Using Temporal Occlusion to Assess Carrier Landing Skills

Webb Stacy, Jeff Beaubien, Sterling Wiggins
Aptima, Inc.
Woburn, MA & Dayton, OH
wstacy@aptima.com, jbeaubien@aptima.com, swiggins@aptima.com

Melissa Walwanis
Naval Air Warfare Center Training Systems Division
Orlando, FL
melissa.walwanis@navy.mil

Amy Bolton
Office of Naval Research
Arlington, VA
Amy.bolton@navy.mil

ABSTRACT

Military skills often have perceptual and motor components that need to be trained and measured. An emerging approach for doing so is temporal occlusion. In this approach, videos are played and stopped just before a key event, and the participant is asked to make a judgment about subsequent events. The term temporal occlusion refers to the blacking out of the time period following the key event, forcing the participant to use available perceptual cues to correctly anticipate what will happen next. The paradigm has been used successfully in research on sports such as tennis (Ward, Williams, & Bennett, 2002) and baseball (Fadde, 2006) but has not yet found widespread use in military training (Williams, Ericsson, Ward, & Eccles2008).

In this paper, we will discuss the use of a temporal occlusion paradigm to assess the perceptual skills of expert and novice pilots as they land on an aircraft carrier. Videos were created from expert landings in a simulator, and subject matter experts (SMEs) identified the situations that would most clearly require either standard or aggressive corrections in order to stay on track for a skilled and safe landing. The resulting stimuli were used in tests administered on an ordinary laptop computer before and after training sessions. The temporal occlusion test was embedded in an overarching experiment concerning the relation of simulator fidelity to training effectiveness, and it was used to assess pilots' implicit perceptual learning during the experiment.

Test results were used to measure the degree to which expert-novice differences on the pretest were reduced in the posttest; that is, the degree to which novice perceptual performance moved in the direction of expert perceptual performance. We will discuss the temporal occlusion results from the experiment and will conclude by discussing several other promising uses of the approach.

ABOUT THE AUTHORS

Dr. Webb Stacy is a Corporate Fellow at Aptima, Inc., where he is responsible for enhancing Aptima's technology portfolio. Dr. Stacy has an interest in using modern Cognitive Science to improve experiential training. His recent work includes investigating the relationship of simulator fidelity to training effectiveness and developing an approach to optimizing the training value of experiential scenarios. Dr. Stacy holds a Ph.D. in Cognitive Science from SUNY/Buffalo, and a B.A. in Psychology from the University of Michigan, and is a Program Chair for the Society for Behavior Representation in Modeling and Simulation.

Dr. Jeff M. Beaubien is a Principal Scientist in the Advanced Cognitive Training Systems division at Aptima, Inc. For the past 15 years, his work has focused largely on training and assessing leadership, teamwork, and decision-making skills. His work has been sponsored by the U.S. Navy, the U.S. Army, the U.S. Air Force, the Federal Aviation Administration, and the Telemedicine and Advanced Research Technologies Center, among others. Dr. Beaubien holds a Ph.D. in Industrial and Organizational Psychology from George Mason University, an M.A. in Industrial and Organizational Psychology from the University of New Haven, and a B.A. in Psychology from the University of Rhode Island.

Sterling Wiggins is a Senior Cognitive Scientist at Aptima, Inc. He leads several projects creating technology and training to support humans working in high risk, safety-critical environments. His research interests include Live, Virtual, and Constructive training, human-automation interaction, and adaptive aiding. He holds an M.A in Education with a focus on learning, design, and technology from Stanford University.

Melissa M. Walwanis is a Senior Research Psychologist at the Naval Air Warfare Center Training Systems Division. She has an integrated research program devoted to transitioning state-of-the-art products to enhance the training and operational capabilities of the nation's war fighters. Her research portfolio, addresses large scale distributed training, scenario authoring, performance assessment, training fidelity needs analysis, and Live, Virtual and Constructive (LVC) training. She has a M.S. in I/O Psychology from the University of Central Florida and is pursuing a Ph.D. in I/O Psychology at the Florida Institute of Technology.

Dr. Amy Bolton is a Program Officer at the Office of Naval Research (ONR). She manages several programs within the Capable Manpower Future Naval Capability. Capable Manpower is a multi-million dollar per year initiative that addresses manpower, personnel, training and human system design Science and Technology challenges for the Navy and Marine Corps. Dr. Bolton’s research interests include adaptive training, human behavior modeling, human system design and Live, Virtual, and Constructive training. She holds a Ph.D. in Applied Experimental and Human Factors Psychology from the University of Central Florida.

