The final version was published as: Mavin, T. J., & Roth, W.-M. (2015). Optimizing a workplace learning pattern: A case study from aviation. *Journal of Workplace Learning, 27*, 112–127.

Introduction

Preparing people for the workplace is a significant and enduring task for society. Whereas there have been broad and ongoing critiques on how learning within training settings transfers to the workplace (e.g. Billett and Choy, 2013; Eraut, 2010; Schön, 1983; Tennant, 1986) and how training prepares people for specific team workplace settings, such as the maritime industry (e.g. Emad and Roth, 2008), medical industry (e.g. Harteis *et al.*, 2012; Smith *et al.*, 2010) and aviation (e.g. Roth *et al.*, 2014a), there still remain unanswered questions connecting the best methods of instruction with work.

As a means of investigating how teams learn and how people conduct continuous professional development, recent research has begun a series of studies categorizing formal training and informal learning (Jacobs and Parks, 2009). More recently, studies have begun team learning patterns at work (Baert and Govaerts, 2012), where *learning pattern* denotes a range of prescribed "learning activities and the conditions at the workplace that are used for informal learning, this all over a certain period of time in an organisation" (p. 539). In these studies, which to date have focused on public service employment, five team learning patterns have been defined: the Olympic learning pattern; the helpdesk learning pattern; entrepreneurial learning pattern; the job performance learning pattern; and the agora learning pattern. One pattern, the *job performance learning pattern*, which focuses training on actual or desired job performance was viewed in a more negative light because the characteristics of work reduced the opportunity to learn. However, the negative aspects of job performance may be specific to public service employment broadly or to the group involved in the study. Might there by industries where the job performance learning pattern leads to positive outcomes and experiences among employees?

In some industries – particularly those involving high risk and substantial impact on the social and natural environment, such as aviation, and the maritime, petrochemical and nuclear industries – job performance is critical. Not surprisingly, therefore, incidents in these industries

have been attributed to "human error", which resulted in significant human, environmental and financial loss, as illustrated by hundreds of commercial aviation accidents, the Chernobyl and Three Mile Island nuclear disasters, and the *Exxon Valdez* oil spill (Flin *et al.*, 2008). This study, carried out in collaboration with an airline, was designed to investigate five different modes of a job performance learning pattern currently in use. The ultimate purpose of the study was to identify one or a combination of modes that best met the needs of the workforce and employer. It therefore responds to this journal's stated goal of "focus[ing] on learning in, from and for the workplace" to investigate "the nature of interventions that might assist the learning process and of the roles of those responsible directly or indirectly for such interventions".

Understanding job performance

A variety of piloting performance dimensions are required to deal safely with both normal flights and emergencies. Some of these are technical dimensions, which are critical not only for pilot performance (e.g. Johnston *et al.*, 2000) but also for employees in other high-risk industries, such as officers on the deck of an aircraft carrier (e.g. O'Connor and Long, 2011) and medical practitioners (e.g. Riley *et al.*, 2011). In the case of aviation, the technical performance aspects include (a) factual knowledge pertaining to the aircraft and standard operating procedures and (b) the physical handling of the aircraft in different situations (for example, crosswind landing or engine loss). Today, most practitioners exhibit high levels of technical skills. More frequently, accidents can be attributed to failure of non-technical skills (e.g. Flin *et al.*, 2008; Helmreich, 2000). Non-technical performance dimensions include decision-making, management, communication, and situational awareness (e.g. Flin and Maran 2004).

Technical and non-technical performance dimensions are often considered separately. However, many airlines in the Tasman region have adopted a model of pilot performance that conceptualizes performance in a holistic manner, including two technical components

(knowledge/procedures and flying within tolerances) and four non-technical components (decision-making, management, communication and situation awareness) (Mavin and Roth, 2014a; Mavin *et al.*, 2013). Given their adopted model of pilot performance, the airlines are challenged to find instructional modes that maximize learning opportunities for the pilots. This is important in the context of the current regulator-imposed practice that requires each pilot to undergo two annual 2-day assessment sessions, the results of which are used for deciding whether the pilot may continue to fly or will be taken off flying duty to receive training. Maintaining high performance levels is in the interest of all stakeholders, because failures may lead to the loss of human lives (including those of the crew and passengers, and bystanders), as well as financial losses to the airline, should an aircraft crash in a densely inhabited zone.

Modes of instruction available for job performance training

In response to the question (and challenge) concerning the optimization of learning opportunities that address all six performance dimensions, how might airlines (or companies in other high-risk industries) make arrangements for the most suitable set of instructional settings that best meet their individual and collective needs? The following review shows that in the context of aviation, various modes of instruction are required to provide pilots with workplace-related learning experiences suitable for improving different performance dimensions.

