Training for Long-Duration Space Missions: A Literature Review into Skill Retention and Generalizability

Marc A. Pieters * **Peter M. T. Zaal** **

* *San Jose State University, NASA Ames Research Center, Moffett Field, CA 94035 USA and Delft University of Technology, Delft, Netherlands (e-mail: marc.a.pieters@gmail.com).*
** *San Jose State University, NASA Ames Research Center, Moffett Field, CA 94035 USA (e-mail: peter.m.t.zaal@nasa.gov).*

Abstract: On long-duration space missions, skill retention and generalizability become ever more important as mission length increases, for it is through these capabilities that astronaut crews achieve autonomy. Because simulators are used extensively in all types of training, the effects of simulator fidelity on skill retention and generalizability are paramount to understand. A literature survey was performed to identify current research gaps in skill retention and generalizability. The survey identified a need for a structured and quantifiable approach to characterize skill decay, for example, using a cybernetic approach. Such an approach would allow for gaining a deeper understanding of the mechanisms through which skill decay operates. Furthermore, the literature survey identified three research gaps and opportunities for future research: (1) developing skill decay functions provides theoretical insights into skill decay and could allow for several practical applications, such as planning refresher training, (2) investigating the effects of simulator fidelity on skill decay functions could allow for better simulator utilization during training, and (3) investigating the generalizability of skills learned in initial training to other tasks could provide space crews with greater autonomy.

Keywords: Training, Spaceflight, Simulator Fidelity, Skill Decay, Skill Generalizability.

1. INTRODUCTION

Ever since the days of the first manned spaceflights, NASA spent meticulous attention on the training of its astronauts. The goal of training is to realize two main objectives: skill acquisition and transfer of training. Skill acquisition is the initial learning to perform a certain task and transfer of training subsequently projects this onto the operational domain. Especially for long-duration space missions, two additional goals become apparent: retention of skills and generalizability across tasks.

To illustrate this, the NASA Human Research Program evidence report on training deficiencies presents some of the training issues involved with long-duration space missions. One of the main issues is the current training-time-to-mission-time ratio, which is 10 to 1 for International Space Station missions [Barshi and Dempsey (2016)]. A mission to Mars might take up to 32 months [Mars Architecture Steering Group (2009)]: following the current training-to-mission ratio would be unfeasible. Nevertheless, initial training will be a lengthy process and thus skills might decay even before launch. Furthermore, crews will have to be in transit to their destination for several months, which also results in skill decay. This extended transit period gives rise to two more challenges. Firstly, it makes it impossible to send specialized crews upwards in order to perform critical repairs. Secondly, the communication delays that occur between Earth and Mars might last as long as 40 minutes, making real-time support from a mission control center impossible. Complex mission-critical tasks can thus not be handled in the more traditional manners anymore. The role of a specialized crew or real-time ground support has to be taken over by the ability to generalize skills or on-board training.

In order to design the training for long-duration space missions, it thus is important to investigate skill decay and generalizability. Simulators are used for the training of many tasks astronauts will perform. Therefore, understanding the effects of simulation fidelity on skill decay and generalizability is important. The objective of this literature review was to summarize the research to date on skill retention and generalizability, and the impact of simulator fidelity on these two aspects of training. A further aim was to to identify literature gaps relating to these specific areas, and guide future research.

After defining a list of peer-reviewed sources in journals, conference proceedings, or technical reports of relevant organizations, such as NASA, FAA, or similar, these keywords were used in the literature search: skill decay, acquisition, retention, training, transfer of training, simulator fidelity. Furthermore, another choice was made: Nicholas and Foushee (1990) state that space crews for long-duration space missions have to be regarded as groups, instead of individuals. However, they continue by stating that group performance is dependent on three factors, one being the level of skill of its members. Therefore, in this literature review, it was chosen to focus on the individual-astronaut skills, instead of on team skills.

