Examining cue recognition across expertise using a computer-based task

Ben W. Morrison

b.morrison@uws.edu.au

Mark W. Wiggins

mark.wiggins@psy.mq.edu.au

Nigel W. Bond

n.bond@uws.edu.au

Michael D. Tyler

m.tyler@uws.edu.au

The School of Psychology and MARCS Laboratories
University of Western Sydney

ABSTRACT

Motivation – The study examined whether experts and novices differed in their recognition of decision-making cues. **Research approach** – To test cue recognition, the authors developed and tested a computer-based cue recognition task on a group of expert and novice offender profilers. **Findings/Design** – Recognition performance was assessed in relation to cue classification agreement and recognition response latency among and between the two groups. The findings revealed superior performance on both measures by the experts compared to the novices. **Research limitations/Implications** – The findings have implications for the cue selection process in the design of computer-based training, and decision support systems. **Originality/Value** – The research offers an objective means of: 1) identifying cues; 2) gauging relative cue stability/strength; 3) comparing cue recognition across expertise; and, 4) selecting a valid cue-set for use in training and support systems. **Take away message** – There are significant differences in cue recognition across expertise that may, in part, differentiate decision-making performance.

Keywords

Decision-making, Cues, Cue Recognition, Expertise, Cue-Based Training.

INTRODUCTION

In many workplace environments, performance is heavily reliant upon the cognitive skills of the operator, which often includes a capacity to acquire and interpret information and form a view as to the nature of the events that transpired. Differences in this form of decision performance are invariably correlated with a difference in operational experience, whereby highly experienced or *expert* decision-makers tend to demonstrate a relatively superior level of decision-making skill which, in many workplaces, is critical (e.g., medical diagnosis, vehicle collision investigation, fire-fighting, violent crime investigation, etc.).

Many organisations within these domains do not possess the time, nor the resources, to enable operators to progress from novice to expert through day-to-day operational experience (see "Junior Police Recruited to Elite Command Post", 2003, June 5). Therefore, less experienced operators are being placed in positions where a high degree of expertise is required, leading to the potential for an increase in human error. As a result, much research has been aimed at identifying the cognitive factors which underlie the difference in decision-making performance across expertise, in an attempt to reduce the novice-expert divide.

There is an increasing body of evidence to suggest that expert and novice performance is distinguished by the capacity to acquire, identify and utilise cues as a basis for diagnosis and response (Klein, 1997). Evidence to support this observation can be drawn from a range of domains in which experts demonstrate a capacity to rapidly recognise and respond to changes in the state of the system (Wiggins & O'Hare, 2003). Invariably, cognitive interviews involving subject-matter experts reveal that their rapid responses are associated with the examination of relatively few features, each of which hold some form of association or meaning for the expert (Klein, Calderwood, & MacGregor, 1989; O'Hare, Wiggins, Williams, & Wong, 2000).

Effectively, cues are presumed to represent relationships, held in memory, between environmental or situational features, and events (Wiggins, 2006). These associations are typically acquired through experience, and provide a basis to selectively attend to those stimuli that are most likely to hold value and facilitate the efficient and accurate interpretation of the situation. In this sense, cues represent an effective attention management strategy that obviates the requirement for a time-consuming, deliberative process of information acquisition and management (Einhorn & Hogarth, 1981).

Feature-event relationships, in the form of cues, are closely associated with the application of heuristics in judgement and decision-making research. Where a cue specifies the co-occurrence of observed phenomena, heuristics represent the process through which these observed phenomena are integrated to form meaning and, thereby predict future events (Gigerenzer & Todd, 1999). For example, in the case of the representativeness heuristic, a particular stimulus may trigger a series of cues in memory. However, it is the frequency and types of triggered cues that lead to the categorisation of the stimulus through the activation of a mental model. The mental model allows predictions of future events based on the pattern and types of cues that are engaged (Griffin, Kahneman, Aspinwall, & Staudinger, 2003). Therefore, where cue activation might be regarded as a *bottom-up* process of activation, the application of a heuristic might be regarded as a *top-down* process.

