	Continue		*Divert*	
Weather Parameter	*M*	*SD*	*M*	*SD*
Visibility (nautical miles)	1.41	1.05	0.00	1.00
Cloud ceiling (feet)	2,269.5	1,206.45	2,212.50	994.90

TABLE 3
Mean Rankings of Importance of Factors Pertaining to Self and Others as a Function of Decision Group

	Continue		*Divert*	
	M	*SD*	*M*	*SD*
Factors pertaining to self				
The convenience of getting your own car and going home	3.91	2.14	2.40	1.65
The possibility of damaging your aircraft	5.95	1.21	5.90	1.10
The opportunity of exercising skill and judgment in difficult conditions	5.82	1.37	4.80	1.93
The possibility that you might not be able to handle the conditions	6.23	0.92	6.60	0.52
"Self" average	5.48	0.73	4.93	0.80
Factors pertaining to others				
Getting back so that the next person hiring the aircraft is able to take the aircraft on time	2.32	1.32	1.50	0.97
The loss of revenue to the club if the aircraft is taken off line for repair	3.14	1.55	2.70	1.89
Your passengers would be impressed by your skills as a pilot if you continued on	1.86	1.21	1.90	1.20
Your passengers would be disappointed in your abilities as a pilot if you diverted	2.05	1.09	1.40	0.52
For other pilots to think positively of your flying skills	3.32	1.81	3.30	1.77
For you to keep a good reputation as a pilot in general	4.86	1.70	4.00	2.54
Not to kill anyone	6.82	0.50	7.00	0.00
Not to damage public or private property other than the airplane	5.73	1.52	6.00	1.41
"Other" average	3.76	0.72	3.48	0.92

Discriminant Analysis

The relative contribution of the factors affecting pilots' decisions whether to continue or divert was examined using a multivariate discriminant analysis. A step-wise entry of the following eight factors was carried out: estimates of weather-related estimates (personal and general accidents), estimates of pilot-error related accidents (personal and general accidents), skill and judgment ratings, willingness to take risks, frequency of risk-taking behavior, and visibility estimates. These factors were selected based on the previous analyses, which showed that these factors were able to differentiate between the groups.

The results of the discriminant analysis revealed that visibility estimates, skill and judgment ratings, and frequency of risk-taking behavior were most important in predicting whether a pilot would continue or divert from a VFR into IMC situation as simulated in this study, $\chi^2(3, N = 31) = 18.72, p < .001$. Visibility estimates had the highest discriminant function coefficients (.929), followed by skill and judgment (.602) and frequency of risk-taking behavior (–.562). Combined, these three factors were able to predict whether a pilot would continue or divert with 87.1% accuracy.

DISCUSSION

The simulation used in this study was successful in identifying pilots who would continue and those who would divert in a VFR flight into IMC scenario. The results also suggest that these two groups of pilots have different profiles in terms of confidence in their own skill and judgment, willingness to take risks, and frequency of risk-taking behavior; estimates of pilot error and weather as hazards in general aviation; and accuracy of situation assessment.

Pilots who chose to continue with the flight had higher ratings of skill and judgment, suggesting that they had greater confidence in their abilities to control the aircraft than the pilots who chose to divert. Furthermore, pilots who chose to continue rated themselves as more willing to take risks than pilots who chose to divert. Together, these two group differences suggest that because of greater confidence in their own piloting abilities, the pilots who continued were more willing to risk flying into the adverse weather. The pilots who continued also indicated that they took risks less frequently than pilots who chose to divert. This may imply that the pilots who continued considered themselves to be "tactical decision makers" in that they were willing to take risks but were more careful of which risks to take.

The two groups of pilots also differed in their estimates of weather and pilot error as hazards in general aviation. The group that continued judged these hazards as less likely causes of accidents than the group that diverted, indicating that pilots who continued considered these hazards as less of a threat to flight safety compared to the pilots who diverted. The differences in these estimates for general and personal accidents for those who continued also suggest that this group of pilots felt they were less vulnerable to these possible causes of accidents. For example, although they estimated that approximately 56% of accidents in general were caused by weather, they estimated that there was only a 45% chance that weather would be a cause of an accident they were involved in. Similar trends were observed for ratings of pilot error, and this conforms to the finding that pilots who continued had greater confidence in their piloting abilities than those who chose to divert. It should be noted that no actual differences in experience and training were observed between the groups of pilots who chose to continue or divert. Pilots in both groups had similar flight hours and certification. Therefore, the differences in self-judgment on skill and ability appear to reflect differences in self-awareness or metacognitive estimates of one's own abilities.

Another difference found between the two groups of pilots was the accuracy of the assessment they made of the situation and, in particular, their estimates of visibility. Pilots who chose to divert were more accurate in their visibility estimates than the pilots who chose to continue. The importance of this factor was revealed by the discriminant analysis, which showed accuracy of visibility estimates to be the most important factor in predicting whether pilots would continue or divert from a VFR flight into IMC situation. This result suggests that pilots who continued with the flight may have done so because they did not perceive visibility conditions to be below VFR minimums. With a more accurate perception, these pilots' decisions may have been more appropriate to the situation, according to Rasmussen's (1982) notion of rule-based behavior. Rasmussen described rule behavior as the selection of appropriate responses based on a stored set of "if (condition) then (action) rules." In the case of encountering adverse weather, if pilots accurately assess visibility to be too low, they would then choose to divert from the situation, as was observed of the divert group of pilots in this study.

