



Cue-utilisation typologies and pilots' pre-flight and in-flight weather decision-making



Mark W. Wiggins^{a,*}, Danielle Azar^a, Jake Hawken^a, Thomas Loveday^a, David Newman^b

^aCentre for Elite Performance, Expertise and Training, Macquarie University, NSW 2109, Australia

^bDepartment of Aviation, Swinburne University, Australia

ARTICLE INFO

Article history:

Received 14 June 2013

Received in revised form 7 December 2013

Accepted 18 January 2014

Available online 11 February 2014

Keywords:

Weather-related decision-making

Cues

Situation assessment

Plan-continuation errors

ABSTRACT

In complex, high consequence environments such as aviation, the capacity to acquire, integrate, and respond to task-related cues is critical for accurate situation assessment and to avoid plan-continuation errors. The aim of the present study was to establish whether differences in performance on a series of aviation-related, cue-based tasks corresponded to differences in decision selection during simulated pre-flight and in-flight weather-related decision-making. In Phase 1 (pre-flight decisions), 57 participants were categorised into one of two typologies based on their performance on the cue-based tasks. These typologies reflected behaviour that was consistent with relatively greater or lesser levels of cue utilisation, and corresponded to whether the pilots elected to make an immediate decision or wait for additional information during a simulated pre-flight decision task. In Phase 2, a cohort of 20 pilots was selected on the basis that they represented one of the two cue-based typologies established in Phase 1. They undertook a simulated flight during which the weather conditions deteriorated progressively en-route. Those pilots who demonstrated a relatively greater level of cue utilisation were more likely to continue the flight as planned, while those pilots who demonstrated a relatively lesser level of cue utilisation were more likely to descend or divert from the planned track. The implications are discussed in terms of targeted training and explanations of plan-continuation errors in the context of weather-related decision-making.

© 2014 The Authors. Published by Elsevier Ltd. Open access under [CC BY-NC-ND license](https://creativecommons.org/licenses/by-nc-nd/4.0/).

1. Introduction

Inadvertent or deliberate visual flight into instrument meteorological conditions, and the resultant collision with terrain, continues to account for a disproportionate number of fatalities amongst general aviation pilots (Groff and Price, 2006; Hunter et al., 2011). This is due largely to the fact that pilots who are authorised to fly under Visual Flight Rules (VFR) lack the psychomotor and cognitive skills necessary to maintain control of the aircraft in the absence of visual reference to the horizon. Once visually-related pilots have lost visual reference to the horizon (such as occurs when flying in cloud), they can lose control of the aircraft within a few minutes (Bryan et al., 1955).

Like other decision-making tasks, weather-related decision-making involves the acquisition of information from a range of sources and a comparison between the options available, each of which carries a degree of uncertainty (Knecht, 2005). Prior to a

flight, this process involves the acquisition and interpretation of actual and forecast weather-related information, often from a number of different sources, including meteorological weather reports and aeronautical charts (Wiggins et al., 2002). During a flight, weather-related decision-making involves the assessment of the weather-related information available from the cockpit of the aircraft and its integration with the existing state of the aircraft. Both prior to, and during the flight, the information available provides the foundation for an assessment of the situation, which is the precursor to the selection of a particular option.

The significance of accurate and efficient situation assessment in decision-making under uncertainty is illustrated by Kaempff et al. (1996) in their analysis of tactical decision making in military operations. They noted that the accuracy of experienced United States Navy officers' responses was dependent upon a process of feature matching in which elements were compared to a prototype in memory. This enabled the rapid assessment of the system state (e.g. type of aircraft, speed, and altitude), together with an understanding of the significance of any changes that had occurred.

* Corresponding author. Tel.: +61 2 9850 9705; fax: +61 2 9850 8062.

E-mail address: mark.wiggins@mq.edu.au (M.W. Wiggins).

From a theoretical perspective, accurate and efficient situation assessment forms the foundation of the Recognition-Primed Model of decision-making (Klein, 1993). Incorporated within this model is the proposition that effective situation assessment involves the recognition and response to a familiar pattern of environmental features (Noble, 1993; Wiegmann et al., 2002). Indeed, Klein (1997) argues that this process is the basis of expertise, since it enables accurate and rapid responses, even in situations involving high cognitive load.

The Recognition-Primed Model also proposes that the recognition of patterns of environmental features derives from the availability of a repertoire of cues in long-term memory that can be triggered in response to specific stimuli (Klein, 1993; Salas et al., 2010; Wiggins and O'Hare, 2003). Cues are thought to represent a relationship between a feature and event or object that has been established through repeated association in the past (Shanteau, 1992; Wiggins, 2012). Following exposure to repeated pairings, the relationship between features and events/objects may become non-conscious, so that the response is both rapid and difficult for the operator to articulate (Zacks et al., 2007).

The utility of cues lies in their capacity to reduce the demand on working memory and enable a rapid and accurate interpretation of a scene. This increases the time and the cognitive resources available for subsequent decision-making (Fadde, 2009; Schriver et al., 2008). At the highest levels of performance, expert decision-makers rely on relatively fewer cues to form a diagnosis, having identified specific relationships that are optimally predictive of the changes that occur in the system state (Shanteau, 1992). For example, Schriver et al. (2008) have demonstrated that the capacity to create efficiencies in diagnosis is associated with shorter response latency amongst expert pilots in a dynamic flight simulation task, thereby enabling more appropriate and timely decisions.