Differing instructional settings constitute different modes of learning. For example, if we use the performance dimension of *technical knowledge* as an area of training focus – like knowledge of aircraft systems, flight rule and regulation, or company procedures – which mode of instruction provides learning experiences that actually transfer to the flight deck (that is, the workplace)? One of the choices would be classroom-based instruction. After all, technical knowledge is factual, and generally stated in the relevant operations manual. Through a method often referred as *didactic instruction*, the instructor, viewed to be an authority, disseminates information in various forms to the learner, who is often viewed as passive

recipient (Roessger, 2012). Forms of didactic instruction include lectures, facilitated discussions, demonstrations or written notes on a board, either a white/blackboard or electronic presentation. Not withstanding economic benefits of classroom-based instruction, a common critique of didactic instruction as a *stand-alone* method is the constructivist notion that learners create meaning from and about the material presented (Knowles, 1973). In this instructional mode, there tends to be little if any opportunity for *actual* practice.

On the other hand, if the training is to teach a pilot to fly an aircraft manually (rather than using the automated pilot), then classroom-based instruction may have considerable limitations. Whereas an obvious choice may be an aircraft, a far safer and more economical means of instruction is a simulator. Over recent decades, aviation simulators have reached a level of fidelity and sophistication where a pilot no longer requires *any* flight-time experience to operate a large commercial aircraft with passengers. Although high-fidelity, full-motion simulators provide the same kind of experience that flying a real aircraft provides, there are suggestions that it may actually make crews less resilient to deal with unforeseen problems (Dahlström *et al.*, 2008).

Simulation is also used for other performance dimensions, including the crew resource management factors of *situational awareness*, *management*, *decision-making* and *communication* (Dahlström *et al.,* 2008). For example, an airline captain may be given a simulated flight scenario to test her ability to deal with a hydraulic failure, which has affected primary flight controls and other important secondary aircraft systems. During training/assessment, a flight examiner will introduce a malfunction – generally at the most inconvenient time for the pilot – requiring the captain to identify the malfunction, determine implications for the safety of the flight and decide a course of action (decision-making), and coordinate other pilots and cabin crew to assist in dealing with the malfunction (management). Whereas, company policy (aviation knowledge) may assist in determining appropriate decisions and crew duties, some scenarios require extensive inter-crew dialogue

(communication) to determine the type of problem and the implications. For example, in a situation where there is a gravely ill passenger, extensive and continuous pilot-to-cabin crew, pilot-to-pilot and pilot-to-air traffic control communication is required. This would be occurring at the same time as the captain continuously maintains awareness of flight parameters, terrain, weather and other aircraft (situational awareness). The high-fidelity simulator provides an appropriate platform for teaching these other performance dimensions (Mavin & Murray, 2010).

Whereas it is important – indeed, mandatory in aviation – to practice scenarios like those outlined, improvements to learning not only occur with practice, but also require reflection on the performance dimensions of practice. In aviation, debriefing pilots following training is integral to practice. Post-training debriefing studies have been shown to be a "multi-layered process of retrospective sense-making, detection and correction of error, social comparison, social control, socialization, and bonding" (Ron *et al.*, 2006, p. 1069). However, pilots may have difficulty entirely and accurately recalling the entire scenario, how it unfolded and, due to high workloads, how other crew performed. To enhance these known difficulties, other methods are introduced into the debriefing session in the form of videos.

Common uses of videos are to complement didactic teaching – for example, by disseminating workplace footage – to improve the rate of psychomotor skills development (Smith *et al.*, 2011). However, recording a learner's performance for later critique is another use of video. Whereas this approach has been used in teacher education for years, it continues to gain support in elite sport (e.g. Richards *et al.*, 2012), commercial aviation (e.g. Mavin and Roth, 2014b), and medical training (e.g. Ward *et al.*, 2003). Such research suggests that when students watch recordings of their own performances, future improvements in performance follows with further gains possible if self-assessment/reflection is conducted with peers and/or senior staff.

Five modes of pilot instruction: Evidence for identifying the right balance

This study was designed jointly with an airline interested in identifying an optimum workplace learning mode, or set of modes, that would provide the best context for workplace learning opportunities, focusing on performance. The five modes that the company reviewed included: (1) traditional classroom-based instruction; (2) a new classroom-based strategy where pilots *individually* assessed the performance of peers in pre-recorded video scenarios; (3) equally ranked pilots assessing the performance of pilots in pre-recorded videos; (4) traditional four-hour simulator sessions followed by 30- to 60-minute debriefing sessions; and (5) a two-hour simulator session followed by a two-hour debriefing session in which pilot performance was replayed by means of a video/software combination ("debriefing tool") (see Table 1). In the following sub-sections, we briefly describe the particulars of each instructional mode, provide a sketch of the research method used and briefly report the findings concerning the effectiveness of each mode as perceived by the pilots.

The investigation as a whole situates itself methodologically in the tradition of the *design experiment* (Brown, 1992) that involves cycles of academic investigation of practice leading to a changed practice that is studied in subsequent research. The method of the design experiment was created to respond to the theoretical and methodological challenges in creating complex interventions. Here we sketch the progression of five cycles of this method to improve learning opportunities of pilots in one airline.