Using Temporal Occlusion to Assess Carrier Landing Skills

Webb Stacy, Jeff Beaubien, Sterling Wiggins
Aptima, Inc.
Woburn, MA & Dayton, OH
wstacy@aptima.com, jbeaubien@aptima.com, swiggins@aptima.com

Melissa Walwanis
Naval Air Warfare Center Training Systems Division
Orlando, FL
melissa.walwanis@navy.mil

Amy Bolton
Office of Naval Research
Arlington, VA
Amy.bolton@navy.mil

The cortex is an organ of prediction.
--Jeff Hawkins, 2004.

What do soccer players learn as they gain experience executing a soccer play? What do physicians learn as they gain experience reading real Echo Cardiograms (ECGs)? What do pilots learn as they gain experience landing on a carrier? In the laboratory, the evidence (Williams et al., 2008; Reber, Beeman, & Paller, 2013) shows that a major component of what is learned while practicing such complex perceptual motor skills is implicit, that is, learned outside of the awareness of the individual. Soccer players learn implicit relative locations of other players, physicians learn patterns in ECGs that are not explicitly taught beforehand, and pilots learn to perceive and respond to cues that they may not have encountered in flight school.

Psychologists have long known that skill learning involves multiple stages that are increasingly autonomous and decreasingly cognitive (Fitts & Posner, 1967) Recently, it has become clear that implicit memory—which corresponds most directly to the autonomous level—involves a different brain network than explicit memory (Chien & Schneider, 2012; Kim, Ritter, & Koubeck, 2011; Sanchez & Reber, 2013;). This means that people who have developed and who use perceptual motor expertise, such as those in the examples above, exercise different brain networks than those without the expertise. The implication is that the acquisition, decay, and measurement of perceptual motor expertise is different from the acquisition, decay, and measurement of more explicit forms of knowledge.

Mental life regularly involves prediction (Bar, 2011; Hawkins & Blakeslee, 2004), and this is particularly true about perceptual and perceptual motor expertise (Cheung & Bar, 2012; Grossberg, 2011). Perceptual motor expertise involves implicit learning of patterns in the environment, and recognizing those patterns generally involves prediction of a complete pattern from a partial pattern. The soccer player will implicitly know where an opposing player will be headed on the field, the physician will implicitly know that certain features of the ECG are likely to be followed by certain others if the patient has heart disease, and the pilot will learn that certain configurations of cues will result in the need to manipulate the throttle in a certain way. F/A-18 pilots regularly talk about "proactive flying" for this very reason.

The experts' ability to (implicitly) predict these patterns gives rise to an opportunity for measuring that expertise, namely to tap into the experts' ability to predict what is going to happen. An emerging approach for doing so is *temporal occlusion*. In this approach, videos are played and stopped just before a key event and the participant is to make a judgment about subsequent events. The term temporal occlusion refers to the blacking out of the time period following the key event, forcing the participant to use available perceptual cues and implicit patterns to correctly anticipate what will happen next. The paradigm has been used successfully in research on sports such as tennis (Ward et al., 2002) and baseball (Fadde, 2006), but has not yet found widespread use in military training (Ward et al., 2008, Williams et al., 2008).

LANDING A JET ON A CARRIER

Carrier landing is one of the most difficult and dangerous tasks a pilot must perform. The runway is very short and it moves. There is very little room for error. Although pilots who have successfully landed on carriers represent a variety of experience levels, even the least experienced of them is very skilled. Thus, it is extremely likely that they are relying on implicit expertise when they land.

Figure 1. Carrier Landing Situation. Calm seas, wake churn, and bow waves indicate that the pilot will need to keep the right wing low during landing.

As an example of the kind of perceptual patterns they internalize, consider Figure 1. It shows a typical view of a carrier. Non-pilots might notice the aircraft on deck or the antennas on the tower. To an expert pilot, however, other features are more salient when landing. In particular, the sea around the carrier is calm, there is significant churn in the wake, and there are four or five bow waves. This indicates that there is little wind in the environment and that the carrier is moving at 25-35 knots to create its own wind in order to make it easier for pilots to land. The carrier-generated wind over the deck, however, has a direction parallel to the carrier's travel and not down the angled landing area. Experienced pilots will know that this will require them to keep their right wing low during landing to ensure proper lineup with the landing area.[1]

Learning Carrier Landing in a Simulator

Carrier landing skills require a fair amount of implicit skill, and currently, a great deal of that implicit skill is acquired in live aircraft. Live aircraft, of course, present many of the implicit perceptual and motor patterns that the pilots will encounter after training, so in many ways present a useful training environment. However, there are three challenges with live aircraft as a training environment. First, the opportunity to practice landing on a carrier is limited; carriers have many missions to accomplish and training pilots to land is just one of them. Second, there is a limited ability to control environmental and other factors. Wind, sea-state, visibility, equipment status, and a variety of other factors are generally not under the instructor's control. This may present students with situations that are either too easy or too challenging for their current level of development. Finally, flying live aircraft is expensive, both because of fuel costs and equipment fatigue.