There are a number of key elements that form the foundation of situation assessments, including the capacity to accurately identify task-related features from an array, the capability to differentiate relevant from less relevant feature-event/object associations, and the capacity to implement a structured process of information acquisition in response to a task-related problem (Wiggins, 2006, 2012).

In the context of a particular domain, the effective acquisition and utilisation of cues differs depending on the nature of individual and the domain-related experiences that have been acquired. Therefore, it is not necessarily possible to identify a single set of cues that are optimal for a particular context since different operators may use different cues to equal effect when resolving a problem (Patrick et al., 1999). What can be established is the extent to which an operator acquires and responds to information in a form that is characteristic of the effective use of cues.

The formation of feature-event/object relationships in the form of cues involves an iterative process whereby cues are modified or discarded as it becomes clear that there are more predictive and/or more efficient associations that might be available (Shah and Oppenheimer, 2008). This process of cue formation is a risky period during the process of skill acquisition, since it is during this period where mistakes are most likely to occur (O'Hare et al., 1994). Indeed, analyses of both automotive and aircraft accident statistics indicate that severe accidents are most likely to occur during the period immediately post-training, when operators begin honing their skills (Duncan et al., 1991; O'Hare et al., 1994).

Weather-related decision-making amongst pilots is an unusual context in which to examine the role of cue utilisation since the features associated with deteriorating weather conditions are dynamic, may present in different forms and, in the case of in-flight decisions, the speed of the aircraft often requires assessments within very short periods of time. There is also a strong motivational component associated with weather-related decision-making, and this is most evident during in-flight decision-making

where pilots can be subject to plan-continuation errors (Bearman et al., 2009).

Plan-continuation errors occur where operators continue to execute a planned behaviour, despite the presentation of information suggesting that an alternative response is warranted (Orasanu et al., 2001). The incidence of plan-continuation errors has been demonstrated experimentally amongst pilots who were confronted with deteriorating weather conditions during a simulated flight (Wiegmann et al., 2002). Where pilots had already completed a significant proportion of the flight, there was a tendency amongst some participants to continue the planned flight to the destination despite a deterioration in the weather conditions that rendered this option inadvisable (Wiegmann et al., 2002).

During the early stages of skill acquisition, learners tend to develop relatively imprecise associations between features and events or objects (Ellis, 1996; Klayman and Ha, 1989) (see Fig. 1). For learner pilots, the association between deteriorating weather and the safety and security of the aircraft is particularly salient so that even the mere presence of cloud may dissuade a visual pilot from undertaking a flight. However, through experience, a greater level of precision is acquired so that different types of weather conditions may be associated with different levels of risk to the aircraft and the likelihood of reaching the destination safely.

For visual pilots who are in the more advanced stages of cue development, the availability of a range of cues may result in a conflict. One of the most significant of these conflicts concerns the situation where the aircraft is in relatively close proximity to the destination but the weather conditions warrant a diversion to either an alternate destination or, in some cases, the original point of departure. It is in this type of situation that the plan-continuation error is most likely to occur since the proximity to the destination appears to be a particularly salient cue that may over-ride the cues associated with the deteriorating weather conditions.

To test this proposition in the present study, pilots were initially evaluated and classified into one of two typologies using the Expert Intensive Skills Evaluation (EXPERTise) Situational Judgement Test (SJT). The EXPERTise SJT classifies participants based on their composite scores across three tasks that are associated with cue utilisation:

1. The *Feature Identification Task*, whereby the participants must identify a key features from an array. The speed and accuracy with which participants are able to acquire that feature is indicative of the strength of their cue associations in memory (Ratcliff and McKoon, 1995).
2. The *Feature Association Task*, whereby participants must rate the association between feature-event/object pairs. The speed and variance of the participant's ratings is indicative of their capacity to distinguish related from unrelated features and events/objects (Morrison et al., 2013).

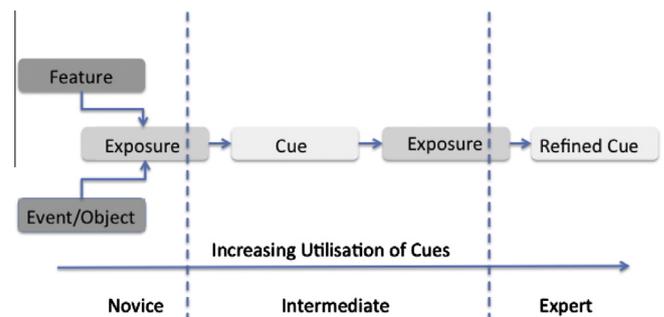


Fig. 1. A conceptual illustration of the formation, progression and refinement of cues and the correspondence with levels of expertise.

3. The *Transition Task*, whereby participants must acquire additional information about a scenario from a list of information categories. The sequence in which the participants acquire the information is indicative of their ability to prioritise key cues, rather than simply attending to features as they are presented (Wiggins and O'Hare, 1995).

The EXPERTise SJT has been designed to assess operators' utilisation of cues during task-related activities (Wiggins et al., 2010). The concurrent validity of the typologies generated by performance on EXPERTise has been demonstrated in the context of both power control (Loveday et al., 2013a) and paediatric diagnosis (Loveday et al., 2013b). Test–retest reliability has also been demonstrated at six-month administrations of the test (Loveday et al., 2013c).