Together, these findings suggest that pilots who continued VFR flight into IMC made errors early in the decision-making process in the form of inaccurate assessments of visibility, and this was compounded by other factors such as their greater willingness to take risks, greater confidence in their flight skills, and a reduced sense of vulnerability to weather hazards and pilot error.

Although differences were found in the factors, the pilots did not differ in their risk awareness. The lack of a group difference in risk awareness could have been a result of the insensitivity of the measure that was used. The pilots were asked to rate the riskiness of general and commercial aviation flights, questions that may have been too general to detect differences between the groups. More specific questions related to risks involved in VFR flight into IMC scenarios may have been more appropriate for the purposes of this study.

Ratings of the importance of personal and social pressures on the pilots' decisions also did not differ between groups. A likely explanation for this is that the questions did not describe situations that the pilots were experiencing during the simulation. For example, pilots were asked to rate the importance of "Getting back so that the next person hiring the aircraft is able to take the aircraft on time" and "That your passengers would be impressed by your skills as a pilot." These items did not match the experimental situation, and hence, the pressures may not have been "real" enough for a valid response to be given. However, although group differences were not significant, the pilots rated personal pressures as more important than social pressures. Thus, the pilots' decisions were influenced more by factors that were related to tangible gains and losses for themselves, as well as self-approval and disapproval. Although these results are similar to findings by O'Hare and Smitheram (1995), these authors claimed that social and peer influences are still important in a pilot's decision to engage in risky flights. They based their claim on the observation of "numerous examples in the air crash files of low flying, 'beat-ups' and 'buzzing' that have led to disaster" that would not have occurred "without the presence or anticipated presence of an audience to observe the maneuvers" (p. 366). Clearly, more realistic simulated situations are needed to further clarify the relative impact of social and personal pressures on pilots' decisions.

More research also needs to be done to examine other personal and situational factors that contribute to a pilot's decision to either continue or divert a flight when encountering adverse weather. These include decision framing, the effects of flight experience, familiarity with the geographic region, and different magnitudes and rate of change in weather conditions. Such effects may combine to give a better picture of how pilot and environmental variables interact to precipitate VFR flight into IMC. Potentially, then, effective intervention strategies may be more readily forthcoming.

ACKNOWLEDGMENTS

This work was supported in part by a grant from the Federal Aviation Administration (DTFA–00–G–010). The contract technical monitor was Dr. David Hunter. The views expressed in this article are those of the authors and do not necessarily reflect those of the FAA. We acknowledge the invaluable contributions of Jonathan Sivier in the development of the simulation used in this study.

REFERENCES

Civil Aviation Authority. (1988). *General aviation accident review 1987* (CAP 542). Cheltenham, England: Civil Aviation Authority.

Goh, J., & Wiegmann, D. A. (2001). Visual flight rules (VFR) flight into instrument meteorological conditions (IMC): A review of the accident data. *Proceedings of 11th International Symposium on Aviation Psychology.* Columbus: Ohio State University.

Jensen, R. S. (1995). *Pilot judgment and crew resource management.* Brookfield, VT: Ashgate.

Jensen, R. S., & Benel, R. A. (1977). *Judgment evaluation and instruction in civil pilot training* (Final Rep. No. FAA–RD–78–24). Springfield, VA: National Technical Information Service.

Kahneman, D., & Tversky, A. (1982). Choices, values, and frames. *American Psychologist, 39,* 341–350.

Klein, G. A. (1997). The recognition-primed decision model (RPD): Looking back, looking forward. In C. E. Tsambok & G. A. Klein (Eds.), *Naturalistic decision making* (pp. 285–292). Mahwah, NJ: Lawrence Erlbaum Associates, Inc.

Laminar Research Incorporated. (1998). *X-plane.* Columbia, SC: Author.

O'Hare, D. (1990). Pilots' perception of risks and hazards in general aviation. *Aviation, Space, & Environmental Medicine, 61,* 599–603.

O'Hare, D. (1992). The "artful" decision maker: A framework model for aeronautical decision making. *The International Journal of Aviation Psychology, 2,* 175–191.

O'Hare, D., Batt, R., Wiggins, M., & Morrison, D. (1994). Cognitive failure analysis for aircraft accident investigation. *Ergonomics, 37,* 1855–1869.

O'Hare, D., & Owen, D. (2000). *Continued VFR into IMC: An empirical investigation of the possible causes* [Final report on preliminary study]. Unpublished manuscript, University of Otago, Dunedin, New Zealand.

O'Hare, D., & Smitheram, T. (1995). "Pressing on" into deteriorating conditions: An application of behavioral decision theory to pilot decision making. *The International Journal of Aviation Psychology, 5,* 351–370.

Rasmussen, J. (1982). Human errors: A taxonomy for describing human malfunction in industrial installations. *Journal of Occupational Accidents, 4,* 311–333.

Sanders, M. S., & McCormick, E. J. (1993). *Human factors in engineering and design* (6th ed.). New York: McGraw-Hill.

Wickens, C. D., & Hollands, J. G. (2000). *Engineering psychology and human performance* (3rd ed.). Upper Saddle River, NJ: Prentice Hall.

Wiegmann, D. A., & Shappell, S. A. (1997). Human factors analysis of postaccident data: Applying theoretical taxonomies of human error. *The International Journal of Aviation Psychology, 7,* 67–81.

Wright, P. (1974). The harassed decision-maker: Time pressures, distraction and the use of evidence. *Journal of Applied Psychology, 59,* 555–561.

Manuscript first received November 2000