Using simulators as training environments for carrier landing could address many or all of these challenges. Simulators are (or could be) generally more available than live aircraft, the environment and the state of the aircraft are certainly controllable and are considerably less expensive than live aircraft. But there is uncertainty about the level and quality of fidelity required to support training carrier landing. That is, there is a question about what kind of simulator environment would provide a meaningful subset of the implicit patterns that are encountered in the live aircraft (and would avoid implicit patterns that would interfere with skill transfer.

THE INVESTIGATION

The Overarching Experiment

Stacy, Walwanis, Wiggins, and Bolton (2013) provided a layered framework for assessing the effects of fidelity on training effectiveness. Using this framework, an experiment was designed to assess the effects of improved image generation and of providing motion in the simulator. The study described here is not for the purpose of evaluating temporal occlusion in pilot training; rather it uses a temporal occlusion paradigm as one of the techniques to measure the outcome of fidelity manipulations.

To understand the way that the temporal occlusion measurement technique was deployed, though, it is first important to understand the overarching experiment that involved investigating training effectiveness as a function of visual fidelity, motion fidelity, and condition difficulty.

The overarching experiment involved a 2 (improved visuals vs. standard visuals) x 2 (motion vs. no motion) x 3 (low vs. medium vs. high difficulty) experimental design. Video resolution and scenario difficulty were within-subjects, experimental manipulations and motion was a between-subjects experimental manipulation. Scenario

---

[1] This account is from personal communication with a subject matter expert (SME). Astute readers will wonder why the SME could articulate the account if the knowledge is implicit. The best answer is that the SME had been a pilot instructor, and so had previously reconstructed this particular kind of reasoning in order to explain it to students. The SME believed, consistent with implicit expertise, that in the field, "right wing low" pops into the pilots' heads quickly and does not result from elaborate reasoning.

difficulty was manipulated by systematically altering the environmental conditions of the landing approach. Institutional Review Board approval was obtained for all aspects of the research.

We planned to recruit a sample of 16 active duty F/A-18 C/D pilots, but logistical and other factors required us to delay running eight of them. Hence the results reported below are for the eight pilots we were able to run. There were equal numbers of expert and novice pilots. The entire experiment lasted approximately three hours, decomposed into two blocks of equal time (approximately 90 minutes per day x 2 consecutive days).

At the beginning of day one, participants were provided with informed consent and completed a series of pre-experimental measures in an empty office. The measures included a series of demographic items as well as the first half of the temporal occlusion study, described below. Next, pilots moved to the F/A-18 simulator, where they were outfitted with a series of physiological monitors, such as heart rate monitors and eye-tracking devices. Participants then engaged in "free flight" for five minutes and practiced three simulated carrier landings. Each participant then flew two blocks of six landings each, with a five-minute break in between each block. With the exception of the simulated night landings (which are always flown as "straight in" approaches for safety reasons), pilots started their landing passes from a position abeam of the carrier. This allowed them to set up the approach and land as they would normally.

On day two, as before, pilots flew two blocks of six landings, with a five-minute break in between. At the end of day two, participants completed a total of 12 landings. Unlike day one, however, there was no free flight period prior to commencing the landing runs. The order of presentation for the visual display manipulation was counterbalanced across days. All participants completed a post-simulation questionnaire and the second portion of the temporal occlusion study. Finally, all participants were debriefed about the purpose of the study and thanked for their time.

In addition to the pre- and post- questionnaires and the temporal occlusion study, three categories of pilot performance measures were collected: 1) Observer-based ratings from expert Landing Signals Officers, whose duties on the carrier include providing such ratings; 2) Simulator-based data describing the pilot's deviations from the ideal path during landing; 3) EKG-based workload measures; and 4) eye-tracking measures. We do not discuss the results of the experiment here in order to focus on the temporal occlusion study, but Beaubien, Stacy, and Wiggins (submitted) provides a good account of them.

THE TEMPORAL OCCLUSION STUDY

The temporal occlusion study involved showing pilots brief videos of simulated landing situations and asking for their judgment about whether the situation called for a standard correction or for an aggressive correction. This addressed the pilots' ability to discriminate nominal landing situations from those that were somewhat off-nominal. With experience, this discrimination improves, as discussed in the results section. For the overarching experiment, the question was whether the discrimination improves with training, and, if so, how much the improvement depends on simulator characteristics and on prior experience. For this reason, the temporal occlusion test was administered prior to training and then again after the end of the two-day experimental training session.