Table 1. **Overview of five modes of instruction**

[Insert Table 1 about here]

Mode 1: Traditional classroom-based instruction – didactic

Description. For years the airline adopted a variety of classroom based instructional methods for annual pilot training. Diverse teaching methods were used, incorporating: (a) direct

delivery of content; (b) scenario-based discussions with pilots talking about topics such as situational awareness, decision-making and communication; and (c) accident and incident investigation of other airlines, where pilots would tease out reasons behind the incident and the proposed likelihood that it may occur in their own airline.

Method. The group comprised 98 pilots consisting of nine senior flight examiners (FE), 46 captains and 43 first officers (FO). A survey was administered using a 1–5 Likert-type scale and short-answer questions. The purpose of the survey was to investigate the workshop participants' preferences for one or the other of the three delivery modes. The survey was designed for the practical purpose of the company to design (more) appropriate delivery modes. The survey, which was designed prior to the academically oriented research began, has limitations as a research tool because it did not undergo rigorous design and validation. However, the results were revealing to the airline and the research team as a starting point for subsequent efforts in designing better learning environments.

Findings. Just under half of all pilots considered previous classroom-based training worthwhile, while almost half responded to it have been average (see Table 2). Possible improvement strategies for classroom-based instruction centered on two general themes. The first related to an over-complication of theory, with a recommendation from a captain being that, as a "simplification – break the training material down to five or six key human factor competencies and focus on training those competencies" and from another captain, "non-technical skills training in the past had often been over-complicated that crewmembers find it hard to find relevance". The second theme was relevance to their job, with suggestions to improve training by "showing how it relates to the way we perform our routines . . . methods of working through problems", "practical subject matter presented in a clear manner", and "working examples".

***Table 2*. Results of traditional classroom-based instruction for three levels of pilots**

[Insert Table 2 about here]

Mode 2: Individual assessment of peers followed by group discussion

Description. In an attempt to create more effective and practically orientated classroom-based instruction, a strategy of *assessment-based training* was instigated. During flight examiner training, a number of junior pilots indicated having learned a tremendous amount while involved in a training module focusing on assessment. After extended discussion and planning between the airline and a university team, a one-day program was developed. The first step was to develop a new simplified assessment form based on a *holistic model of performance* that was available from the literature (Mavin and Roth, 2014a). To assess workplace performance realistically, the company pre-recorded numerous videos in the company simulator. The professionally developed videos used (a) company pilots' and air traffic controllers' voices to increase fidelity, and company cabin crew used during pilot/cabin communication during emergencies; (b) a variety of weather combinations; and (c) a variety of emergencies, from simple malfunctions to complex issues. Prior to assessing three extended videos in a classroom-based session, the pilots were given coaching on performance assessment using the company's human factor based assessment metric. A joint classroom-based discussion followed, facilitated by a company pilot. Discussions focused on "what went right/wrong", "why", and "improvement strategies" for pilots in the video.

Method. On completion of the training day, the pilots, as in Mode 1, completed a survey that complied with the standards of psychological research (Mavin *et al.*, 2013).

Findings. Even though Mode 2 was classroom based, there was improvement in how pilots received the training over Mode 1. For example, two-thirds of the pilots believed assessing pilot performance like a flight examiner improved their learning. Second, over 75 per cent of the pilots had been stimulated to think about their own practice. Third, two-thirds of the pilots developed enhanced action plans for their own practice (Table 3). Pilots' comments from the

survey included: “It gave me the ability to calibrate my observations – some of which were wrong” (flight examiner), “I can relate to videos of our own procedures and simulator easier” (captain), and “Good to sit back and do an analysis of a problem and to understand a better way to manage this situation” (first officer).

***Table 3*. Results from assessment-based training for three levels of pilots**

[Insert Table 3 about here]

Mode 3: Assessment of peers with a fellow pilot of similar rank

Description. This mode of instruction was similar to Mode 2. However, rather than having pilots assess videos individually and carry out discussions in a larger classroom environment – with pilots of different rank in the room such as captain and first officer – the training manager suggested that pilots should be of equal rank and assess in pairs to delimit possible rank/power issues that still dominate the industry. To determine its effectiveness, a trial with a smaller group of eighteen pilots was conducted (six flight examiners, six captains and six first officers). Each pair of pilots used *one* assessment form for each pilot assessed. A review of the research literature on small groups working towards the joint production of a written form showed that this increased the level of interaction (Roth and McGinn, 1998).

Method. The study implemented the think-aloud protocol (Roth and Mavin, 2014), a standard method for investigating expertise. Eighteen pilots (six flight examiners, six captains and six first officers) worked in pairs of same-ranked pilots (e.g. two first officers). Each pair assessed three pre-recorded video scenarios not previously seen. Pairs were asked to arrive at an agreement for their assessment and to support their assessment with reasons. Each session took three hours to complete the assessments. All sessions were video-recorded and transcribed for subsequent analysis.