By definition, highly experienced, expert decision-makers within a domain have observed a vast number of previous decision cases. As a result, they are likely to possess a comprehensive store of cues in memory that enables the recognition and interpretation of information within the operational environment (Wiggins & O'Hare, 1995). In comparison, less experienced practitioners have had less exposure to the operational environment and, therefore, will have acquired a relatively limited range of cues. Arguably, it is this distinction between the frequency and the specificity of cues that contributes to the less rapid, more deliberate and linear process that is typically associated with novice decision-making (Larkin, McDermott, Simon, & Simon, 1980). If such differences in cue recognition exist across expertise, they may, to some extent, differentiate expert from novice decision performance.

The proposed relationship between decision-makers' experience and their capacity for cue recognition is consistent with Anderson's (1993) *principle of strength* in relation to cognitive productions in memory. In this case, the repeated exposure and successful application of a production increases the likelihood that it will be used in the future. The strength of a production, that is, its capacity to be recognised or retrieved from memory, is determined by the number of successful applications; its level of *activation*. Like productions, it might be argued that both the recollection and the strengthening of cues are based on an individual's capacity to activate the relevant pathways to the information stored in memory (Eveleth, 1999).

Consistent with this perspective, Ackerman and Rathburn (1984) investigated the effect of episodic experience on the acquisition and retrieval of cue-based information. When presenting participants with word pairings, the experimenters varied the participants' recognition experience by presenting word pairings that were categorically related to each other and other pairs that were unrelated. They demonstrated that the recall of related words was not only faster, but that recall could be further strengthened as a result of additional exposure to the target stimuli. Similarly, the identification of potential differences in the strength of feature-event relationships may provide an indication of the extent to which cues are more or less established in an operator's long-term memory. Consistent with this evidence, we propose that if a relationship exists between a cue's strength, or *stability* in memory, and the speed at which recognition will occur, a decision-maker's capacity for cue recognition could potentially be assessed by observing response latency when responding to valid cue-based associations.

Although the isolation of specific cues on the basis of cue recognition performance is a relatively novel concept in experimental psychology, the principles on which a cue recognition test can be developed are well established. For example, the Lexical Decision Task (LDT) is a procedure that involves measuring the response latency of participants in their ability to recognise and classify stimuli (Meyer & Schvaneveldt, 1971). Participants usually make a forced choice as to whether an item is a member of a group or not, for example, word vs. non-word. This task has often been used to examine participants' capacity to recognise associations between related stimuli so that the response latency for the recognition of a word or non-word can be influenced by the use of primes which possess a high degree of relatedness. The current experiment was designed to adopt this method to test participants' response latency to a series of paired, feature-and-event-related concepts. To identify those concepts which possess a high degree of relatedness, a task similar to the LDT was designed and is referred to as the *Feature Event Paired Association Task* (FEPAT).

The FEPAT

The FEPAT is a forced-choice, computer-based presentation task which, theoretically, can be used to: 1) identify a range of associations (cues) between feature- and event-related concepts; 2) gauge the strength of participants' recognition of cues based on their response latency; 3) identify cues that possess a high degree of relatedness across a

particular sample (e.g., experts); and 4) monitor and assess the rate at which cue acquisition occurs over a period of time during a training-based intervention.

The task involves presenting text-based pairings of feature- and event-related information to participants on a computer screen, and asking them to decide whether the concepts presented share an association in a specified context/domain. Participants are briefly presented an event description at the centre of the screen. This screen is then replaced with the presentation of the same event description on the left side of screen coincident with a feature description presented on the right of screen. By limiting exposure, over-processing can be limited to promote recognition memory, and to discourage a deliberate, analytical process. During each presentation, participants are to respond either *Yes* or *No*, as quickly as possible using the computer keyboard. Two labels may be posted at the top of the computer screen differentiating the event of interest from the offender-related feature, as a means of aiding orientation.

In the current use of the FEPAT, an event description is presented prior to the feature-event pairing to act as a *prime*. Priming occurs when the presentation of one item (a prime) accelerates the response to a succeeding item (a target) (McNamara, 1994). Here, the processing of the event-related information prior to the feature-related information should prime the user to a limited number of potentially related items in memory, thereby promoting an accelerated rate of recognition if a relationship is present between the two concepts. The use of either the event or the feature as a prime will depend on the decision context from which the feature-event pairings are drawn.

Decision Context

The process of offender profiling involves the examination of information encompassing a criminal act (a series of events), and the inference of offender traits for the purpose of constructing a potential offender description or *profile* (a series of features) (Hazelwood, Ressler, Depue, & Douglas, 1995). As such, it is suggested that profilers, as decision-makers, will use cues to construct a description of the perpetrator. The clear use of associations to solve decision tasks, combined with the diagnostic nature of the task, makes the practice of offender profiling an ideal context for observing cue use within naturalistic decisions.