In the overarching experiment, two simulator characteristics were varied; 1) the quality of the visual images on the screen and 2) the presence or absence of motion. Motion was a between-subjects variable, so it was easily incorporated in the experimental design of the temporal occlusion study. Unfortunately, since visual quality was a within-subject variable in the overarching experiment, every pilot had encountered both high- and normal-quality visual conditions by the end of the second day. For this reason, it was not possible to use the temporal occlusion study to analyze the effects of visual quality in the simulator.

Participants

Participants were 8 active-duty Navy F/A-18 pilots with a variety of levels of experience. The four novice pilots had a mean service time of 5.75 years and an average of 27 landings on the carrier in the F/A-18, while the four expert pilots had a mean service time of 8.75 years and an average of 112 landings on the carrier. No novices were instructor qualified, but one was qualified as an LSO. Two experts were instructor qualified, and all were qualified as an LSO.

Materials

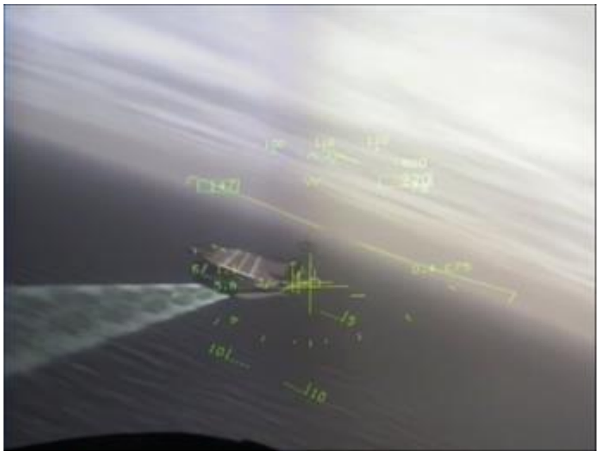

Figure 2. Screenshot from a video clip used in the temporal occlusion study. Pilots were asked whether the situation called for a standard or for an aggressive correction.

Videos for the temporal occlusion study were created by recording the visual simulator output of a number of landings flown by a SME. The SME was asked to fly some landings as if he were a novice, and to fly others as an expert would. The SME was a test pilot and was accustomed to commenting while flying, so he was asked to comment about the situation while landing for both the “novice” landings and the “expert” landings. The session resulted in 90 minutes of video (including the audio commentary.)

This video was then edited into 54 video clips (with audio removed) that were approximately eight seconds in length. Four SMEs subsequently rated each video clip for: 1) whether the clip represented a standard or an aggressive correction and 2) how obvious the answer was. Clips about which there was not a consensus answer or for which half or more of the SMEs did not think the answer was obvious were discarded. To preclude the effects of prior exposure to the stimuli, different clips were used in the pretest and in the posttest. There were equal numbers of clips for standard and for aggressive corrections; and there were an equal number of clips in the pretest and the posttest. Figure 2 shows a screen shot from one of the clips.

Procedure

Video clips were presented in an office setting, away from the simulator, on a laptop PC using an open source software package for building and running experiments called Open Sesame (Mathôt, Schreij, & Theeuwes, 2012). Subjects were briefed on study procedures and detailed instructions were presented on the screen. Participants were asked to respond using the left or the right mouse button (depending on their judgment) at the end of each clip as quickly and as accurately as possible and were given four practice trials to allow them to understand the task fully. There were 17 standard correction and 17 aggressive correction video clips for each session. Clips were presented in random order.

Data Cleaning

Both reaction time (RT) and the correctness of the pilot’s response were recorded. There were two between-subject conditions: 1) whether or not the pilots were trained on the motion-based simulator, called the Motion condition, and 2) whether they were novices or experts, called the Expertise condition. There were also two within-subject conditions: 1) the pre-training vs. the post-training condition, called the Pre/Post condition, and 2) whether the correction was standard or aggressive, called the Correction Type condition.