2. TRAINING

2.1 Types of Training

On long-duration space missions, astronauts will train before launch, as well as on board. The application of training on board is dual. On one side, it could serve as refresher training, to ensure the astronauts possess the necessary skill to perform their mission critical task. On the other, just-in-time training would occur for low-likelihood events that are time critical. However, because of the radically new architecture of long space missions, a situation where astronauts would perform initial training on board could be imagined. Training type is thus a more useful variable to consider than mission time line. In initial training, high fidelity simulators are often used. In a situation where initial training is performed on board, however, it is less likely that a high fidelity simulator is present due to design restrictions associated with spacecraft. This necessitates investigating the effects of simulator fidelity on training, which is discussed later in this literature review.

The goal of refresher training is to ensure skill levels are adequately high to perform tasks after a period of non-use. In ground based refresher training, the original simulators can be used. However, on board the same simulator restrictions apply. Gardlin and Sitterley (1972) propose three levels of complexity in on-board refresher training: The lowest would entail nothing more than a verbal or mental review of the task (sometimes referred to as symbolic rehearsal). The median level would be using the real systems in a safe training mode. Finally, Gardlin states: "Beyond this would be the application of more sophisticated combinations of software, computers, and simulation/training hardware to provide high fidelity reproductions of spacecraft system dynamics and the operational visual environment associated with critical mission operations, phases, and maneuvers."

The third form of training is just-in-time training, which is characterized by an unanticipated task which needs to be trained in a timely manner, for instance by making use of previously learned generalizable skills. An example of such a task is a failure of a critical system which requires an extra vehicular activity to repair. The critical nature of the failure would allow for only a short period (a day, for instance) to train. Barshi and Dempsey (2016) state that the ability to design just-in-time training requires expecting the unexpected. They continue: "Because not all such events can be anticipated in advance, methods for the crew to develop their own training for such occasions must be developed for cases when communication delays prevent the up-link of such training from the ground." Apart from this being a demanding task on its own, it would have to be executed under time pressure. Thus, from this point of view, a low-fidelity simulation is preferred. Only by allowing the crew such flexibility, could the highest autonomy be reached. Caldwell and Onken (2011) define five levels of autonomy, the highest one being "goal determination". In long duration space missions, the need for this high level of crew autonomy is especially pertinent.

2.2 Modeling Training

Kim et al. (2013) state: "To better address the issues related to learning and long-term retention, it is necessary to predict the learners future cognitive states." In order to do so, it is necessary to construct a cognitive model of the human. This section discusses a common taxonomy: Rasmussen's S-R-K [Rasmussen (1983)]. The manner in which the definitions are used, deviates slightly from the original intention by Rasmussen. Rather than identifying different levels of how one operator might execute a certain task based on skill progression, here the definitions are used to point to a division in task types. This ties directly into how the operator internally processes these tasks.

Several authors have proposed models of representing learning in humans. Kim et al. (2013) provide an excellent overview in their work. Fitts and Melton (1964), for instance, proposed three levels in the learning process. This is built upon by Anderson (1982), who also proposed a three-level taxonomy for cognitive skill development. Rasmussen (1983) formulated the Skills, Rules and Knowledge model. All these models have the division of three stages in common. Firstly, "acquiring declarative and procedural knowledge", secondly, "consolidating the acquired knowledge" and thirdly, "tuning the knowledge towards overlearning" [Kim et al. (2013)]. As a subject is learning a new skill, he or she progresses through these stages. The initial stage of learning involves generating the first mental processes in executing a task. Afterwards, these processes are interconnected in order to form a memory of executing the task. When this is completed, a stage called overlearning is commenced, where the connections harden and muscle memory continues to strengthen.

