Skill Acquisition and Retention in Training:
DSTO Support to the Army Ammunition Study

Christina Stothard and
Robin Nicholson

DSTO-CR-0218

20020405 049

Skill Acquisition and Retention in Training: DSTO Support to the Army Ammunition Study

Christina Stothard and Robin Nicholson

Land Operations Division
Electronics and Surveillance Research Laboratory

DSTO-CR-0218

ABSTRACT

This work was undertaken within the Land Preparedness Studies Task (ARM 01/059) within the Land Operations Division. It is in response to the request for DSTO support contained within the Terms of Reference for the Army Ammunition Study (AAS). This request was for DSTO to assist in developing a better understanding of skill degradation (retention) and acquisition, with the view that this would enable accurate prediction of training requirements and, in the longer term, develop tools to predict levels of proficiency provided by different training regimes.

In summary we:

- Took a systems approach to the training problem and assessed the impact of the treatment of skill acquisition and degradation of predictions of training frequency requirements.
- Undertook a literature survey, concentrating on the Cognitive Psychological literature, to ascertain the current thinking on how people learn and forget.
- Assessed the gaps in the literature and scoped the work needed to address these gaps.
- Made some progress in developing a new approach to modelling retention.
- Assessed how DSTO, or other S&T agencies, could further address the issues raised by the AAS in both the short and long terms, through the development of appropriate R&D programmes.

RELEASE LIMITATION

Approved for public release

AQ F02-06-1078

Skill Acquisition and Retention in Training: DSTO Support to the Army Ammunition Study

Executive Summary

This work is aimed at developing a better understanding of skill acquisition and degradation (retention), to enable accurate prediction of training requirements. It also provides a start for the development of tools to predict levels of proficiency provided by different training regimes.

The representation of skill acquisition and retention, within training/resource models, is shown to greatly influence model outputs. Therefore, at the heart of any modelling effort is an understanding of how proficiency is achieved and lost (skill acquisition and retention).

We review the cognitive psychology literature on learning and retention, in order to establish what is the current understanding of skill acquisition and retention. The review covers two broad areas. First, the factors affecting skill acquisition and retention, and second, the mathematical description of the rate of skill acquisition and retention.

The rate of skill retention, within the field of psychology, is currently accepted to be best fit by a power function. However, there are several significant difficulties encountered when using the power function in a dynamic model. We suggest an alternative model of retention, using a System Dynamic framework, which is suitable for dynamic modelling. This model requires more development and validation to assess its robustness.

Gaps found in the literature, in areas pertinent to the understanding skill retention and acquisition in the Army context, include relearning and the change in skill retention from novice to expert, and moving from individual to collective skills. The System Dynamics model potentially addresses the relearning issue. Of interest is the effect of initial over-training on long term retention, and subsequent time to train to competency. There may be a significant advantage in initially training individuals beyond their required level of proficiency to reduce future training requirements. The area of collective training requires R&D attention and we propose some initial work.

The many factors affecting skill retention, including task complexity, degree of over-learning and time since last practice, are discussed. A method of assessing which factors have the greatest impact is presented in section 7. If these can be identified then more work can be carried out on designing training programmes that lever off this knowledge.

The work we have carried out so far provides qualitative support for the judgements made by Army trainers. To extend this to quantitative support requires more work, in the development of experiments and measures for proficiency, and models that predict proficiency of certain skills. The aim is to augment the qualitative and subjective measures of the Army Subject Matter Experts with quantitative, objective measures. As well as providing Army training planners with information on how individuals (and teams) learn and forget, this will allow the effective planning of training regimes.

Contents

Glossary

AAS	Army Ammunition Study
ARI	Army Research Institute
DPRM	Defence Preparedness Resource Model
DSTO	Defence Science and Technology Organisation
EEG	Electroencephalogram
LOD	Land Operations Division
MAP	Maximum Allowed Proficiency
MRP	Minimum Required Proficiency
R&D	Research and Development
RAR	Royal Australian Regiment
S&T	Science and Technology
SD	System Dynamic
SME	Subject matter Expert
STM	Simple Training Model
UDA	User Decision Aid
WTSS	Weapons Training Simulation System

1. Introduction

This work was undertaken within the Land Preparedness Studies Task (ARM 01/059) within the Land Operations Division. It is in response to the request for Defence Science and Technology Organisation (DSTO) support contained within the Terms of Reference for the Army Ammunition Study[1] (AAS). This request was for DSTO to assist in developing a better understanding of skill degradation (retention) and acquisition, with the view that this would enable more accurate prediction of training requirements and in the longer term the development of tools to predict levels of proficiency provided by different training regimes.

In summary we:
- Took a systems approach to the training problem and assessed the impact of the treatment of skill acquisition and degradation of predictions of training frequency requirements.
- Undertook a literature survey, concentrating on the Cognitive Psychological literature, to ascertain the current thinking on how people learn and forget.
- Assessed the gaps in the literature and scoped the work needed to address these gaps.
- Assessed how DSTO could address the issues raised by the AAS in both the short and long terms, through the development of appropriate Research and Development (R&D) programmes.