As a first step, pilot responses were filtered in the following ways:

1. RTs quicker than 200 milliseconds were discarded from all analyses because they almost certainly represent blindly anticipatory responses and, necessarily, guesses. RTs longer than 2.5 times the standard deviation of all RTs were discarded because they almost certainly represent some decision process other than the one under study. Elimination of these outliers follows advice from Ratcliff (1993).
2. Clips for which the aggregate correctness over all pilots was worse than chance were discarded from all analyses because there was probably an ambiguity that rendered them inappropriate for the study.
3. RTs for error responses were eliminated from the RT analysis, leaving only correct RT. This is common practice in the analysis of RT. For obvious reasons, these responses were not eliminated in the analysis of error rates.
4. A logarithmic transformation was applied to RTs. RTs generally have a distribution that is never smaller than 0 and that has a long tail. One distribution that has these characteristics is a Log-Normal distribution, and variables with that distribution become normally distributed under a logarithmic transformation, yielding some beneficial statistical properties. Discussion of these issues may be found in Luce (1986).

As mentioned above, for logistical reasons these results represent the first eight subjects of an experiment that will eventually involve 16 subjects. As a result, the power of the analysis based on these initial data is not what it will eventually be. For this reason, below we report both the traditionally significant results (that is, those whose probability of a Type II error is less than .05) as well as trends that might develop into significant results when data from the additional eight pilots are available. In these cases, the computed probability of a Type II error is reported directly. Further, to avoid stretching the analytic technique inappropriately, we omitted the least theoretically interesting variable (Correction Type) from the analyses below.

RESULTS and DISCUSSION

The results discussed here represent one of the outcomes of the overarching experiment that is investigating the relationship between simulator fidelity and training effectiveness. Temporal occlusion was not used as a training technique in the experiment; instead, it provided one of the dependent variables in the experiment. This means that the results described here do not bear directly on the value of temporal occlusion training. However, as will be seen, temporal occlusion can be a meaningful approach to the measurement of perceptual motor skill development.

An Analysis of Variance (ANOVA) revealed a significant effect of expertise on both RT (F (1,139) = 4.81, p < .05) and error rate (F (1,175) = 7.01, p < .01). Figure 3 shows the nature of this effect. Unsurprisingly, experts were faster and more accurate than novices. The error bars in the graph (and in all subsequent ones) represent the associated standard error of the mean. It appears that both experts and novices became more accurate after training, though in the analysis the interaction between Pre/Post and expertise does not approach significance.

The fact that the study shows that experts are faster and more accurate than novices indicates that the study is getting at an expert perceptual skill of interest, and in that sense it serves to validate the study. Further, there was a trend in the Pre/Post error rates (F (1,175) = 2.45 , p = .119) that showed improvement in accuracy after training, further validating the idea that the two-day training in the simulator had an effect that showed up in the temporal occlusion study.

The ANOVA also revealed a significant expertise by motion interaction (F (1,139) = 4.30, p < .05*)* in RT and a trending expertise by motion interaction in error rates (F(1,175) = 2.133, p=.146.) The nature of this interaction can be seen in Figure 4. RTs for experts in both motion and no motion conditions and for novices in the no motion condition were about the same, but novices in the motion training condition were considerably slower. Error rates add an interesting dimension: They were approximately the same for experts in both training conditions, but were very low for novices in the no motion condition but

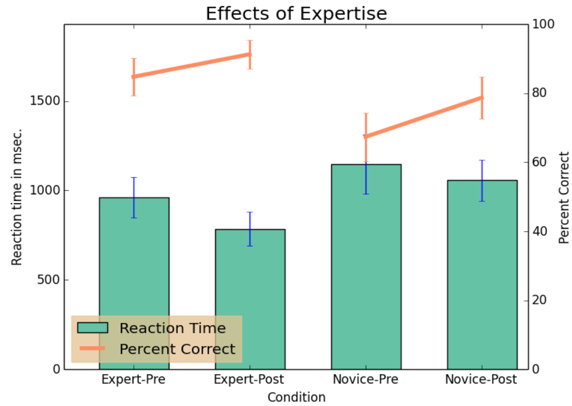

Figure 3. Effects of Expertise on RT and error rate. Unsurprisingly, experts were faster and more accurate. Error bars represent the associated standard error of the mean.

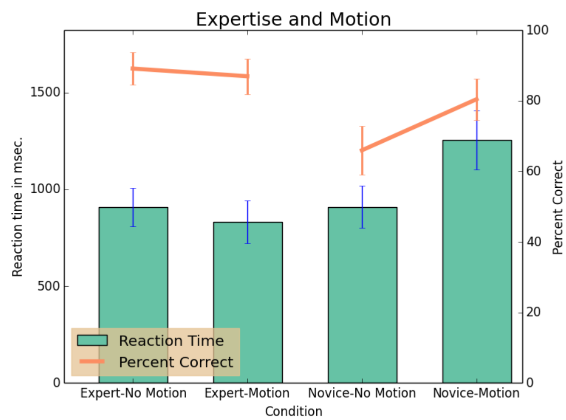

Figure 4. Effects of Expertise and Motion. Experts were about the same in RT and error rate in both training conditions. Novices in the no motion condition were about as fast but less accurate, while novices in the motion condition were slower but quite a bit more accurate. As explained in the text, no effect was due to a speed-for-accuracy tradeoff.