Findings. There was substantial variance in how the pairs assessed the performances in the scenarios, sometimes reaching from 1 to 5 on the same performance component (e.g. situation awareness or decision-making). However, in each case the pairs provided ample, good reasons for their assessment. Figure 1 shows how pilots of differing rank assessed three flight scenarios for the six performance dimensions of technical skills (aircraft flown with tolerance [TOL] and technical knowledge [KN]) and non-technical skills (situational awareness [SA], decision making [DM], management [MA], and communication [CO]). Each pair (flight examiner, captain and first officer) not only gave different assessment scores but also undertook the assessment process differently. For example, all three flight-examiner and two captain pairs rated the performance of the captain as a fail, yet one of the captain pairs and all of the first officer pairs considered the performance acceptable. In addition to differences in performance grading, it was found that pilot rank also varied the assessment process (Figure 1). For instance, more experienced pilots talked more competently about overall performance, connecting observed causal factors to actual performance outcomes (e.g. success or failure of a flying procedure). Less experienced pilots had a tendency to identify facts rather than causal relationships between factors – for example, by indicating how poor management performance leads to low situation awareness.

[Insert Figure 1 about here]

Figure 1. **How pilots of different rank assess technical and non-technical skills**

Mode 4: Traditional four-hour simulation training with one-hour debriefing

Description. In the first three instructional modes, pilots did not actually engage in the practice of flying. Learning opportunities to fly an aircraft without exposing passengers or aircraft to risk may be arranged in a high-fidelity, full-motion flight simulator. Traditional simulator training generally incorporates a one-hour briefing and a four-hour simulator session, followed by a half-hour to one-hour debriefing (depending on the instructor and student performance).

During training and assessment in the simulator, pilots deal with emergencies in varying flight conditions. The flight instructor, seated behind the pilots, operates the simulator using variety of panels, and communicates directly with the pilots or through an interphone arrangement that simulates cabin crew and air traffic control. Flight examiners make notes on the performance, and either discuss this immediately following an event in the simulator or deal with the issues (both good and bad) in more detail in the subsequent debriefing session. To facilitate the debriefing, which occurs in a room remote from the simulator, a video recorder positioned in the simulator records pilots and aircraft flight parameters. These video devices – debriefing tools – can be replayed during the debriefing session.

Method. The study was designed according to the principles of cognitive anthropology, combining controlled psychological methods and anthropological fieldwork (Roth *et al.*, submitted). The authors observed and recorded 29 simulator sessions. Each of the two pilots and the training captain/flight examiner pilot responsible for the sessions were interviewed individually at the two-hour break (half-time), on completion of the four-hour simulator session, and upon completion of the debriefing. To obtain further data, the entire debriefing session was video-recorded and transcribed for later examination. All interviews and the sessions were transcribed for analysis.

Findings. Several broader themes emerged from this study. First, it was apparent that pilots were showing signs of fatigue after four hours of simulated flight instruction. This delimited what and how they learned. Pilots experienced difficulties recalling what had occurred in the preceding four-hour simulator session, therefore impairing reflection on their practice. That is, whereas, the flight simulator sessions put the pilots through a rigorous testing regime, as required by the regulator, the pilots tended to benefit less in terms of the second purpose of these sessions: reflection.

Mode 5: Two-hour simulation with detailed three-hour debriefing

Description. Traditional approaches to training in the simulator used a 1:4:1 ratio (brief:simulation:debrief). Based on the outcomes of the Mode 4 study, the airline wanted to explore whether shortening the simulator time with a simultaneous increase in the debriefing time would lead to greater learning outcomes.

Method. The study was designed according to principles of the cognitive anthropology of work (Roth *et al.*, 2014b). A two-day trial was planned with two pilots flying a number of short scenarios. On completion of each scenario, they would move back to the debriefing room to watch their performance. The session was video recorded and transcribed for later analysis.

Findings. Asking the pilots to watch longer lengths of video clips in the debriefing set up an unusual dynamic, the most striking of which was that the instructor played a very minor role in the evaluation. The pilots led the debriefing by stopping the video numerous times and discussing, either with each other or the instructor, their thoughts on issues that arose. The captain found the process very good: "It's the perfect way to do it . . . they can tell you exactly what's going through their mind . . . I think that's probably how we should do the training". In general, the pilots had adopted a more critical, interrogative approach to their own performance. In many instances, the two pilots watched a sequence and asked for the video to be stopped, which allowed them to discuss what was going through their mind at the time or why they did what they did. When asked to specifically comment on the instructional method, the first officer replied:

> For me, very beneficial. I look at this and I think of a few obvious points and go *ah* that's when I went to the wrong gauge or searching. While I watch the video here I think a whole lot less about the buttons and more about the manner of the whole thing, frame of mind, the [non-technical skills] interaction.

Lessons learned from the five modes of instruction

The pilots identified both strengths and weaknesses for each of the modes of instruction. It was clear that pilots were experiencing some difficulty transferring Mode 1 (traditional classroom-based instruction) into praxis, that is, onto the flight deck of the aircraft. In contrast, Mode 2 (classroom-based instruction) with assessment-based training, showed improvement in how pilots viewed the classroom workplace transfer. Issues associated with over-complication of theory appeared to reduce considerably. In relation to applicability to practice, a positive move towards pilots changing, or at least considering, their own future practice occurred.