Having been classified into typologies on the basis of their performance on EXPERTise, participants were asked to complete a pre-flight decision scenario in which they acquired information as they would in the operational environment to determine whether they would: (a) undertake a planned flight, (b) seek additional information prior to conducting the flight, or (c) not undertake the flight. A smaller cohort, representing participants from the two cue-based typologies was also asked to complete a simulated, in-flight decision scenario during which the weather conditions deteriorated progressively to a point 15 miles from the destination where the conditions were below the requirements for flight under Visual Flight Rules (VFR). It was hypothesised that participants who demonstrated a relatively lesser level of cue utilisation would select the option to seek additional information during the pre-flight decision scenario and be more likely to divert during the in-flight decision scenario. By contrast, it was hypothesised that pilots who demonstrated a relatively greater level of cue utilisation would be more likely to undertake the flight in response to the pre-flight decision scenario and would be more likely to continue the flight during the in-flight decision scenario.

2. Phase One: Pre-flight decision-making

2.1. Method

2.1.1. Participants

Fifty-seven pilots participated in Phase One of the study (55 male, 2 female). They ranged in age from 18 to 70 years, with a mean 41.46 years ($SD = 14.74$). Among the participants, 60% held a private pilot's licence, 18.2% held a commercial pilot's licence, and 21.8% held an Airline Transport Pilots Licence (ATPL). A summary of participants' flight experience is provided in Table 1. Across the cohort, three pilots self-reported having unintentionally flown into deteriorating weather conditions over the course of their flying careers. All three held a private pilot's license.

2.1.2. Instruments

Cue-based performance was assessed using a modified version of the EXPERTise SJT (Wiggins et al., 2010). It comprised a series of tasks that were designed to examine different aspects of the

utilisation of cues, including the feature identification task (two stages), the feature association task (two stages), and the transition task. The tasks were oriented around weather decision-making amongst pilots and the specific features were developed in consultation with a subject-matter expert.

The feature identification task involved two stages, both of which utilised specific meteorological weather reports (METAR) as stimuli. These METAR reports were sourced from an Australian Government website (www.bom.gov.au/aviation) over a five-week period. Using an international standard, METAR reports describe the weather conditions at a particular airport and consists of information such as the location, date, time, surface wind, visibility, and cloud.

For the purposes of the study, the reports were grouped into sets of five, and a subject matter expert assisted with the identification of the METAR in the group of five that displayed the poorest weather conditions relative to the others. In the first of the feature identification stages, the five METAR reports were displayed on the screen simultaneously and the participant was required to select the METAR with the poorest weather conditions (e.g. ceiling, visibility, cloud-base) as quickly as possible. Response latency was recorded for accurate selections across 16 trials. During the second feature identification stage, the METARs were presented for 1.5 s before the program progressed to the next screen. The participant was required to choose the METAR with the poorest weather conditions from a list of corresponding airport locations that were presented subsequently. The frequency of accurate responses was recorded for 16 trials.

The feature association task involved two stages, both of which displayed 17 pairs of randomised task-related feature and event/object terms to the participant (e.g. forecast – cloud-base, escape route – turbulence). The terms were derived from previous research involving weather-related decision making among pilots, where cues associated with deteriorating weather conditions were used to assist pilots to recognise and respond to deteriorating weather conditions (Wiggins and O'Hare, 2003). In the first stage of the feature association task, feature and event/object-related pairs were each shown separately for 1.5 s, after which the participant was asked to rate the 'strength' of the perceived relationship between the terms on a 10-point likert scale. The variance in ratings was calculated across the trials.

In the second stage of the feature association task, the feature and event-related terms were presented simultaneously. Consistent with stage one, the perceived strength of the relationship was rated and, across the trials, the variance in ratings was calculated.

The third task (Feature Transition Task) involved a simulated pre-flight, weather-related decision scenario in which the additional information needed to formulate a complete assessment of the scenario was accessible through a 'drop down' list of features. Twenty features were 'hidden' from the participants, and were visible only if the participant chose to access the additional information. For example, by 'clicking' on 'fuel', participants were given more information regarding the fuel available for the aircraft. The acquisition of this information, in addition to the order of information acquisition, was recorded using a process tracing strategy. This

Table 1
Phase One: Summary of the flying experience of pilots.

	PPL (60%) M (SD)	CPL (18.2%) M (SD)	ATPL (21.8%) M (SD)	Total (100%) M (SD)
Total experience in aviation (years)	15.97 (12.77)	24.56 (16.72)	18.25 (8.15)	18.31 (12.83)
Total flying experience (h)	3870.61 (6427.95)	4166.67 (5851.09)	6689.17 (5042.43)	4120.25 (5805.58)
Total flying experience as pilot in command (h)	1749.16 (3913.31)	3801.11 (5735.62)	2854.17 (2032.63)	2209.61 (3907.48)
Total flying experience in the previous 90 days (h)	39.98 (65.72)	61.22 (81.01)	109.58 (70.47)	56.27 (72.78)

enabled the calculation of the ratio of sequentially accessed pairs of information screens over the total pairs of information screens available (cf. Wiggins and O'Hare, 1995).

2.1.3. Stimuli

The stimulus for Phase One of the study comprised a pre-flight, weather-related scenario that incorporated a flight plan, an area forecast, and a World Aeronautical Chart (WAC). The task required participants to assess the information available and determine which of three actions they would be most likely to undertake (e.g. conduct the flight as planned, wait and acquire additional information, or not conduct the flight). On the subsequent screen, they were asked to rate, on a 10-point scale, the utility of each of nine features (see Table 2) when deciding on the action taken. The responses to the scenario, along with the feature utility ratings, were recorded.