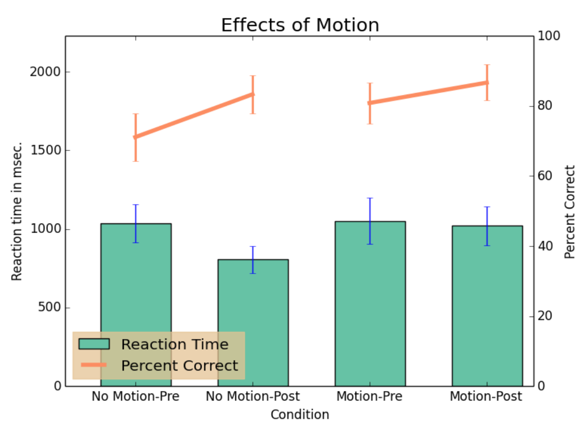

Figure 5. Effects of Motion. There appears to be a trend such that pilots in the no motion condition get faster and more accurate with training, but there appears to be no such effect in the Motion condition.

improved considerably for novices in the motion condition.

The interaction of expertise by motion by Pre/Post conditions was not significant. An examination of the Pre-only RTs and error rates showed that there were no significant differences in RTs or error rates between the pilots assigned to the motion and the no motion training condition, so the interaction is probably not due to pre-existing differences.

One explanation for the increase in accuracy accompanied by a slower response might have been a speed-for-accuracy tradeoff (Fitts, 1954)—pilots may have been more careful in responding. However, as discussed more below, the evidence does not support this hypothesis. In all conditions in this experiment, RTs of error responses were slower than RTs of correct responses--there was actually a speed *disadvantage* to less accurate responses.

To cast more light on this interaction, we now turn to the effects of Motion itself. Though there were no significant or trending effects of motion on RT or errors, it is still interesting to look at the effects of motion as shown in Figure 5. The trend appears to be that pilots in the No Motion condition—those whose training sessions in the simulator did not involve simulated motion—got faster and more accurate, while pilots in the Motion conditions did not. If this result holds up when more data are collected, it will be of great interest, because there will two competing hypotheses with wildly different implications.

The first hypothesis is that the training sessions are, to some extent, modality-specific. Under this hypothesis, the performance of the pilots in the Motion condition may have improved a great deal, but since they had both motion and visual cues present, implicit visual cues played a reduced role. If this hypothesis holds up, then it will be especially important to identify and control the implicit learning required to develop skilled perceptual motor performance, and the fidelity of the entire system (Stacy et al., 2013) is critical for improving performance.

The second hypothesis is that the motion base interferes with learning. Under this hypothesis, the pilots in the Motion condition simply had additional distracting, motion-based stimuli to deal with. If this hypothesis holds up, it will not necessarily imply that motion based simulators do not help improve carrier landing performance; after all, it could have been something about the specific configuration of this motion system—but rather that additional research will be needed to understand more of the specifics of the relationship between motion-based systems in simulators and perceptual motor performance.

To some extent, the results of the overarching experiment will begin to help test these hypotheses. Performance improvement in the motion conditions—as measured by critical variables in the simulator, by LSOs, and by physiological measures of operator state—will support hypothesis 1, and its lack will support hypothesis 2. Ultimately, however, the competing hypotheses can be strongly tested only by performing a temporal occlusion study in the simulator itself, using either the same or different conditions from the ones the trainee has already encountered during training.

No Speed-for-Accuracy Tradeoff

A notable feature of RTs in the temporal occlusion study is that error RTs were consistently slower than correct RTs. Overall, the mean correct RT was 973.95 milliseconds and the mean error RT was 1334.19 milliseconds This difference was about 360 milliseconds large and significant ($t(184) = 2.466$, $p < .05$, two-tailed), and similar results obtained for almost all conditions. The phenomenon has been deemed “slow errors” (Luce, 1986) and it usually happens when the required discrimination is a difficult one. This is very likely the study described here.

Slow errors preclude the more common possibility that errors represent a speed-for-accuracy tradeoff. The reason is that there is nothing to trade off: tolerating a higher error rate will simply make the response slower. This makes taking both RT and error rate into account at the same time more difficult. The most likely model is that pilots have an internal deadline for responding. If they generate an answer they think is correct before the deadline (and it usually will be correct if they believe it so), they respond normally. However, if they cannot generate an answer before the deadline, they simply guess.