Focusing on Rasmussen's S-R-K, the first level is skill-based behavior, which comprises of behavior where a task is performed without exerting laborious mental effort [Rasmussen (1983)]. In an experimental research environment, a tracking task is a typical skill-based task. In real life, an example of a tracking task is to follow another aircraft in a formation flight, or following a flight director. The second level is rule-based behavior, where the essence is the presence of a stored rule which acts as a "feedforward control" in achieving a certain goal. The rules may have been formed by the operator himself, but may also have been communicated in a different manner. The separation between rule-based and skill-based tasks is not always distinct, as a person might move from applying rules to learning the corresponding skill. The third and final level is called knowledge-based behavior. Here, Rasmussen argues that performance is fully goal controlled. The goal is formally known and a mental model is used to attain the goal. This level represents new situations where an operator has to exert mental effort to achieve a solution. The taxonomy will be used later in the discussion on the effects of simulator fidelity on retention and generalizablity.

3. SKILL DECAY

3.1 Relevant Variables

There is some notion as to why skills decay. However, the question of how they decay remains pertinent. Several studies have been performed to identify the variables which might influence the retention of skill over time. For instance, Schendel et al. (1978) suggest the following variables: “degree of proficiency attained by the learner during initial training; amount and kind of refresher training; transfer of skills from one task to another; interfering activities; scheduling of practice during training; use of part-task versus whole-task training methods, and; introduction of extra test trials prior to final testing”. In a meta-analysis by Arthur, Jr. et al. (1998) all variables influencing skill decay were divided into two categories: methodological or task-related. “Here, the task-related variables are inherent characteristics of the task and are not amenable to modification by the researcher”, as stated by Villado et al. (2013), which is one of the studies put forward in Arthur, Jr. et al. (1998). The meta-analysis used 53 articles.

The first objective was to find the correlation between the length of the retention interval and the amount of skill decay. A positive correlation was found; longer intervals result in more decay. This seems like a rather trivial result. However, the strength of the decay varied over different studies, which points to other influencing variables. One of those influencing variables is the degree of overlearning, although its precise effects are not fully clear. One of the studies used in the meta analysis concluded that “overlearning does not prevent, but only somewhat reduces, decrements of performance with time” [Hammerton (1963)]. Arthur, Jr. et al. (1998) concluded that although some correlation indeed exists, the current literature only supports a limited range of overlearning and consequently, only a limited effect is observed. Overlearning thus turns out to be a variable of which the exact effects remain under discussion.

Apart from initial training and time, task characteristics also play a role. The meta-analysis hypothesized that open-loop tasks would display less decay than closed-loop tasks, which is a result that follows from other sources on skill decay [Farr (1986)]. However, the results of the meta-analysis indicate a finding that is in contrast with this hypothesis: closed-loop tasks are retained better. This might be due to some contamination effects over the different studies used however.

Another variable is the nature of the task; physical or cognitive tasks. The results show that physical tasks show less decay than cognitive tasks. Driskell et al. (1994) argue that this might be due to the fact that even though mental action allows for the construction of words and images to aid recall, it does not provide any direct feedback on performance. Physical tasks are more likely to provide this feedback in the form of visual or tactile knowledge of the results and thus, a higher performance is easier to attain.

The meta-analysis hypothesized that natural tasks would display less decay than artificial tasks, which would favor using as realistic tasks as possible in training. The results marginally favoured this statement, although the difference was small. The difference could be explained by varying levels of motivation for learning the tasks. In general, participants express more motivation to master tasks that appear to be natural [Arthur, Jr. et al. (1998); Stefanidis et al. (2005)]. This conclusion, however, is based on a questionnaire, which inherently features some subjectivity. However, perhaps there is some subconscious truth in the question of motivational differences.

Finally, the meta-analysis found a negative correlation between the original learning and retention environment; if the two differed more, more decay was present. This relates directly to generalizability, where a certain skill must be adapted to fit a new task. However, this conclusion was drawn from only four data points.

3.2 Modelling Skill Decay

The decay of skill is more than simply a function of time [Arthur, Jr. et al. (2013)]. The variables identified in the previous section attest to this. Modelling the decay of skill is a daunting task. Nevertheless, many have tried.