In the following sections, we present the results of a sensitivity analysis. This demonstrates the importance of an accurate understanding of skill acquisition and retention, when modelling the relationship between proficiency required and the consequent training resources required.

We then present the results of a literature survey. This survey was undertaken to establish the current understanding of skill retention in the psychological sphere.

Throughout we will discuss the gaps in the literature in areas pertinent to understanding the problem of skill retention and acquisition in the Army context.

Finally, we will discuss future research directions that may be taken by DSTO, in collaboration with other R&D agencies, in order to augment the Army's current understanding of skill retention and acquisition, with the aim of enhancing the current training processes.

---

[1] 98-32726-1 DCA OUT/2001/

2. Importance of Skill Retention and Acquisition

2.1 Background

The initial problem we were looking at, within the Force Preparedness Analysis task (ARM 99/088), was Readiness. Upon looking at previous attempts at modelling readiness, using System Dynamics and attaching resources to it (e.g. the Defence Preparedness Resource Model (DPRM)), and attempting a small scale problem ourselves (4th Royal Australian Regiment – 4 RAR, Ammo Usage Model), it was realised that the heart of any modelling effort was an understanding of how proficiency is achieved and lost (skill acquisition and retention). In short, the results of these models were very sensitive to our representation of skill acquisition and retention.

2.2 Sensitivity Analysis

To make clear the sensitivity of the outputs of readiness/training models to the way in which skill acquisition and retention are represented, we have constructed a simple training model (STM) that simulates the treatment of skill acquisition and retention in the above-mentioned DPRM and 4 RAR models. The influence diagram for the STM is given below in Figure 2-1.
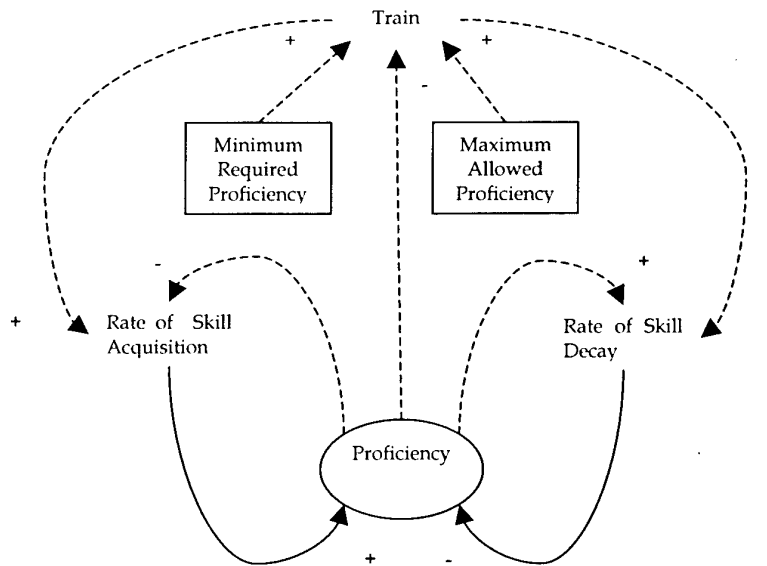
Figure 2-1 : Influence Diagram representation of Simple Training Model

The simple premise here is that training increases proficiency, and in the absence of training, proficiency decays. The model is set to automatically instigate training when the level of proficiency drops below the Minimum Required Proficiency (MRP), and to cease training when the proficiency rises above the Maximum Allowed Proficiency (MAP). There is also an assumed (feedback) relationship between Proficiency and the rates of Skill Acquisition and Skill Decay. It is the form of these relationships that our initial study aims at uncovering. However, to justify investing time in such a study we first investigated the sensitivity of the above simple model to the form of these relationships. The initial results of this sensitivity analysis are presented below.

The simple experiment we used was to look at the number of simulated-days-training required to maintain different levels of proficiency, with varied forms of Skill Acquisition and Decay Rates. At each level of required-proficiency, defined by the parameters MRP and MAP, the model was run and the percentage of model runtime spent training was recorded. We then varied the representation of the Acquisition and Decay Rates and repeated the runs over required-proficiency. Figure 2-2 below is an example of the output of Proficiency from one of the model runs. What we see is Proficiency varying between the MRP and the MAP parameters. The sections of the chart where proficiency is increasing represents when training is occurring, and where proficiency is decreasing training is not occurring.
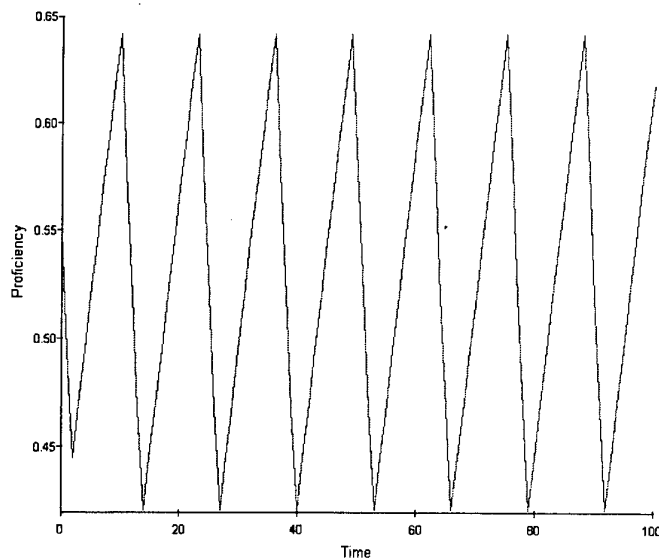
Figure 2-2: Example of results of run with MRP=0.45 and MAP=0.65