As profilers are required to undertake a diagnostic task by first familiarising themselves with the event, and subsequently formulating connections with features of the perpetrator, the ordering of the feature-event pairing during the current use of the FEPAT will represent a diagnostic cue, that is, event → feature. For example, the presence of a *blitz attack*, a sudden and unprovoked attack usually from behind (Hazelwood, et al.), may indicate an offender's low level of social competency. However, the presentation of the FEPAT can be manipulated to be used for domains which engage anticipatory cues, whereby environmental features are examined in an attempt to anticipate the occurrence of an event. For example, in the decision-making context of fire-fighting, the environmental feature may be a gas leak, and this may be associated with an explosion. An example of the sequence of one FEPAT trial is illustrated in Figure 1.

Figure 1. An illustrated example of the sequence of screen presentations within FEPAT

Aim and Hypotheses

The aim of the present study is to test whether differences in cue recognition exist across expertise using a computer-based task. In doing so, it is anticipated that the findings will yield a number of cues which are predominantly used by expert decision-makers that may later be embedded in a cue-based training system.

The development of cue-based systems has proven problematic in previous attempts, as it is difficult for designers to agree on the most appropriate cues to embed within the system (Wiggins, 2006). Although the use of cognitive interviews is an accepted form of identifying concepts of interest during a decision task (i.e., features and events of interest), a form of analysis which attempts to objectively identify cognitive associations between such concepts, (i.e., the composition of the cues themselves), and their relative degree of engagement across expertise, remains largely unexplored. The development and validation of an objective cue recognition task, such as the FEPAT, would be a significant aid in both the design of training initiatives based on relevant cue acquisition, and in providing a means for assessing such programs' effectiveness in aiding novice performance.

It is anticipated that the FEPAT will provide an indication of: 1) which features and events share an association (exist as a cue) in participants' long-term memory; and, 2) the strongest cues, or those that demonstrate the highest level of activation across two levels of decision-maker proficiency (expert and novice). Several predictions are made regarding these findings. Specifically, when presented with feature-event pairings in controlled conditions, we hypothesise: 1) greater agreement among experts than novices concerning the cues recognised; 2) low agreement between experts and novices; and 3) lower response latencies for experts than novices for the cues identified by the experts.

METHOD

Design

The study employed a between-subjects design with a single independent variable (IV), expertise, comprising two levels: novice and expert. The dependent variables (DV) were: 1) participants' classification of the associations as valid or invalid (Yes or No responses); 2) the level of classification agreement amongst each sample; and, 3) participants' response latency in relation to their classification.

Participants

As a vast range of profiling perspectives exist, it was essential to draw cues from a range of profiling proponents with education and training from a range of profiling backgrounds and philosophies. Expert decision-makers ($n = 8$) were recruited from a pool of offender behavioural analysts, forensic/clinical psychologists, and forensic investigators with an extensive level of operational experience in the domain of offender profiling. Novice decision-makers ($n = 20$) were drawn from several learning backgrounds consistent with those of the experts.

As participants' decision performance was not assessed, expertise was defined in accordance with Shanteau's (1984) approach to the classification of expertise where the novice is identified and the expert is defined relative to the novice on a continuum of expertise. This approach assumes that expertise is a level of proficiency to which novices are able to progress. The level of proficiency can then be grossly assessed in terms of years of operational experience and appropriate qualifications within the domain.

Stimuli/Apparatus

Preceding the current experiment, a cognitive interview method was employed to examine novice and expert profilers and identify and extract the key components of cues – features and events. The technique involved the use of a semi-structured interview protocol which presented a critical incident and several probe questions to elicit the information pertinent to the identification of the unknown offender.

Using content and thematic analyses, and two independent coders, this process yielded a total of 41 event-related and 28 feature-related concepts from both the novice and expert populations. An inter-coder agreement rating of 88.2% was established between each coder for the entire population overall. This result demonstrates a high level of inter-coder reliability (>80%, see Kurasaki, 2000).