It started with Ebbinghaus (1885), who noticed that a decay curve is most often negatively accelerated: the curve falls most quickly immediately after initial training. McGeoch and Irion (1952) attempted constructing these curves as well: their work focused on reciting words from memory. These are just two examples of research in decay of memory. More examples can be found, such as Wixted and Carpenter (2007); Wixted and Ebbesen (1991); Meeter et al. (2005); Bahrick (1992); Baldwin et al. (1976). Most of these studies suggest that, although it is possible to construct a decay curve for specific types of tasks, there is no such thing as a universal skill decay curve [Naylor and Briggs (1961)]. There are several speculated methods of creating limited skill decay curves; for example, as illustrated by Gronlund and Kimball (2013), who discuss using computational modelling to predict skill decay curves as a function of some of the variables identified in the previous section. However, the authors did not identify any work into the actual construction of these curves using a well-defined systematic approach. Baldwin and Ford (1988) go as far as defining five possible scenarios of “maintenance of training curves”, four of which show possibilities along which skill can decay and a final one which indicates a skill growth after training, because the skill is used extensively. A point to note in their classification is that there is no period of non-use. Rather the term “maintenance” is coined, which indicates that the skills could also be maintained by executing them in operational environments. Furthermore, skill decay is only classified, and no quantifications are proposed in Baldwin and Ford’s literature study.

Farr (1986) stated 11 issues for further research and development in his 1986 report on skill retention. The construction of skill decay curves was one of them, but firstly he stated that it is necessary to define measures to quantitatively describe this phenomenon. Arthur, Jr. et al. (2013) look back on these needs and conclude that “not much empirical research has been devoted to the study of knowledge and skill retention outside the more basic, cognitive-experimental work on memory and motor learning”. Even on these two subgroups, the research and its applicability is limited. Two reasons are put forward: Firstly, experiments in retention of skill are logistically very challenging, as long retention periods are involved. Secondly, the need for understanding skill decay in detail is not present, as it is commonly regarded as “a matter of interference with retrieval processes rather than sheer non-use". The detailed working mechanisms of skill decay are not fully understood, which impedes the issue to gain a relevance in the scientific community.

It thus appears to be the case that in order to fully understand skill decay, there is a need for a clearly structured approach which would shed light on the underlying processes. With this approach, skill decay curves could be developed, which in turn find many applications in practical operations.

4. SIMULATOR FIDELITY

Training simulators come in many different varieties, each placing importance on different aspects of representing the real world, leading to different levels of fidelity. Some fidelities commonly found in literature are physical, psychological, behavioral, and face fidelity. Each of them describes some of the effects the simulator might have on the operator. For a long time, the paradigm in designing simulator training was to facilitate as high fidelity on all aspects as possible. Caird (1996) described this with the statement: "For decades, the naive but persistent theory of fidelity has guided the fit of simulation systems to training." This notion was especially pertinent in the design of training simulators. However, in recent times a shift towards a deeper understanding of the different types of fidelity is observed [Roscoe (1991)]. Some simulators might focus on a specific type of fidelity and hence this choice is not trivial anymore in the design. This section attempts to correlate the type of fidelity a simulator might have on the retention variables identified before.