2.1.4. Procedure

The participants were invited to take part in the study using an email that included instructions for participation. The instructions directed the participant to the domain address (expertise.cadre.com.au) and provided them with login details to access the study. Having completed the demographics questionnaire, the EXPERTise program initiated. Full instructions were provided throughout the program, and a debrief sheet was available for download following the completion of the experimental tasks. The study required approximately 25 min to complete.

2.2. Results

2.2.1. Cluster analysis

The initial aim of the present study was to establish whether typologies could be differentiated on the basis of pilots' performance across the tasks that comprised EXPERTise. To demonstrate that there were distinct levels of cue utilisation performance and that these levels of performance were consistent across the tasks, the K-means cluster procedure within the SPSS statistical package was employed. As required by the K-means cluster procedure, the cue utilisation measures were converted to standardized z-scores prior to clustering. z-Scores reflect how many standard deviations an observation is above or below the mean score for each measure and, consequently, are able to account for differences in scale between the measures (Tabachnick and Fidell, 1996). K = 2 clusters was entered as the most likely fit for the data, based on sample size and the outcomes of previous investigations (Loveday et al., 2013b).

Table 3 summarises the results of the cluster analysis, including the mean centroid for each cluster on each of the variables that comprise EXPERTise. Performance on the EXPERTise tasks resulted in the formation of two typologies that would form the basis of the subsequent analyses. Cluster 1 comprised 28 participants who recorded relatively high response latencies in response to the first of the feature identification tasks and the second paired association task, greater accuracy in response to the second feature identifica-

Table 2

The nine features that were rated as part of the pre-flight decision scenario.

Fuel limitations
Forecast
Cloud-tops
Terrain
Cloud-base
Visibility
Escape routes
Turbulence
Wind direction

Table 3

Participant cluster means for the measures that comprise EXPERTise, distributed across the two cue utilisation typologies.

	Greater	Lesser
Feature identification task (Stage 1) – response latency	.46	–.48
Feature identification task (Stage 2) – accuracy	.64	–.59
Feature association task (Stage 1) – response latency	.62	–.60
Feature association task (Stage 1) – ratings variance	.10	–.28
Feature association task (Stage 2) – ratings variance	.41	–.57
Feature transition task – ratio of sequential pairs	–.34	.30

tion task, greater variance in response to the two feature association tasks, and a relatively lower ratio in response to the transition task. This pattern of responses is broadly consistent with a greater reliance on the utilisation of context-related cues to interpret and respond to the information available. The remaining 29 participants comprised the second typology, the performance of which was less consistent with the utilisation of cues.

2.2.2. Typologies and pre-flight decision response

In comparing the two performance typologies (greater or lesser levels of cue utilisation), it was necessary to establish whether an association existed between cluster assignment and outcome performance on the simulated pre-flight decision task. A Chi-square test for independence revealed a significant association between cluster membership and response outcomes, $\chi^2(2,57) = 7.62, p = .022$. Specifically, those pilots whose performance was more consistent with greater levels of cue utilisation recorded relatively dichotomous responses either to conduct or to not conduct the flight (see Table 4). Pilots whose performance was more consistent with lesser levels of cue utilisation tended to select the option to acquire additional information prior to the flight.

To determine whether the typologies were differentiated on the basis of the total hours of flight experience or hours of recent flight experience, a multivariate analysis of variance was conducted with typology as the independent variable and total hours of flight experience and hours of recent flight experience as the dependent variables. The analysis revealed that the typologies identified were not differentiated on the basis of the total hours of flight experience accumulated by pilots, $F(1,57) = 1.133, p = .292$, nor by their hours of recent flight experience, $F(1,53) = .000, p = .998$.

A Chi-square test for independence was conducted subsequently to determine whether there was a relationship between typology and the level of the pilot's license. No relationship was evident between typologies and the licence held by pilots, $\chi^2(3,59) = 2.79, p = .43$. The overall results of Phase One suggested that the performance on EXPERTise was a product of factors other than experience and level of qualification.

3. Phase Two: In-flight decision-making

3.1. Method

3.1.1. Participants

The participants in Phase Two were 20 visually-rated, male, general aviation pilots. These pilots were those from Phase One

Table 4

Phase One: Cross-tabulation of cue utilisation typology against the decision to the conduct the flight during the pre-flight decision task.

	Conduct the flight	Acquire additional information	Do not conduct the flight
Greater cue utilisation	13	4	11
Lesser cue utilisation	8	14	7

who agreed to participate in the follow-up study. Nine pilots demonstrated performance on EXPERTise that was consistent with greater levels of cue utilisation, while the remaining 11 pilots demonstrated performance that was consistent with relatively lesser levels of cue utilisation. An multivariate analysis of variance between the groups revealed that there were no significant differences between them in relation to age, $F(1, 19) = .08, p = .78$, years of experience in aviation, $F(1, 17) = .24, p = .63$, the total number of flight hours accumulated, $U = 36, N_1 = 9, N_2 = 11, p = .31$, the number of flight hours accumulated as pilot in command, $U = 30.5, N_1 = 9, N_2 = 11, p = .15$, or the number of flight hours accumulated in the 90 days preceding the study, $U = 30, N_1 = 9, N_2 = 11, p = .83$.

The majority of participants in both typologies held private pilots licences (67% for performance consistent with higher cue utilisation and 72% for performance consistent with lower cue utilisation), while the remaining pilots held commercial or an airline transport pilot's licence. Overall, 15% of the pilots, $n = 3$, held an instrument rating, two of whom were classified in the lesser cue utilisation typology. The remaining IMC-rated pilot was classified in the greater cue utilisation typology.