The FEPAT was presented to participants on laptop computer via DMDX software (Forster & Forster, 2003), which collected responses and response latencies. Two labels were posted at the top of the computer screen differentiating the event of interest (left of screen) from the offender-related feature (right of screen) to aid the orientation of participants. An additional two labels were used to designate each shift key as either the Yes (right shift key) or No (left shift key) response.

Procedure

The participants were seated in front of the laptop computer. Prior to the test trials they completed 12 familiarisation trials, each consisting of a single feature-event pairing. Familiarisation trials consisted of several associations that were

based on assumed common knowledge (e.g., sky + blue). Several incongruent pairs were also used (e.g., sky + crocodile). However, no context-relevant exemplar pairs were provided to avoid exposure bias. Prior to commencing the test condition, participants were provided with the list of profiling-related feature and event concepts to ensure that they were familiar with the terminology used.

Pilot testing was conducted to determine suitable presentation durations and intervals. The initial crime-related event prime had the longest duration to promote the activation of a number of potential associations in participants' memory. For each test trial, participants were briefly presented with a crime-related event description at the centre of the screen for 3292 milliseconds (ms). Subsequently (after a time interval of 1646ms), this screen was replaced with the presentation of the same event description on the left side of screen simultaneously with an offender-related feature description presented on the right of screen for 1646ms. The next trial commenced after a time interval of 1646ms.

For each presentation, participants were instructed to decide whether each pairing was associated by striking either the right (Yes) or left (No) shift keys on the computer provided. The test phase consisted of four blocks of presentations. Each block consisted of 280 or 308 trials that constituted the combination of 10 or 11 events paired with 28 features, a total of 1148 trials in each experimental session. Participants were given the opportunity to rest between blocks, if desired. Each event and its corresponding feature presentation were counterbalanced to control for practice effects.

RESULTS

Data Reduction

Participant association classifications and response latencies (from the point of paired presentation onset), were recorded for each concept pairing presented. Due to the large number of feature-event pairs, many of which were deemed invalid by both groups of participants, it was necessary to reduce the data sample. Each data sample, expert and novice, was reduced for analysis based on two factors: 1) Agreement. First, the pairs were ranked in order of agreement between participants, that is, on whether they recognised an association between a feature and event. The top 20% of pairs ($n = 230$) was selected and then the second factor was applied; 2) Response latency. The remaining feature-event pairs were ranked in order of mean response latency, and the bottom 20% (46) (i.e., the fastest) were retained. This process was performed separately on both the novice and expert data sets. The remaining 46 cues within each set were termed the novice and expert *target cue* samples and were subject to analysis. To aid conceptualisation, the top five cues from each target sample (novice and expert) are provided in Table 1.

Expert Target Sample			Novice Target Sample	
No.	Event (crime-related)	Feature (offender-related)	Event (crime-related)	Feature (offender-related)
1	Destruction of Evidence	Criminal Sophistication	Level of Organisation	Degree of Intelligence
2	Method of Body Concealment	Degree of Intelligence	Destruction of Evidence	Criminal History
3	Method of Approach	Social Competency	Victim Similarity	Relationship with Victim
4	Injury Type/Location	Relationship with Victim	Victim Selection	Criminal History
5	Staging/Undoing/Personation	Degree of Intelligence	Level of Organisation	Criminal Sophistication

Table 1. The top five cues (Event-Feature pairings) from each target sample (expert and novice), based on agreement and response latency.

Differences in Cue Recognition

The aim of the analyses was to determine whether differences in cue recognition existed across the two samples (expert and novice profilers), in relation to cue agreement, cue composition, and response latency.

Cue agreement. The aim of the *cue agreement* analysis was to determine the level of agreement for cue classification amongst both groups to identify each target sample. A frequency analysis revealed that agreement ranged from 87.5 to 100% among the expert sample, demonstrating a relatively high degree of agreement amongst experts. In comparison, within the novice target sample, agreement ranged from 75 to 100%. The findings demonstrate an overall greater level of agreement within the expert population compared to novices in relation to which feature-event pairings represented valid cues.

Cue Composition. The aim of the *cue composition* analysis was to determine whether a significant difference existed between the cues which appeared in the expert and novice target samples. A one-way chi square was used to test

whether a difference existed between the expected and observed frequencies for the number of cues appearing in both the expert and novice target samples. The chi-square analysis revealed a statistically significant difference, $\chi^2(2, N = 98) = 55.88, p < .001$. These results indicate that a less than expected proportion of expert cues appeared in the novice target sample. Indeed, of the 46 cue pairings from the expert cue sample, only six cues (13%) occupy a position within the novice target sample. This demonstrates a substantial difference in cue recognition across expertise, based on the cues which may be targeted within the operational environment.