The knowledge level in S-R-K assumes that a situation is completely new and thus no specific training of it can be performed. Rather, attention must be paid to set up training procedures to allow for high grades of generalizability. As Dahlstrom et al. (2009) put it: "Crews can effectively counter many threats by replicating or slightly varying the technical skills learned during their training." He continues by arguing that to achieve such creativity of solution, a lower fidelity simulator would be beneficial, since it shifts focus from procedural knowledge to the pure skill-based task. This can subsequently be varied to counter unforeseen circumstances. For the middle level of the S-R-K taxonomy, rule-based behavior, there exists evidence that high physical fidelity simulators accommodate retention of the task for at least a year [Boet et al. (2011)]. Also with low fidelity simulation or even without any simulation at all, these kind of tasks can remain on high performance levels. Without simulation, for instance, a technique called "symbolic rehearsal" is suggested by Kluge and Frank (2014) and Kluge et al. (2015). This is a technique where a person internally visualizes performing a skill, without actually doing it. They conclude that symbolic rehearsal effectively supports knowledge retention, but skill retention in a lesser manner. Procedure-based tasks can be practiced and maintained in several ways, using simulators of varying fidelity levels, with little effort over longer-spread periods. O'Hara (1990) proposes 30-minute refresher training every six months. This specific recommendation follows from an experiment with marine cadets; their watch standing skills were tested over a prolonged period of time with no use. After quick refresher training, though, their performance levels could quickly be increased again. Grimsley (1969) found no difference in retention between subjects trained on a low-fidelity simulator and subjects trained on a high-fidelity one. In this research, subjects had to execute a procedural task. It appears that rule-based tasks require relatively low levels of simulator fidelity to maintain.

On the skill-based task level, several studies suggest that practice refresher interventions support skill retention in the best way. The experiments treated in Kluge et al. (2015) used a retention period of two weeks (or one week with an intervention in the middle). Three types of retention interventions were investigated: skill practice, skill testing and symbolic rehearsal. In skill practice, a practice session like in the initial training was repeated. In skill testing, a performance test was used to retain the skill and finally symbolic rehearsal was investigated. The results show that both skill practice and testing support retention. Symbolic rehearsal could not fully prevent skill decay. In another study on the topic, Sauer et al. (2000) were not able to identify an influence between two different training methods (limited task-focused knowledge versus detail process-focused knowledge). A conclusion was drawn that continuous skill-based control tasks showed little skill decay over the 8 month retention period. The control task in this study was a process control task of a life support system of a spacecraft. It thus mainly dealt with low-frequency control.

In Merbah and Meulemans (2011), the step to motor skills is made. There are some suggestions to increase retention and generalizability by performing the training phase in a random order; instead of treating one subject per session, instructors are urged to mix it up. During the training phase participants will perform worse, but the skill is retained for longer and is more generalizable. Roscoe (1991) notes on the retention of motor skills: "Research has shown that innovations in training strategies, in some cases involving intentional departures from reality, can have stronger effects than high simulator fidelity on the resulting quality of pilot performance". Caird (1996) adds to this: "There is some evidence from flight simulation that higher levels of fidelity have little or no effect on skill transfer and reductions in fidelity actually improve training. Reductions of complexity may aid working memory and attention as skills and knowledge are initially acquired."

To conclude this discussion on simulator fidelity and skill decay, it appears that rule and knowledge based tasks can be retained for prolonged periods of times using frequent practice sessions of various nature. Use could be made of low-fidelity simulators or of mental techniques such as symbolic rehearsal, as entailed by Gardlin and Sitterley (1972). For skill-based tasks, however, the story is somewhat more complicated. There is evidence that continuous motor tasks are retained for prolonged periods of time [Casner et al. (2014)]. However, Mulder et al. (2004) stress the need for frequent refreshing training with accurate training simulators. Prophet (1976) warns that many manual control tasks that are executed in flight or in space missions are more complex and therefore feature more decay than the simple motor skills. He supports this statement with the results of several investigations. Firstly, Hammerton (1963) found a significant decrement in control skills after a six month period. In the same fashion, Sitterley and Berge (1972) and Sitterley (1974) found a large decay after a three month retention period, with errors being 2-3 times as large as at the end of training. Sitterley even used a complex spaceflight control task in one of his experiments, the other one being based on visually landing an aircraft. Furthermore, Cotterman and Wood (1967) found similar results in their 1967 study on lunar landing manual control skills. These works show varying results across experiments and tasks. A more thorough investigation into the fidelity for the retention of manual control skills is thus needed.

5. DISCUSSION

Considering the previous findings uncovered with this literature study, several research gaps were identified. Firstly, as became clear in Section 3 there is a need for skill decay functions to be developed using an analytic and structured approach.