When Mode 2 is compared with Mode 3 (assessment-based training with a peer) there were fundamental implications. First, one assumption was that classroom-based discussions about video performance would be more or less the same among pilots – irrespective of rank and experience. However, this did not turn out to be the case. We found that similar ranked assessment-based training may actually be problematic, with a polarization occurring during discussions, as apposed to a discussion with differently ranked pilots. This polarization could be exemplified with more junior pilots. The implication for instruction was that training once limited to senior pilots – assessment-based training – appeared to be a training technique overlooked as a method suitable for training all pilots. However, using only similarly ranked pilots may not be the most appropriate combination, thus supporting Mode 2. Yet a consequence for Mode 2 would be that, during facilitated debriefing, disagreement about how the performance is viewed between pilots might occur.

In relation to Modes 4 and 5, there were important differences. First, during traditional simulator training, the intensity of the training had pilots showing signs of fatigue, finding post-simulator recall of previous events difficult, whereas, in shorter simulator session, pilots appeared alert and happy to discuss the implications of their actions in detail. Second, given the opportunity – as in the Mode 5 extended debrief – pilots were able to examine their own performance in greater detail. Whereas the differences between Mode 4 and Mode 5 appeared

to show that "*less* (flying) *is more* (learning through reflection)", the limited nature of the study makes the results less definitive, thus requiring further investigation.

Discussion

Recent research focusing on public administrative employment identified five learning patterns that categorized team professional development. Within these categories, it had been demonstrated that professional development centering on job performance was viewed as negative, with the suggestion that "job characteristics here reduce the opportunities to learn" (Baert and Govaerts, 2012, p. 544). Whereas this view may be well supported – and even appropriate – within the particular organizational setting that these authors researched, it may not be generalizable to other types of organizations. Organizations in high-risk industries must place a significant focus on job performance during professional development to increase the resilience of operators to potentially catastrophic failures (e.g., Dahlström *et al.*, 2008). This study was not designed to invalidate the findings of learning patterns, but rather to demonstrate what modes, or combinations of modes, are appropriate for job performance in a particular type of (high-risk, high-stakes) industry that requires continuous high performance levels.

This study exemplifies the use of the method of *design experiment* for improving practice. The airline involved called upon academic researchers to investigate their current practice of instruction, which was changed based on the research results. The new practice was studied leading to changes in or addition of delivery modes. The five such cycles featured here have had tremendous impact on the reconfiguration of pilot training in this airline, and other airlines who have begun to change the delivery of instruction based on the lessons learned.

In traditional classroom-based methods of instruction, there has been significant research and debate on the classroom-to-workplace transfer (e.g. Billett, 2001; Emad and Roth, 2008). However, simply removing instruction from the classroom may not be the answer. In this study, we identified that classroom-based instruction showed varying levels of support with

pilots. By providing real workplace examples – in the form of videos – where pilots were forced to actually assess, it was possible to improve participants' engagement. In this study, this mode of instruction – often referred to as *anchored instruction* – is fixed by the use of videos creating a scenario whereby individuals can critically review complex problems through multiple lenses (Merriam *et al.*, 2007).

Pilots preferred anchored instruction to traditional methods. However, important issues regarding classroom dynamics required resolution. In the present study, this line of thinking was supported, with pilots of similar rank not only assessing performance differently, but also going about performance assessment differently (Mavin *et al.*, 2013; Roth and Mavin, 2014). The implications for the airline training appeared clear, with the most preferable classroom based instruction shown to be Mode 2 (anchored instruction with a variety of ranked pilots). These findings are aligned with research suggesting that during self-assessment, especially poor-performing individuals have difficulty assessing performance accurately (e.g. Dunning *et al.*, 2004; Sitzmann *et al.*, 2010). To explain these observations, it was theorized that individuals unable to *perform* specific tasks were more likely not to possess skills required to *assess* these tasks referred to as the *double paradox* (see Dunning *et al.*, 2003). In a study into expert flight examiners, it was observed that when pilots assess performance, differing skills appear to have varying levels of importance (Mavin and Roth, 2014a). These authors suggest that pilots might use *compensatory* and *non-compensatory* models to assist them during performance assessment.

Modifying classroom-based instruction appears to be important, although there continues to be an abundance of literature supporting the occurrence of the greatest learning when individuals are given guided learning opportunities that are both physically and socially authentic (e.g. Ericsson, 2008). In this study, the last two modes of instruction delivered authentic workplace practice – simulation and debrief – albeit using a mixture of delivery styles. Even though pilots were not critical of the four-hour simulator session, there was

sufficient objective evidence that pilots here, as in other studies (Baltzley *et al.*, 1989), were tired and had forgotten much of what they had been through during the preceding four hours. In contrast, shorter simulator sessions followed by extended self-directed debriefing, were extremely effective for the pilots.