Often, in behavioral experiments that involve both speed and accuracy, one or the other is the primary focus. In many situations, for example, error rates don’t change very much from condition to condition, so they can be ignored when comparing conditions. In other situations, the responses are involved enough that error rates predominate over speed.

Because of the focus on quick and accurate reactions, it is necessary that investigations into perceptual motor expertise consider both speed and accuracy. For this purpose, it will be interesting to develop a single metric that accommodates both. In the present study, having a single measure of “goodness” would simplify our understanding of, for example, motion effects, and it would increase the power of any statistical tests as a side benefit. It is not simple to develop a single measure however, because both speed-for-accuracy tradeoffs and slow errors must be addressed. Glickman, Gray, and Morales (2005) and Plesac and Busmeyer (2010) have begun to create models that approach this ideal, and application of those models to temporal occlusion, and to perceptual motor expertise in general, will be rewarding.

Using temporal occlusion to measure learning during this fidelity experiment was a success. In particular, the fact that pilots became both faster and more accurate in the no-motion condition than in the motion condition suggests that either the implicit learning during training was modality specific or that motion hampered learning. These competing hypotheses, with important but opposite implications for fidelity would not likely have surfaced without the temporal occlusion paradigm. Subsequent analyses of other performance measures gathered during the experiment will provide additional insight.

FUTURE WORK

The immediate next step will be to run the remaining eight pilots in the temporal occlusion study to increase its power and to be able to discover whether the main results—the effects of expertise and the expertise by motion interaction—hold up, and to discover whether the effects of motion itself become significant. If so, it will be important to understand whether the relatively larger gains in the no motion condition are due to better implicit learning of visual cues (possibly because the motion pilots had better implicit learning of motion-based cues), or whether the motion condition somehow interfered with implicit learning in general. As mentioned, some of this can be tested by examining other performance measures in the overarching experiment, but a careful follow-on temporal occlusion study performed in the simulator itself will provide the clearest test of the two hypotheses.

The notion of modality-specific implicit learning has important implications for fidelity. Sanchez, Yarnik, and Reber (in press) have quantified the transfer-of-training when the temporal, perceptual, and motor elements of perceptual motor skills were disrupted relative to how they were implicitly trained, and found that while transfer was not all-or-none, there were definite performance deficiencies. Of course, outside the laboratory, the implicit patterns encountered by people learning perceptual motor skills implicitly are not rigidly controlled, and it will be interesting to discover empirically whether varying the implicit patterns during training leads to better transfer perceptual motor skills. The implications for simulator-based training and the fidelity of the training environment that includes the simulator are enormous. To maximize transfer, it will be important to know not only what class of implicit patterns must be reproduced faithfully, and what systematic variations of those patterns leads to the most flexible and transferrable skill acquisition. These effects will best be assessed by taking into account the training environment and human response factors discussed in the Layered Fidelity Framework as discussed in Stacy et al. (2013).

This paper discussed the use of temporal occlusion for skill assessment. Temporal occlusion can also be used directly for training. For example, Fadde (2006) reports successfully teaching baseball batters to recognize pitches before the ball is thrown using temporal occlusion. In the training, he showed the batters videos of a pitcher throwing a ball, stopped the video just before the pitcher was released, and asked the batter to identify the category of pitch. College baseball players showed significant improvements in their batting average and a number of other measures of batter performance. Similarly, it would be possible to use stimuli such as those generated for this study for training in a kind of part-task trainer. For example, imagine that there were a large set of such videos available on a laptop. Pilots in training could watch, respond, and receive feedback. If the implicit patterns they responded to were even partially transferrable to real-world perceptual (and perceptual motor) skills, this could provide an inexpensive, convenient, and low-cost approach.

CONCLUSION

The development of perceptual motor skills is an important part of military training and key components of those skills are implicitly learned perceptual motor patterns. Training environments for these skills should strive to reproduce key implicit patterns in a way that maximizes their transfer to skills in the field. Temporal occlusion is a tool to measure the acquisition of these skills. In this study we have used that tool to begin to understand the nature of those patterns as generated in a simulator and in the field. We hope it has advanced the understanding of perceptual motor skills in some small way, and we look forward to the scientific discovery of new ways to enhance their acquisition effectively and efficiently.