Secondly, the link between simulator fidelity and skill decay functions is of interest. Section 4 shows that there is a need for research into fidelity of simulators for manual control tasks. Typical space mission tasks falling into this category are landing or maneuvering a spacecraft, or driving a planetary rover. Multiple sources suggested that this kind of skill is subject to strong decay, more so than simple manual motor tasks [Prophet (1976); Hammerton (1963); Sitterley (1974); Sitterley and Berge (1972); Cotterman and Wood (1967)]. Adding to the problem are sources stating that this kind of skill is best trained in high fidelity simulators [Mulder et al. (2004)]. This is challenging, because of the constrained mass and power budgets commonly found in spacecraft. In this vein, an experiment building on the quantification of the skill decay functions is suggested. By setting up the decay functions for a task executed on different types of simulators, the link between fidelity and retention can be established.

Thirdly, the generalizability of skills is a relevant topic for future research since it directly relates to just-in-time training. To handle unanticipated situations and make the astronaut-machine system more resilient, this is paramount. Initial skills must be sufficiently generalizable to facilitate just-in-time training.

Farr (1986) states: "We need an operational definition of task complexity that will inform us of the memorability of the total task, as well as its components taken individually. We further require a means for deriving an index, preferably a quantitative one, that we can effectively use for twin purposes: (a) for determining the ease of learning the task; and (b) for predicting the decay rate of the task or any of its major components." A cybernetic approach would provide a highly useful tool when conducting research into skill decay functions and generalizability, for it allows to model a human operator in an analytic and structured manner [Pool and Zaal (2016)].

REFERENCES

Anderson, J.R. (1982). Acquisition of cognitive skill. *Psychological Review*, 89(4), 369–406. doi:10.1037/0033-295X.89.4.369.

Arthur, Jr., W., Day, E.A., Bennett, Jr., W., and Portrey, A.M. (2013). *Individual and Team Skill Decay: The Science and Implications for Practice (Applied Psychology Series)*. Routledge.

Arthur, Jr., W., Winston Bennett, J., Stanush, P.L., and McNelly, T.L. (1998). Factors that influence skill decay and retention: A quantitative review and analysis. *Human Performance*, 11(1), 57–101. doi:10.1207/s15327043hup1101_3.

Bahrick, H.P. (1992). Stabilized memory of unrehearsed knowledge. *Journal of Experimental Psychology*, 121(1), 112–113. doi:http://dx.doi.org/10.1037/0096-3445.121.1.112.

Baldwin, R.D., Cliborn, R.E., and Foskett, R.J. (1976). The acquisition and retention of visual aircraft recognition skills. Technical report, U. S. Army Research Institute for the Behavioral and Social Sciences.

Baldwin, T. and Ford, J.K. (1988). Transfer of training: A review and directions for future research. *Personnel Psychology*, 41, 63–106. doi:10.1111/j.1744-6570.1988.tb00632.x.

Barshi, I. and Dempsey, D.L. (2016). Evidence report: Risk of performance errors due to training deficiencies. Technical report, National Aeronautics and Space Administration, Lyndon B. Johnson Space Center, Houston, Texas.

Boet, S., Borges, B.C.R., Naik, V.N., Siu, L.W., Riem, N., Chandra, D., Bould, M.D., and Joo, H.S. (2011). Complex procedural skills are retained for a minimum of 1 yr after a single high-fidelity simulation training session†. *BJA: British Journal of Anaesthesia*, 107(4), 533–539. doi:10.1093/bja/aer160.

Caird, J. (1996). Persistent issues in the application of virtual environment systems to training. In *Proceedings Third Annual Symposium on Human Interaction with Complex Systems. HICS'96*, 124–132. IEEE Comput. Soc. Press. doi:10.1109/HUICS.1996.549502.

Caldwell, B. and Onken, J. (2011). Modeling and analyzing distributed autonomy for spaceflight teams. In *41st International Conference on Environmental Systems*, 1–8. American Institute of Aeronautics and Astronautics, Portland, Oregon. doi:10.2514/6.2011-5135.