Conclusion

Developing a broader understanding of team learning patterns is important for understanding how teams undergo professional development. As outlined by Baert and Govaerts (2012, p. 538), given the importance of professional learning, organizations "cannot rely on uncoordinated and incidental learning in the workplace alone". To further develop frameworks of workplace learning patterns requires ongoing research within and across organizational settings. Expanding studies into other professions to determine similarities and differences that may exist will assist in developing a more robust learning patterns typology, enabling an examination of the types of learning patterns that describe individual and team professional development within and across occupational settings.

This study potentially has major practical implications, as can be seen from the reactions of a number of airlines in the area to whom we communicated the results presented above. Five airlines indicated interest in a study of debriefing, with a particular focus on the functions of the debriefing tool mentioned above and an assessment model. Our partner airline used the results to make classroom-based video assessment an integral part of practice, not only for all pilots (including new pilots to the airline) but also cabin crew. The airline also used those results to explore and study other modes (e.g. Mode 1 → Mode 2 and Mode 3). Another airline requested a study of the effect of one-on-one instruction using Mode 3 as a model, but pairing pilots who have performance problems with an experienced pilot as a coach. Based on the results of the Mode 4 and Mode 5 studies, our partner airline is in the process of planning a

larger study investigating the simulator time-debriefing time ratios, with this study reducing simulator session to three hours and increasing debriefing to two hours.

In this study, we report how one airline in collaboration with a university-based research team tested different instructional contexts for the purpose of identifying a combination of contexts that maximize workplace learning geared towards performance. The different modes investigated were a function of the particulars of the industry. A decision about which mode of instruction might be transferable (generalizable) to other industries will depend on job characteristics. However, studies focusing on human factors show that some categories, such as situation awareness, management, communication and decision-making, are common across industries (Roth *et al.*, 2014). This may suggest that, in other industries, instructional modes may include the same modes that were investigated in this study. Studies specifically designed to investigate the particular requirements for optimal learning environments in these workplaces are required.

As a whole, our study has implications at another level. Our experience shows industry partners and researchers designing and conducting research, with outcomes that benefit each. Thus the industry partners obtain data and analyses that they can use for evidence-based decision-making regarding an optimal workplace learning model or the identification of a mix of models. Because research publications require high-quality standards to pass the peer-review process, partner airlines prefer final peer-reviewed research publication rather then internally generated reports. The researchers benefit because they can publish studies alone or in collaboration with members of their industry partners. Whereas industry partners can often be secretive – as we know from research with software development companies – our airline partners have come to recognize the correlation between the quality requirements of the peer review process and the hardness of the evidence on which they base decisions.

Acknowledgments

This research was funded from a variety of industry-based (Timothy J. Mavin as Principal Investigator) and university-based (Wolff-Michael Roth as Principal Investigator) projects.

References

Baert, H. and Govaerts, N. (2012), "Learning patterns of teams at the workplace", *Journal of Workplace Learning*, Vol. 24 Nos 7/8, pp. 538–50.

Baltzley, D.R., Kennedy, R.S., Berbaum, K.S., Lilienthal, M.G. and Gower, D.W. (1989), "The time course of postflight simulator sickness symptoms", *Aviation, Space, and Environmental Medicine*, Vol. 60 No. 11, pp. 1043–8.

Billett, S. (2001), *Learning in the Workplace: Strategies for Effective Practice*, Allen & Unwin, Sydney.

Billett, S., and Choy, S, (2013), "Learning through work: emerging perspectives and new challenges", *Journal of Workplace Learning*, Vol. 25 No. 4, pp. 264–76.

Brown, A. (1992), "Design experiments: Theoretical and methodological challenges in creating complex interventions in classroom settings", *Journal of the Learning Sciences*, Vol. 2 No. 2, 141–78.

Dahlström, N., Dekker, S.W.A., van Winsen, R. and Nyce, J.M. (2008), "Fidelity and validity of simulator training", *Theoretical Issues in Ergonomics Science*, Vol. 10 No. 4, pp. 305–14.

Dunning, D. Heath, C. and Suls, J.M. (2004), "Flawed self-assessment implications for health, education, and the workplace", *Psychological Science in the Public Interest*, Vol. 5 No. 3, pp. 69–106.

Dunning, D., Johnson, K., Ehrlinger, J. and Kruger, J. (2003), "Why people fail to recognize their own incompetence", *Current Directions in Psychological Science*, Vol. 12 No. 3, pp. 83–7.

Emad, G. and Roth, W.-M. (2008), "Contradictions in practices of training for and assessment of competency: a case study from the maritime domain", *Education + Training*, Vol. 50 No. 3, pp. 260–72.

Eraut, M. (2010), "Knowledge, working practices, and learning", in Billett, S. (Ed.), *Learning through practice*, Springer, Dordrecht, pp. 37–58.

Ericsson, A.K. (2008), "Deliberate practice and acquisition of expert performance: a general overview", *Academic Emergency Medicine*, Vol. 15 No. 11, pp. 988–94.

Flin, R. and Maran, N. (2004), "Identifying and training non-technical skills for teams in acute medicine", *Quality and Safety in Health Care*, Vol. 13 No. Suppl. 1, pp. i80–i84.