3.1.2. Flight simulation

The participants each completed a simulated cross-country flight using Redbird FMX flight simulators located at Macquarie University and at Swinburne University. The Redbird flight simulator uses an enhanced version of Microsoft Flight Simulator as a software platform, has 180° wrap-around visibility, responds with three degrees of freedom, and is capable of being registered as an Advanced Aviation Training Device by the Federal Aviation Administration. The program 'Insight' was used to record the longitude, the latitude, the altitude, and the heading of the aircraft for each half-second of the flight. The simulated flight was planned from Bathurst, New South Wales, to Cowra via Orange, New South Wales, a distance of 63 nautical miles.

The weather conditions deteriorated gradually during the flight at pre-determined distances from the point of departure. The variables that changed during the flight included the cloud base, visibility, breakout (change in visibility), and cloud coverage. The changes in the features of the weather conditions, together with the distance from the point of departure at which the changes occurred, are provided in Table 5.

The dependent variable during the simulation was whether pilots continued the flight without a change in altitude or heading as the conditions deteriorated, or whether they initiated a descent or a change in heading, the latter of which might incorporate a diversion to an alternate destination. The nature of the changes in weather conditions, as well as rate of deterioration, was intended to make the decision as to whether to continue or deviate from the flight plan as subjective as possible.

3.1.3. Risk perception

Risk perception was considered a potential covariate in the analysis, since there is evidence to suggest that differences in pilots' decision-making reflects differences in their perceptions of the risk associated with particular activities (Hunter, 2002). The risk perception questionnaire employed during the study was developed by Hunter (2002) and consists of a series of 26 scenarios that describe higher and lower risk aviation activities, in-flight weather conditions, and non-aviation activities. The participants were asked to rate on a scale of 1–100, each of the 26 scenarios in terms of the level of risk that they perceived to be involved in each situation. A value of 'one' corresponded to low risk and a value of 100 corresponded to high risk, with 50 as the midpoint. An exemplar question from the risk perception questionnaire is 'Make a two-hour cross-country flight with friends, after checking

Table 5

Deterioration of weather conditions during the simulated flight.

Distance into flight (nm)	Cloud base (feet)	Visibility (nm)	Breakout	Cloud coverage
0	7500	Unlimited	7387	4/8
10	7500	40	7500	4/8
15	7500	30	7500	5/8
25	7500	20	7500	7/8
30	7100	10	7500	7/8

your weight and balance'. In the present study, the reliability of the instrument across the four dimensions was 0.817.

3.1.4. Procedure

Having completed EXPERTise in Phase One, the participants were advised that they would be undertaking a simulated flight under Visual Flight Rules (VFR). They were also advised that they should undertake the flight as they would within the operational environment. Participants were supplied with a flight plan, current weather reports, including an Area Forecast, Terminal Area Forecast, Meteorological Reports, and a World Aeronautical Chart (WAC) with the intended flight marked. They were also provided with a kneeboard and a pen, and were permitted 15 min to familiarise themselves with the route prior to the flight.

The simulated flight was initiated with the aircraft ready for take-off at Bathurst on Runway 35. The participants were required to depart Bathurst and climb to a cruising altitude of 6500 feet on a heading of 267° to Orange. On reaching Orange, they turned the aircraft to a heading of 220° towards Cowra. The simulation was terminated when the participants had landed the aircraft at either Orange or Cowra, or if the participants appeared to have lost control of the aircraft.

3.2. Results

The primary aim of this phase of the study was to determine whether a relationship existed between different cue utilisation typologies and the decisions that pilots made in response to deteriorating weather conditions during a simulated flight. Therefore, the variable of most interest in the present study was whether pilots continued the flight without a change in altitude or heading as the conditions deteriorated, or whether they initiated a descent or a change in heading, the latter of which might incorporate diversions to an alternate destination. Overall, 50% pilots elected to continue the flight, despite the fact that the visibility was below that necessary for operations under Visual Flight Rules (VFR). A Chi-square analysis established a significant relationship between cue utilisation typology and the decision to continue the flight or initiate a response, either to descend or initiate a diversion, $\chi^2(1, 20) = 5.05, p = .025$. Inspection of the contingency table indicated that 77% of pilots (7) whose performance was consistent with higher levels of cue utilisation continued the flight. By comparison, only 27% of pilots (3) whose performance was consistent with lower levels of cue utilisation elected to continue the flight.

3.2.1. Risk perception and measures of experience

A one-way ANOVA was used to determine whether the cue utilisation typologies could be differentiated on the basis of the perceptions of risk associated with aviation activities in particular, and with life events more generally. The Risk Perception Questionnaire developed by Hunter (2002) comprises four dimensions, although the level of reliability between the scales provided some justification for the calculation of a composite score. The ANOVA failed to reveal a statistically significance difference between cue utilisation typologies and composite scores on the Risk Perception

Questionnaire, $F(1, 18) = .00, p = .97$. Nor was there any significant difference between pilots' risk perception scores on the basis of whether they continued or diverted from the planned flight, $F(1, 18) = .10, p = .75$.

Consistent with the results pertaining to risk perception, no differences were evident between pilots who diverted from the planned route and those who continued in terms of the total number of flight hours accumulated, $F(1, 19) = 2.48, p = .13$, the number of hours accumulated as pilot-in-command, $F(1, 19) = 1.64, p = .22$, and the number of flight hours accumulated in the 90 days preceding testing, $F(1, 15) = .00, p = .94$. Only three of the participants in Phase Two held instrument ratings, one of whom diverted, while the remaining two pilots continued the flight.