Casner, S.M., Geven, R.W., Recker, M.P., and Schooler, J.W. (2014). The retention of manual flying skills in the automated cockpit. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, 56(8), 1506–1516. doi:10.1177/0018720814535628.

Cotterman, T.E. and Wood, M.E. (1967). Retention of simulated lunar landing mission skills: A test of pilot reliability. Technical report, Aerospace Medical Research Laboratories, Wright-Patterson Air Force Base, Ohio. URL http://www.dtic.mil/cgi-bin/GetTRDoc?Location=U2&doc=GetTRDoc.pdf&AD=AD0817232.

Dahlstrom, N., Dekker, S., van Winsen, R., and Nyce, J. (2009). Fidelity and validity of simulator training. *Theoretical Issues in Ergonomics Science*, 10(4), 305–314. doi:10.1080/14639220802368864.

Driskell, J.E., Copper, C., and Moran, A. (1994). Does mental practice enhance performance? *Journal of Applied Psychology*, 79(4), 481–492. doi:10.1037/0021-9010.79.4.481.

Ebbinghaus, H. (1885). *Memory: A Contribution to Experimental Psychology*. Teachers College, Columbia University, New York.

Farr, M.J. (1986). The long-term retention of knowledge and skills: a cognitive and instructional perspective. Technical report, Institute for Defense Analysis.

Fitts, P.M. and Melton, A.W. (1964). Perceptual-motor skill learning. *Categories of human learning*, 47, 381–391.

Gardlin, G.R. and Sitterley, T.E. (1972). Degradation of learned skills. a review and annotated bibliography. Technical report, Boeing Co., Seattle, WA, United States. URL https://ntrs.nasa.gov/archive/nasa/casi.ntrs.nasa.gov/19730001425.pdf.

Grimsley, D.L. (1969). Acquisition, retention, and retraining: Effects of high and low fidelity in training devices. Technical report, George Washington Univ., Alexandria, VA. Human Resources Research Office.

Gronlund, S.D. and Kimball, D.R. (2013). Remembering and forgetting: from the laboratory looking out. Book Chapter in "Individual and Team Skill Decay: The Science and Implications for Practice (2013)".

Hammerton, M. (1963). Retention of learning in a difficult tracking task. *Journal of Experimental Psychology*, 66(1), 108–110. doi:10.1037/h0040296.

Kim, J.W., Ritter, F.E., and Koubek, R.J. (2013). An integrated theory for improved skill acquisition and retention in the three stages of learning. *Theoretical Issues in Ergonomics Science*, 14(1), 22–37. doi:10.1080/1464536X.2011.573008.

Kluge, A. and Frank, B. (2014). Counteracting skill decay: four refresher interventions and their effect on skill and knowledge retention in a simulated process control task. *Ergonomics*, 57(2), 175–190. doi:10.1080/00140139.2013.869357.

Kluge, A., Frank, B., Maafi, S., and Kuzmanovska, A. (2015). Does skill retention benefit from retentivity and symbolic rehearsal? – two studies with a simulated process control task. *Ergonomics*, 59(5), 641–656. doi:10.1080/00140139.2015.1101167.

Mars Architecture Steering Group (2009). Human exploration of mars: Design reference architecture 5.0. Technical report, NASA Headquarters.

McGeoch, J.A. and Irion, A.L. (1952). *The psychology of human learning*. Longmans, Green & Co., 2 edition. URL http://psycnet.apa.org/psycinfo/1952-05377-000.

Meeter, M., Murre, J.M.J., and Janssen, S.M.J. (2005). Remembering the news: Modeling retention data from a study with 14,000 participants. *Memory & Cognition*, 33(5), 793–810. doi:10.3758/BF03193075.

Merbah, S. and Meulemans, T. (2011). Learning a motor skill: Effects of blocked versus random practice a review. *Psychologica Belgica*, 51(1), 15. doi:10.5334/pb-51-1-15.