ACKNOWLEDGEMENTS

The views expressed herein are those of the authors and do not necessarily reflect the official position of the organizations with which they are affiliated. We would like to acknowledge the support of the Office of Naval Research and the Naval Air Warfare Center Training Systems Division for their support of this work. This study received strong operational support from the Manned Flight Simulation team at Patuxent Naval Air Station, and in particular Christian Riddle and Steve Naylor, whose excellent efforts kept the simulator operational under sometimes challenging conditions. We would also like to acknowledge the strong support from Jennifer Pagan, Bruce Reiner, and Sandra Velez, whose diligent attention ensured that the experiment could be executed successfully. Finally, we would like to acknowledge the terrific support from our experts: Mike "Sting" Wallace, Beth "Gabby" Creighton, RK "Crater" Creighton, and Mark "Jitters" Kircher.

REFERENCES

Bar, M. (2011). *Predictions and the Brain: Using our Past to Generate a Future*. Oxford: Oxford University Press.

Beaubien, J.M., Stacy, E.W., & Wiggins, S.M. (submitted). Differentiating measures of learning (MOL) from measures of performance (MOP) during aircraft carrier landing practice. *Proceedings of the 2014 Interservice/Industry Training, Simulation, and Education Conference,* Orlando, FL.

Cheung, O. S., & Bar, M. (2012). Visual prediction and perceptual expertise. *International Journal of Psychophysiology*, **83**(2), 156-163.

Chien, J. M., & Schneider, W. (2012). The brain's learning and control architecture. Current Directions in *Psychological Science*, **21**(2), 78-84.

Fadde, P. J. (2006). Interactive video training of perceptual decision-making in the sport of baseball. *Technology, Learning, Cognition, and Instruction*, **4**(3), 265-285.

Fitts, P. M. (1954). The information capacity of the human motor system in controlling the amplitude of movement. *Journal of Experimental Psychology*, **47**(6), 381–391.

Fitts, P. M., & Posner, M. I. (1967). *Human Performance*. Brooks/Cole.

Glickman, M. E., Gray, J. R., & Morales, C.J. (2005). Combining speed and accuracy to assess error-free cognitive processes. *Psychometrika*, **70**(3), 405-425.

Grossberg, S. (2011). Cortical and subcortical predictive dynamics and learning during perception, cognition, emotion, and action. In M. Bar (Ed.), *Predictions and the Brain: Using our Past to Generate a Future*. Oxford: Oxford University Press.

Hawkins, J., & Blakeslee, S. (2004). *On Intelligence*. New York: Times Books.

Kim, J. W., Ritter, F. E., & Koubeck, R. J. (2011). An integrated theory for improved skill acquisition and retention in the three stages of learning. *Theoretical Issues in Ergonomics Science*, **14**(1), 22-37.

Luce, R. D. (1986). *Response Times: Their Role in Inferring Elementary Mental Organization*. New York: Oxford University Press.

Mathôt, S., Schreij, D., & Theeuwes, J. (2012). OpenSesame: An open-source, graphical experiment builder for the social sciences. *Behavior Research Methods*, 44(2), 314-324.

Plesac, T. J., & Busmeyer, J. R. (2010). Two-stage dynamic signal detection: a theory of choice, decision time, and confidence. *Psychological Review*, 117(3), 864-901.

Ratcliff, R. (1993). Methods for dealing with reaction time outliers. *Psychological Bulletin*, **114**(3), 510-532.

Reber, P. J., Beeman, M., & Paller, K. A. (2013). Human memory systems: A framework for understanding the neurocognitive foundations of intuition. In D. D. Schmorrow & C. M. Fidopiastis (Eds.), *Augmented Cognition/Human Computer Interfaces International.* Heidelberg: Springer-Verlag.

Sanchez, D. J., & Reber, P. J. (2013). Explicit pre-training instruction does not improve implicit perceptual-motor sequence learning. *Cognition*, **126**, 314-351.

Sanchez, D. J., Yarnik, E. N., & Reber, P. J. (in press). Quantifying transfer after perceptual-motor sequence learning: how inflexible is implicit learning? *Psychological Research*, in press.

Stacy, W., Walwanis, M., Wiggins, S., & Bolton, A. (2013). Layered fidelity: an approach to characterizing training environments. *Proceedings of the Interservice/Industry Training, Simulation, and Education Conference*, Orlando, FL.

Ward, P. Farrow, D. Harris, K.R., Williams, A. M., Eccles, D. W., & Ericsson, K. A. (2008). Training perceptual-cognitive skills: Can sport psychology research inform military decision training? *Military Psychology*, **20(Suppl 1)**, S71-S102.

Ward, P, Williams, A. M., & Bennett, S. J. (2002). Visual search and biological motion perception in tennis. *Research Quarterly for Exercise and Sport*, **73**, 107–112.

Williams, A. M., Ericsson, K. A., Ward, P., & Eccles, D. W. (2008). Research on expertise in sport: Implications for the military. *Military Psychology*, **20(Suppl 1)**, S123-S145.