4. General discussion

This study was designed to determine whether cue utilisation typologies differentiated the behaviour of pilots in response to simulated pre-flight and in-flight decision-making scenarios. Pilots were initially classified on the basis of their level of cue utilisation, having completed a weather-related version of the EXPERTise situational judgement test. Two typologies emerged that broadly represented relatively greater or lesser levels of cue utilisation. Consistent with the hypothesis, a relationship was evident between cue utilisation typologies and pilot's pre-flight decision-making in which pilots who demonstrated a relatively greater level of cue utilisation were more dichotomous in their responses, electing either to conduct or not conduct the flight. By contrast, pilots who demonstrated a relatively lesser level of cue utilisation were more likely to select the more ambivalent option, presumably on the basis that there was a perception of insufficient information available to form a decision.

In the case of the in-flight decision scenario, pilots who demonstrated lesser levels of cue utilisation were more likely to divert or descend compared to pilots who demonstrated greater levels of cue utilisation, who tended to continue the flight into deteriorating weather conditions. This occurred despite the fact that the typologies were not significantly different in terms of risk perception, total hours of flight experience, hours of recent flight experience, whether or not they held an instrument rating, or the licence that they held. In combination, the responses to the pre-flight and in-flight scenarios suggest that the acquisition and subsequent application of cues may play a significant role in weather-related decision-making amongst pilots. Importantly, it may, in part, explain the plan-continuation error prevalent amongst visual pilots who faced with deteriorating weather conditions during flight.

According to Schriver et al. (2008), where pilots possess a more sophisticated repertoire of cues in memory, there is an increase in the number of feature-event/object relationships to which an operator may attend. It is through experience that these relationships are then refined and that operators establish those cues that are inefficient and/or inaccurate (Janelle et al., 2003). In the context of the present study, 77% of pilots who demonstrated relatively greater levels of cue utilisation elected to continue the flight during the in-flight scenario, despite the deteriorating weather conditions. Therefore, it is possible that these pilots have progressed beyond the stage of cue acquisition where they are responsive to 'high-level' features such as the mere presence or absence of cloud, and are increasingly responsive to finely tuned features such as visibility, the progression of the deterioration in weather conditions, and/or the proximity to the destination. Consequently, what presents as a plan-continuation error might be explained as a response to a features, albeit inaccurate, that the rate of deterioration and the proximity of the destination are such that it supplants other features that might be associated the safety of the aircraft.

Reacting to these finely tuned, but potentially inaccurate features might also explain the dichotomous responses that were evident in the pre-flight decision scenario, since pilots who demonstrated greater levels of cue utilisation would be responding to different features, depending upon their individual experiences. Since there were no high-level features to which to attend, pilots with relatively lower levels of cue utilisation sought a response that would enable them to acquire additional information and, perhaps, identify the higher level feature/s that would indicate the appropriate response under the circumstances.

4.1. Theoretical implications

The inappropriate or ineffective application of cues is one of a number of explanations for cases of visual flight into instrument meteorological conditions. Other explanations include the motivation to reach the destination, and/or a failure to identify and appropriately assess the risks associated with continued flight (Wiegmann et al., 2002). The results of the present study are broadly consistent with these explanations insofar as they identify a particular cohort of pilots for whom pre-flight and in-flight decision-making is likely to be associated with poor outcomes.

As practitioners develop their experience within a domain, they begin by developing and refining strategies that are intended to improve the accuracy and the efficiency of their performance within the operational environment (Wiggins, 2012). Inevitably, this process of hypothesis testing or effortful learning results in some associations that are inefficient or, in some cases, inaccurate (Roediger and Butler, 2011; Rolison et al., 2011). It is only through further exposure to the domain that the inadequacies of these associations can be established and that refinements can be initiated to redress the issues. However, in the context of high-consequence environments such as weather-related decision-making, the consequence of this type of error-based learning can be significant.

The outcomes of the present research suggest that, where plan continuation errors are evident in transitions into deteriorating weather conditions, these errors may be a consequence of inadequate or inappropriate feature-event/object associations. The relative effectiveness of these associations might also explain inaccurate assessments of risk, together with the apparent desire to reach the destination. In effect, these pilots may be unable to detect the nuances that would otherwise lead them to discontinue the flight under the circumstances.

Support for the role of cue associations in weather-related judgements can be drawn from the outcomes of training strategies that are designed to enable the accurate interpretation of cues. Using independent assessments of the cue-based weather-related training program *Weatherwise*, both Wiggins and O'Hare (2003) and Chansik (2011) were able to demonstrate improvements in pilot performance in response to simulated deteriorating weather conditions. The *Weatherwise* program is designed to expose pilots to key cues that are indicative of deteriorating weather conditions and to which they may not have been exposed during initial training.

4.2. Impact on industry and future research

The outcomes of the present study offer, for the first time, the basis of a psychometric instrument that might enable the identification of those pilots who are at a stage of skill acquisition at which they are most at risk of inaccurate weather-related decisions. It should also be possible to adopt a similar approach across other forms of situation assessment under uncertainty, including the interpretation of weather radar displays and the analysis of the decision height for instrument approaches to landing.

Targeting pilot groups at particular levels of risk will reduce the costs associated with training interventions and will also enable a greater uptake of cue-based information, since the effectiveness of cue-based training appears to be restricted to operators who have reached a level of competence and who have developed a reasonably detailed mental model of the domain (Perry et al., 2012). The role of cue-based training in this case is to capitalise on the feature-event/object relationships already resident in memory and to 'tune' the associations, thereby enabling finer levels of discrimination than might have previously been the case (Anderson, 1982). The result may be an increase in the rate of skill acquisition, since the unstructured process of hypothesis testing is avoided.