Flin, R., O'Connor, P. and Crichton, M. (2008), *Safety at the Sharp End*, Ashgate, Aldershot.

Harteis, C., Morgenthaler, B., Kugler, C., Ittner, K.P., Roth, G. and Graf, B. (2012), "Professional competence and intuitive decision making: a simulation study in the domain of emergency medicine", *Vocations and Learning*, Vol. 5 No. 2, pp. 119–36.

Helmreich, R.L. (2000), "On error management: lessons from aviation", *British Medical Journal*, Vol. 320 No. 7237, pp. 781–6.

Jacobs, R.L. and Park, R. (2009), "A proposed conceptual framework of workplace learning: implications for theory development and research in human resource development", *Human Resource Development Review*, Vol. 8 No. 2, pp. 133–50.

Johnston, A.N., Rushby, N. and Maclean, I. (2000), "An assistant for crew performance assessment", *The International Journal of Aviation Psychology*, Vol. 10 No. 1, pp. 99–108.

Knowles, M. (1973), *The Adult Learner: A Neglected Species*, Gulf, Houston, TX.

Mavin, T.J. and Roth, W.-M. (2014a), "A holistic view of cockpit performance: an analysis of the assessment discourse of flight examiners", *The International Journal of Aviation Psychology*, Vol. 24 No. 3, pp. 210–27.

Mavin, T.J. and Roth, W.-M. (2014b), "Between reflection on practice and the practice of reflection: A case study from aviation", Reflective Practice: International and Multidisciplinary Perspectives, Available from: doi: 10.1080/14623943.2014.944125.

Mavin, T.J., Roth, W.-M. and Dekker, S. (2013), "Understanding variance in pilot performance ratings: two studies of flight examiners, captains and first officers assessing the performance of peers", *Aviation Psychology and Applied Human Factors*, Vol. 3 No. 2, pp. 53–62.

Merriam, S.B. Caffarella, R.S. and Baumgartner, L.M. (2007), *Learning in adulthood: A Comprehensive Guide*. Wiley, San Francisco, CA.

O'Connor, P. and Max Long, W. (2011), "The development of a prototype behavioral marker system for US Navy officers of the deck", *Safety Science*, Vol. 49 No. 10, pp. 1381–7.

Richards, P., Collins, D. and Mascarenhas, D.R.D. (2012), "Developing rapid high-pressure team decision-making skills: the integration of slow deliberate reflective learning within the competitive performance environment—a case study of elite netball", *Reflective Practice: International and Multidisciplinary Perspectives,* Vol. 13 No. 3, pp. 407–24.

Riley, W., Davis, S., Miller, K., Hansen, H., Sainfort, F. and Sweet, R. (2011), "Didactic and simulation nontechnical skills team training to improve perinatal patient outcomes in a community hospital", *Joint Commission Journal on Quality and Patient Safety*, Vol. 37 No. 8, pp. 357–64.

Roessger, K.M. (2012), "Toward an interdisciplinary perspective: a review of adult learning frameworks and theoretical models of motor learning", *Adult Education Quarterly*, Vol. 62 No. 4, pp. 371–92.

Ron, N., Lipshitz, R. and Popper, M. (2006), "How organizations learn: Post-flight reviews in an F-16 fighter squadron", *Organization Studies*, Vol. 27 No. 8, pp. 1069–89.

Roth W.-M. and Mavin, T.J. (2014), "Experts assessing the performance of experts: an expert/expert study among pilots", *Cognitive Science*, DOI: 10.1111/cogs.12152.

Roth, W.-M., Mavin, T.J. and Munro, I. (2014a), "Good reasons for high variance (low interrater reliability) in performance assessment: A case study from aviation", *International Journal of Industrial Ergonomics*, Vol. 44 No. 5, pp. 685–96.

Roth, W.-M., Mavin, T.J. and Munro, I. (2014b), "How a cockpit forgets speeds (and speed-related events): toward a kinetic description of joint cognitive systems", *Cognition, Technology and Work*, DOI: 10.1007/s10111-014-0292-0

Roth, W.-M., Mavin, T.J. and Munro, I. (submitted), "Flight examiners' methods for ascertaining pilot proficiency".

Roth, W.-M., Mavin, T.J. and Dekker, S.W.A. (2014), "The theory–practice gap: epistemology, identity, and education", *Education and Training*, Vol. 56 No. 6, pp. 521–36.

Roth, W.-M. and McGinn, M.K. (1998), "Inscriptions: a social practice approach to 'representations'", *Review of Educational Research*, Vol. 68 No. 1, pp. 35–59.

Schön, D.A. (1983), *The Reflective Practitioner*, Basic Books, New York.

Sitzmann, T., Ely, K., Brown, K.G. and Bauer, K.N. (2010), "Self-assessment of knowledge: a cognitive learning or affective measure?", *Academy of Management Learning and Education,* Vol. 9 No. 2, pp. 169–91.

Smith, A.R., Cavanaugh, C. and Moore W.A. (2011), "Instructional multimedia: an investigation of student and instructor attitudes and student study behavior", *BMC Medical Education*, Vol. 11 No. 1, pp. 38–51.