Cue Response Latency. The aim of the *cue response latency* analysis was to determine whether a difference existed in the response latency for cue recognition across expertise, within the expert target sample. Descriptive statistics for both groups' response latencies revealed definitive and consistent differences across the two groups across all 46 pairings from the expert target sample.

As the differences across the groups appeared largely consistent, it was considered unnecessary to employ a statistical test to examine whether significant differences existed for each of the 46 variables, and that further, such tests would unnecessarily increase the potential for a Type I error. Instead, the expert target sample was collapsed into four subsets of data. These four subsets were established according to response latency (lowest to highest) and consisted the means for the first set of ten, second set of ten, third set of ten, and final 16 cues, of the sample (total of 46 cues). These subset means were compared across expertise.

A One-Way Multivariate Analysis of Variance (MANOVA) was used to test the difference between experts and novices on response latency for cue recognition across the four subsets, within the expert sample. With alpha set at .05, and homogeneity of variance met, the results revealed significant differences in mean response latency between the expert and novice group for each subset with Wilkes Lambda multivariate test of overall differences among groups being statistically significant $F(10, 831) = 3.95, p < .001, \eta_p^2 = .84$. In combination with the previous results, this suggests that, not only do experts and novices differ in terms of the types of features and event relations that constitute cues, but that the response latency in forming this view also differs, with experts responding much faster than novices.

DISCUSSION

The current study aimed to test whether differences in cue recognition existed across expertise using a computer-based task. The FEPAT was developed to test cue recognition, which was assessed via classification agreement within and between the two groups (experts and novices), and cue recognition response latency across the two groups.

The results showed that the expert group were, as a whole, more consistent in their classification of cues. This is an interesting finding given the apparent difference in profiling philosophies investigated. This suggests that the operational context and the practice of profiling itself may determine the cues recognised and used, rather than the learning background. Experts also appear to target different cues to their novice counterparts. Further, experts' response latency was significantly lower than novices' response latency for the expert target sample. On the basis of these results it can be concluded that cue recognition differs between expert and novice profilers in regard both the cues recognised and the speed of recognition.

Limitations and Future Directions

The primary limitation associated with the present study was the use of text-based labels to represent cue components. A decision-maker will most likely use an array of media in the environment. For instance, a profiler will scan a scene for visual evidence such as blood spatter and, for the same case, listen to witness statements (Hazelwood, et al., 1995); a combination of audio and visual information.

At its extreme, this limitation may misrepresent some forms of cue-based information resulting in misdiagnosis. However, the representation and presentation of numerous features and events identified did not allow for a more naturalistic presentation, as many consisted of more abstract concepts, such as the offender's social competency. As a result, the cues were presented in a standardised format which was consistent with the verbal accounts/descriptions observed within the cognitive interview process. This ensured that all of the relevant cues could be included and that any potential effect of varying the presentation medium was minimised. Future studies may seek to explore alternative representations of cue-based information to determine whether information presentation impacts recognition.

A further limitation exists in relation to the proposed relationship between cue strength in memory and response latency. As asserted previously, the strength of a cue is determined by the number of its successful applications (Anderson, 1993). This would suggest that some cues are inherently stronger than others, based on a number of factors, including their relative level of activation. However, it is important to investigate further the proposed relationship between cue strength and recognition response latency, to determine whether a correlation exists between decision-makers' recognition latency and their perceived cue strength. Such an investigation will provide further support for the use of FEPAT as a valid process for identifying the most relevant cues to experienced decision-makers.

One final limitation associated with the study is that, although definitive differences were identified, it remains relatively unclear whether the demonstrated differences in cue recognition are a reflection of, or a precursor to,

expertise. If the differences are a reflection of expertise, it is possible to use the FEPAT as a form of competency test. On the other hand, if they are a precursor, it may be possible to use the FEPAT as a cue selection process for both computer-based training and decision support systems. The precise role of cues in expertise development and overall decision performance will be investigated in future studies.