4.3. Conclusion

This study sought to examine the relationship between typologies that describe different levels of cue utilisation amongst pilots and their performance in response to simulated pre-flight and in-flight weather-related decision tasks. The results indicated that those pilots whose performance was consistent with relatively greater levels of cue utilisation were more definitive in their decision to either conduct or not to conduct the flight during the pre-flight decision scenario, and were more likely to continue the flight during the in-flight decision scenario. This suggests that, although pilots of this typology may have developed and are applying cues, these cues may be incomplete, inefficient or, in some cases, inaccurate. This may explain the plan-continuation errors that are evident amongst some pilots, but it also suggests that these pilots can be identified so that appropriate training can be initiated.

Acknowledgement

The authors would like to thank Dr. David Hunter for his comments and suggestions on earlier drafts of this manuscript.

References

- Anderson, J.R., 1982. Acquisition of cognitive skill. *Psychol. Rev.* 89, 369–406. <http://dx.doi.org/10.1037/0033-295X.89.4.369>.
- Bearman, C., Paletz, S.B., Orasanu, J., 2009. Situational pressures on aviation decision-making: goal seduction and situation aversion. *Aviat. Space Environ. Med.* 80 (6), 556–560. <http://dx.doi.org/10.3357/ASEM.2363.2009>.
- Bryan, L.A., Stonecipher, J.W., Aron, K., 1955. AOPA's 180° Rating. Aircraft Owners and Pilots Association, Washington, DC.
- Chansik, K., 2011. The effects of weather recognition training on general aviation pilot situation assessment and tactical decision making when confronted with adverse weather conditions. All Dissertations. Paper 754.
- Duncan, J., Williams, P., Brown, I., 1991. Components of driving skill: experience does not mean expertise. *Ergonomics* 34, 919–937.
- Ellis, J., 1996. Retrieval cue specificity and the realization of delayed intentions. *Quart. J. Exp. Psychol.: Sect. A* 49 (4), 862–887.
- Fadde, P.J., 2009. Expertise-based training: getting more learners over the bar in less time. *Technol. Instruct., Cogn. Learn.* 7 (2), 171–197.
- Groff, L.S., Price, J.M., 2006. General aviation accidents in degraded visibility: a case control study of 72 accidents. *Aviat. Space Environ. Med.* 77, 1062–1067.
- Hunter, D.R., 2002. Risk Perception and Risk Tolerance in Aircraft Pilots. Federal Aviation Administration (NTIS DOT/FAA/AM-02/17), Washington, DC.
- Hunter, D.R., Martinussen, M., Wiggins, M., O'Hare, D., 2011. Situational and personal characteristics associated with adverse weather encounters by pilots. *Accid. Anal. Prev.* 43, 176–186. <http://dx.doi.org/10.1016/j.aap.2010.08.007>.
- Janelle, C.M., Champenoy, J.D., Coombes, S.A., Mousseau, M.B., 2003. Mechanisms of attentional cueing during observational learning to facilitate motor skill acquisition. *J. Sports Sci.* 21, 825–838.
- Kaempf, G.L., Klein, G.A., Thorsden, M.L., Wolf, S., 1996. Decision making in complex naval command-and-control environments. *Hum. Factors: J. Hum. Factors Ergon. Soc.* 38 (2), 220–231. <http://dx.doi.org/10.1177/001872089606380204>.
- Klayman, J., Ha, Y.-W., 1989. Hypothesis testing in rule discovery: strategy, structure, and content. *J. Exp. Psychol. Learn. Mem. Cogn.* 15, 596–604.
- Klein, G.A., 1993. A recognition-primed decision (RPD) model of rapid decision making. In: Klein, G.A., Orasanu, J., Calderwood, R., Zsombok, C.E. (Eds.), *Decision Making in Action: Models and Methods*. Ablex, Norwood, NJ, pp. 138–147.
- Klein, G.A., 1997. Developing expertise in decision making. *Think. Reason.* 3 (4), 337–352. <http://dx.doi.org/10.1080/135467897394329>.
- Knecht, W., 2005. Pilot Willingness to Take-Off into Marginal Weather, Part II: Antecedent Overfitting with Forward Stepwise Logistic Regression. Federal Aviation Administration (NTIS DOT/FAA/AM-05/15), Washington, DC.
- Loveday, T., Wiggins, M.W., Harris, J., Smith, N., O'Hare, D., 2013a. An objective approach to identifying diagnostic expertise amongst power system controllers. *Hum. Factors* 55, 90–107. <http://dx.doi.org/10.1177/0018720812450911>.
- Loveday, T., Wiggins, M.W., Searle, B.J., Festa, M., Schell, D., 2013b. The capability of static and dynamic features to distinguish competent from genuinely expert practitioners in pediatric diagnosis. *Hum. Factors* 55, 125–137. <http://dx.doi.org/10.1177/0018720812448475>.
- Loveday, T., Wiggins, M., Festa, M., Schell, D., Twigg, D., 2013c. Pattern recognition as an indicator of diagnostic expertise. In: Latorre Carmona, P., Sanchez, J.S., Fred, A.L.N. (Eds.), *Pattern Recognition – Applications and Methods*, vol. 204. Springer, Berlin, pp. 1–11. http://dx.doi.org/10.1007/978-3-642-36530-0_1.