Mulder, M., Van Paassen, M.M., and Boer, E.R. (2004). Exploring the roles of information in the manual control of vehicular locomotion: From kinematics and dynamics to cybernetics. *Presence: Teleoperators and Virtual Environments*, 13(5), 535–548. doi:10.1162/1054746042545256.

Naylor, J.C. and Briggs, G.E. (1961). *Long-term retention of learned skills: a review of the literature*. Columbus: Ohio State University, Laboratory of Aviation Psychology.

Nicholas, J.M. and Foushee, H.C. (1990). Organization, selection, and training of crews for extended spaceflight - findings from analogs and implications. *Journal of Spacecraft and Rockets*, 27(5), 451–456. doi:10.2514/3.26164.

O'Hara, J.M. (1990). The retention of skills acquired through simulator-based training. *Ergonomics*, 33(9), 1143–1153. doi:10.1080/00140139008925319.

Pool, D.M. and Zaal, P.M.T. (2016). A cybernetic approach to assess the training of manual control skills. *13th IFAC/IFIP/IFORS/IEA Symposium on Analysis, Design, and Evaluation of Human-Machine Systems*. doi:10.1016/j.ifacol.2016.10.588.

Prophet, W.W. (1976). *Long-Term Retention of Flying Skills: A Review of the Literature*. Alexandria VA: Human Resources Research Organization (ADA036077).

Rasmussen, J. (1983). Skills, rules, and knowledge; signals, signs, and symbols, and other distinctions in human performance models. *IEEE Transactions on Systems, Man, and Cybernetics*, SMC-13(3), 257–266. doi:10.1109/tsmc.1983.6313160.

Roscoe, S.N. (1991). Simulator qualification: Just as phony as it can be. *The International Journal of Aviation Psychology*, 1(4), 335–339. doi:10.1207/s15327108ijap0104_6.

Sauer, J., Hockey, G.R.J., and Wastell, D.G. (2000). Effects of training on short- and long-term skill retention in a complex multiple-task environment. *Ergonomics*, 43(12), 2043–2064. doi:10.1080/00140130010000893.

Schendel, J.D., Shields, J.L., and Katz, M.S. (1978). Retention of motor skills: Review. Technical report, U.S. Army - Research Institute for the Behavioral and Social Sciences, Virginia.

Sitterley, T.E. (1974). Degradation of learned skills. static practice effectiveness for visual approach and landing skill retention. Technical report, Boeing Aerospace Co., Seattle, WA, United States. URL https://ntrs.nasa.gov/search.jsp?R=19740024447.

Sitterley, T.E. and Berge, W.A. (1972). Degradation of learned skills. effectiveness of practice methods on simulated space flight skill retention. Technical report, Boeing Co., Seattle, WA, United States. URL https://ntrs.nasa.gov/search.jsp?R=19730001426.

Stefanidis, D., Korndorffer, J.R., Sierra, R., Touchard, C., Dunne, J.B., and Scott, D.J. (2005). Skill retention following proficiency-based laparoscopic simulator training. *Surgery*, 138(2), 165–170. doi:https://doi.org/10.1016/j.surg.2005.06.002.

Villado, A.J., Day, E.A., Arthur, Jr., W., Boatman, P.R., Kowollik, V., Bhapatkar, A., and Bennett, Jr., W. (2013). Complex command-and-control simulation task performance following periods of nonuse. Book Chapter in "Individual and Team Skill Decay: The Science and Implications for Practice (2013)".

Wixted, J.T. and Carpenter, S.K. (2007). The wickelgren power law and the ebbinghaus savings function. *Psychological Science*, 18, 133–134. doi:https://doi.org/10.1111/j.1467-9280.2007.01862.x.

Wixted, J.T. and Ebbesen, E.B. (1991). On the form of forgetting. *Psychological Science*, 2(6), 409–415. doi:10.1111/j.1467-9280.1991.tb00175.x.