Conclusion

The experimental findings suggest that key differences exist in cue recognition and use as a function of one's level of expertise and that these differences may, in part, differentiate novice from expert performance. Further, it would appear that the FEPAT was successful as an objective means of gauging a cue's relative strength, and, consequently, led to the identification of a number of cues that appear to be used frequently by experts, and that novices may not recognise or use as frequently. In enabling novice' acquisition of these high frequency expert cues through the use of cue-based training and decision support systems, decision performance may be improved in a manner which protects the integrity of the system, in this case, the investigative process, by reducing the potential for serious error, while advancing the cognitive skills of the operator to a higher level of expertise.

REFERENCES

- Ackerman, B. P., & Rathburn, J. (1984). The effect of recognition experience on cued recall in children and adults *Child Development*, 55(5), 1855-1864.
- Anderson, J. R. (1993). *Rules of the mind*. Hillsdale, NJ: Lawrence Erlbaum Associates.
- Einhorn, H. J., & Hogarth, R. M. (1981). Behavioral decision theory: Processes of judgment and choice. *Annual Review Psychology*, 32, 53-88.
- Eveleth, D. M. (1999). Analogical reasoning: When a non-expert reasons like an expert. *Journal of Behavioral and Applied Management*, 1(1), 28-40.
- Forster, K. I., & Forster, J. C. (2003). DMDX: a Windows display program with millisecond accuracy. *Behavior Research Methods, Instruments, and Computers*, 35, 116-124.
- Gigerenzer, G., & Todd, P. M. (1999). Fast and frugal heuristics: The adaptive toolbox. In *Simple heuristics that make us smart*. (pp. 3-34): Oxford University Press.
- Griffin, D., Kahneman, D., Aspinwall, L. G., & Staudinger, U. M. (2003). Judgmental heuristics: Human strengths or human weaknesses? In *A psychology of human strengths: Fundamental questions and future directions for a positive psychology*. (pp. 165-178): American Psychological Association.
- Hazelwood, R. R., Ressler, R. K., Depue, R. L., & Douglas, J. C. (1995). Criminal investigative analysis: An overview. In R. R. Hazelwood & A. W. Burgess (Eds.), *Practical aspects of rape investigation: A multidisciplinary approach* (pp. 115-126). Boca Raton, FL: CRC Press.
- Junior police recruited to elite command post. (2003, June 5). *ABC News* Retrieved March 9, 2007, from <http://www.abc.net.au/news/stories/2003/06/05/872395.htm>.
- Klein, G. A. (1997). Developing Expertise in Decision Making. *Thinking & Reasoning*, 3(4), 337-352.
- Klein, G. A., Calderwood, R., & MacGregor, D. (1989). Critical decision method for eliciting knowledge. *IEEE Transactions on Systems, Man, & Cybernetics*, 19(3), 462-472.
- Kurasaki, K. S. (2000). Intercoder reliability for validating conclusions drawn from open-ended interview data. *Field Methods*, 12(3), 179-194.
- Larkin, J., McDermott, J., Simon, D. P., & Simon, H. A. (1980). Expert and novice performance in solving physics problems. *Science*, 208, 1335-1342.
- McNamara, T. P. (1994). Theories of priming: II. Types of primes. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 20(3), 507-520.
- Meyer, D. E., & Schvaneveldt, R. W. (1971). Facilitation in recognizing pairs of words: Evidence of a dependence between retrieval operations. *Journal of Experimental Psychology*, 90, 227-234.
- O'Hare, D., Wiggins, M., Williams, A., & Wong, W. (2000). Cognitive task analyses for decision centred design and training. In J. Annett & N. A. Stanton (Eds.), *Task analysis* (pp. 170-190). London: Taylor & Francis.
- Shanteau, J. (1984). *Some unasked questions about the psychology of expert decision-makers*. Paper presented at the Proceedings of the 1984 IEEE conference on systems, man, and cybernetics, New York.
- Wiggins, M. (2006). Cue-Based processing and human performance. In I. W. Karwowski (Ed.), *Encyclopedia of ergonomics and human factors* (pp. 3262 - 3267). London, UK: Taylor and Francis.

- Wiggins, M., & O'Hare, D. (1995). Expertise in aeronautical weather-related decision making: A cross-sectional analysis of general aviation pilots. *Journal of Experimental Psychology: Applied*, 1(4), 305-320.
- Wiggins, M., & O'Hare, D. (2003). Expert and novice pilot perceptions of static in-flight images of weather. *International Journal of Aviation Psychology*, 13(2), 173-187.