- Morrison, B.W., Wiggins, M.W., Bond, N.W., Tyler, M.D., 2013. Measuring relative cue strength as a means of validating an inventory of expert offender profiling cues. *J. Cogn. Eng. Decis. Making* 7 (2), 211–226.
- Noble, D., 1993. A model to support development of situation assessment aids. In: Klein, G.A., Orasanu, J., Calderwood, R., Zsombok, C.E. (Eds.), *Decision Making in Action: Models and Methods*. Ablex, Norwood, NJ, pp. 287–305.
- Orasanu, J., Martin, L., Davison, J., 2001. Cognitive and contextual factors in aviation accidents: decision errors. In: Salas, E.E., Klein, G.E. (Eds.), *Linking Expertise and Naturalistic Decision Making*. Lawrence Erlbaum Associates, Mahwah, NJ, pp. 209–225.
- O'Hare, D., Wiggins, M., Batt, R., Morrison, D., 1994. Cognitive failure analysis for aircraft accident investigation. *Ergonomics* 37, 1855–1869. <http://dx.doi.org/10.1080/00140139408964954>.
- Patrick, J., Gregor, A., Halliday, P., Handley, J., O'Reilly, S., 1999. Analysing operators' diagnostic reasoning during multiple events. *Ergonomics* 42, 493–515. <http://dx.doi.org/10.1080/001401399185603>.
- Perry, N., Wiggins, M.W., Childs, M., Fogarty, G., 2012. Can reduced processing decision support interfaces improve the decision-making of less-experienced incident commanders? *Decis. Support Syst.* 52, 497–504.
- Ratcliff, R., McKoon, G., 1995. Sequential effects in lexical decision: tests of compound cue retrieval theory. *J. Exp. Psychol. Learn. Mem. Cogn.* 21, 1380–1388.
- Roediger III, H.L., Butler, A.C., 2011. The critical role of retrieval practice in long-term retention. *Trends Cogn. Sci.* 15, 20–27. <http://dx.doi.org/10.1016/j.tics.2010.09.003>.
- Rolison, J.J., Evans, J.S.B., Walsh, C.R., Dennis, I., 2011. The role of working memory capacity in multiple-cue probability learning. *Quart. J. Exp. Psychol.* 64, 1494–1514. <http://dx.doi.org/10.1080/17470218.2011.559586>.
- Salas, E., Rosen, M.A., DiazGranados, D., 2010. Expertise-based intuition and decision making in organisations. *J. Manage.* 36, 941–973. <http://dx.doi.org/10.1177/0149206309350084>.
- Schrifer, A.T., Morrow, D.G., Wickens, C.D., Talleur, D.A., 2008. Expertise differences in attentional strategies related to pilot decision making. *Hum. Factors: J. Hum. Factors Ergon. Soc.* 50, 864–878. <http://dx.doi.org/10.1518/001872008X374974>.
- Shah, A.K., Oppenheimer, D.M., 2008. Heuristics made easy: an effort-reduction framework. *Psychol. Bull.* 134, 207–222. <http://dx.doi.org/10.1037/0033-2909.134.2.207>.
- Shanteau, J., 1992. How much information does an expert use? Is it relevant? *Acta Psychol.* 81, 75–86. [http://dx.doi.org/10.1016/0001-6918\(92\)90012-3](http://dx.doi.org/10.1016/0001-6918(92)90012-3).
- Tabachnick, B.G., Fidell, L.S., 1996. *Using multivariate statistics*, 3rd Edition. Harper Collins College Publishers, New York.
- Wiegmann, D.A., Goh, J., O'Hare, D., 2002. The role of situation assessment and flight experience in pilots' decisions to continue visual flight rules flight into adverse weather. *Hum. Factors* 44, 189–197. <http://dx.doi.org/10.1518/0018720024497871>.
- Wiggins, M.W., 2006. Cue-based processing and human performance. In: Karwowski, W. (Ed.), *Encyclopedia of Ergonomics and Human Factors*, second ed. Taylor and Francis, London, UK, pp. 641–645.
- Wiggins, M.W., 2012. The role of cue utilisation and adaptive interface design in the management of skilled performance in operations control. *Theor. Issues Ergon. Sci.* 1–10. <http://dx.doi.org/10.1080/1463922X.2012.724725>.
- Wiggins, M.W., Harris, J., Loveday, T., O'Hare, D., 2010. EXPERTise (Software Package). Macquarie University, Sydney, Australia.
- Wiggins, M., O'Hare, D., 1995. Expertise in aeronautical weather-related decision making: a cross-sectional analysis of general aviation pilots. *J. Exp. Psychol.: Appl.* 1 (305–320).
- Wiggins, M.W., O'Hare, D., 2003. Expert and novice pilot perception of static in-flight images of weather. *Int. J. Aviat. Psychol.* 13, 173–187. http://dx.doi.org/10.1207/S15327108IJP1302_05.
- Wiggins, M., Stevens, C., Howard, A., Henley, I., O'Hare, D., 2002. Expert, intermediate and novice performance during simulated pre-flight decision-making. *Aust. J. Psychol.* 54, 162–167. <http://dx.doi.org/10.1080/00049530412331312744>.
- Zacks, J.M., Speer, N.K., Swallow, K.M., Braver, T.S., Reynolds, J.R., 2007. Event perception: a mind-brain perspective. *Psychol. Bull.* 133, 273–293. <http://dx.doi.org/10.1037/0033-2909.133.2.273>.