Smith, B.R. et al. (2010), "Educating medical students in laboratory medicine: a proposed curriculum", *American Journal of Clinical Pathology*, Vol. 133 No. 4, pp. 533–42.

Tennant, M. (1986), "An evaluation of Knowles' theory of adult learning", *International Journal of Lifelong Education*, Vol. 5 No. 2, pp. 113–22.

Ward, M., MacRae, H., Schlachta, C., Mamazz, J., Poulin, E., Reznick, R. and Regehr, G. (2003), "Resident self-assessment of operative performance", *American Journal of Surgery*, Vol. 185 No. 6, pp. 521–4.

Table 1. **Overview of five modes of instruction**

	Overview	**Student in practice**	**Student being assessed**
Mode-1	Classical classroom based instruction (didactic)	No	No
Mode-2	Classroom based instruction – pilot involved in individually assessing videos followed by group discussion	No	No
Mode-3	Classroom based instruction – pilot involved in assessment of videos with peer	No	No
Mode-4	Pilot involved in high intensity four-hour simulated practice with shorter one-hour debriefing	4-hour simulator session	Yes
Mode-5	Pilot involved in high intensity practice with detailed three-hour debriefing	2-hour simulator	Yes

Note. Describes modes of practice for the pilot verse the practice of assessment. Modes of learning are occurring over an approximate three to five hour period.

***Table 2*. Results of classical classroom based instruction for three levels of pilots**

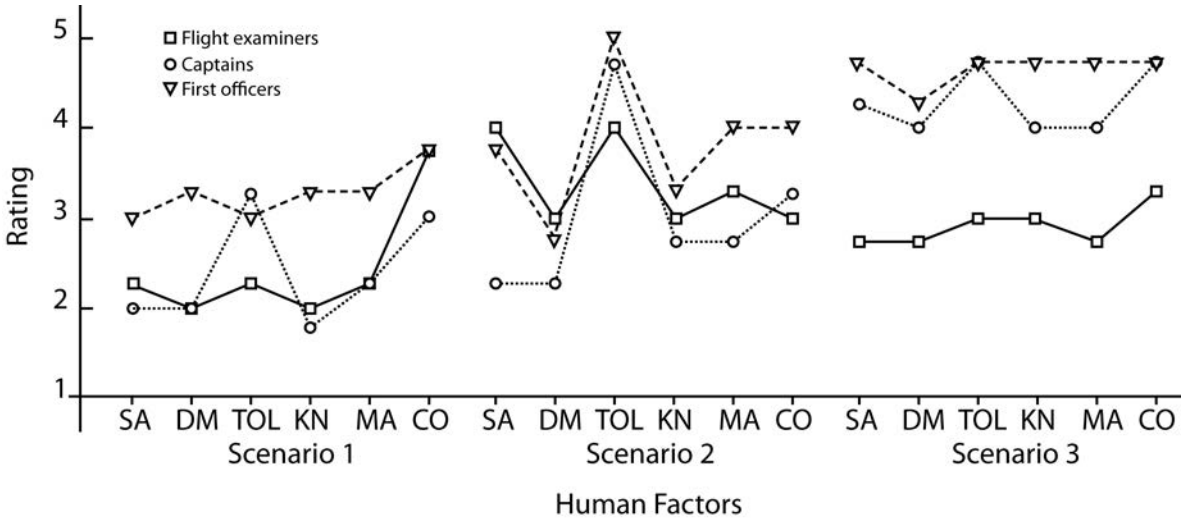
	Not worthwhile %	Middle %	Worthwhile %
1. Classroom-based CRM training has been worthwhile			
Senior flight examiner	0	33%	66%
Captains	5%	52%	43%
First Officers	5%	53%	42%

Results of 1–5 Likert-type scale survey from three groups of pilots. The first group is Senior Flight examiners, pilots given specific regulatory approvals to train and assess pilots within an airline (they are also captains). Captains are pilots designated as first in command of an aircraft. First officers are pilots assigned second in command.

***Table 3*. Results from *assessment-based training* for three levels of pilots**

	Not worthwhile %	Middle %	Worthwhile %
1. Applying the assessment standards to other pilot's performance in the video scenarios improved my learning			
Senior flight examiner	0	0	100
Captains	7.5	32.5	60
First Officers	2	31	67
2. Having to personally apply the assessment standards to the pilot's performance in the video scenario has caused me to think about my own practice			
Senior flight examiner	0	0	100
Captains	4.5	20.5	75
First Officers	2	21.5	76.5
3. Using video scenarios has caused me to develop enhanced action plans for my own future practice			
Senior flight examiner	0	11	89
Captains	4	39	57
First Officers	5	24	71

Figure 1. How pilots of different rank assess technical and non-technical skills



Graphs represent pilots of differing rank assessing three flight scenarios for the six performance dimensions of technical skills [aircraft flown with tolerance (TOL) and technical knowledge (KN)] and non-technical skills [situational awareness (SA), decision making (DM), management (MA) and communication (CO)].