The first example (the linear case) represented by Figure 2-3 to 2-5 below, shows the resource usage in terms of percentage time spent training, given a sigmoidal learning

3

curve (rate of skill acquisition) and a linear forgetting curve (rate of skill retention). Figure 2-6 to 2-8 shows the resource usage where the only change is an exponential decay forgetting curve.
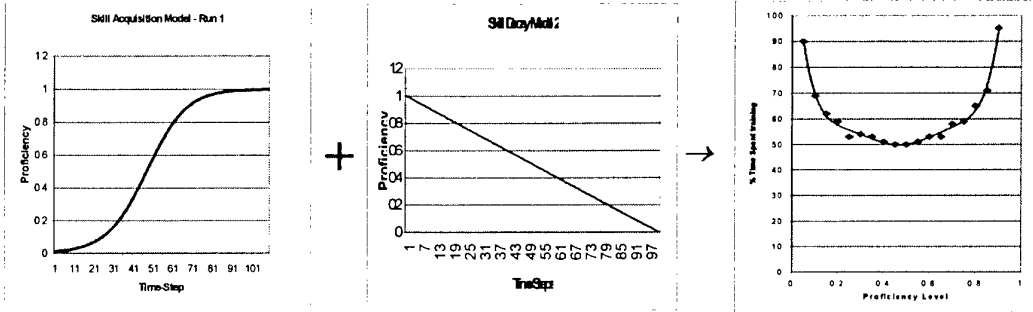
Figure 2-3:Learning Curve 1   Figure 2-4:Decay Curve 1   Figure 2-5: Model Results
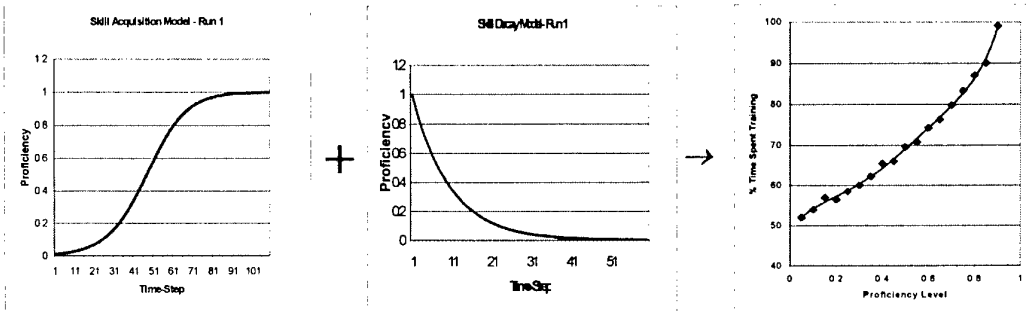
Figure 2-6:Learning Curve 1   Figure 2-7:Decay Curve 1   Figure 2-8: Model Results

The result in case one suggests that if skill-level drops off linearly and is acquired in a manner described by Figure 2-3, then the most efficient training occurs where proficiency is maintained at around 0.5, at the low point in Figure 2-5. It also suggests that maintaining skills at a low level is as expensive as maintaining them at a high level. This unlikely result probably points to the fact that the linear drop off of skills is a poor model.

Upon changing the decay of skills (skill retention curve) to an exponential, we see a dramatic change in the resources required to maintain proficiency. The percentage of time spent training increases with proficiency, in a non-linear manner. That is, doubling the required proficiency more than doubles the required time spent training. Note these results are only to demonstrate a point and are not based on real retention curves.

The point, that is clearly demonstrated, is the sensitivity of the resource usage, Figures 2-5 and 2-8, in terms of required percentage time spent training, to changes in the treatment of skill retention. Hence, to move to a position where the modelling effort is useful, and to a point where we can optimise training schedules, to minimise cost and maximise

effectiveness, we need to have an accurate understanding of skill retention and acquisition.

The model used here is very basic and the forms of retention and acquisition are hypothetical, and simply show how different assumptions affect the output of the model. What is clearly illustrated is that the only way forward is a more detailed study in the domain of experimental psychology.

In the next section we present the findings of the literature survey we undertook, with the aim of establishing the current position of the thinking on skill retention and acquisition, within the social sciences.

3. Literature Survey

3.1 Summary

Upon carrying out a literature survey, we found strong support for a power function representation of forgetting. Figure 3-1, below, shows a generic power curve with the form $\alpha t^{-\beta}$.

Power Function- Skill Retention
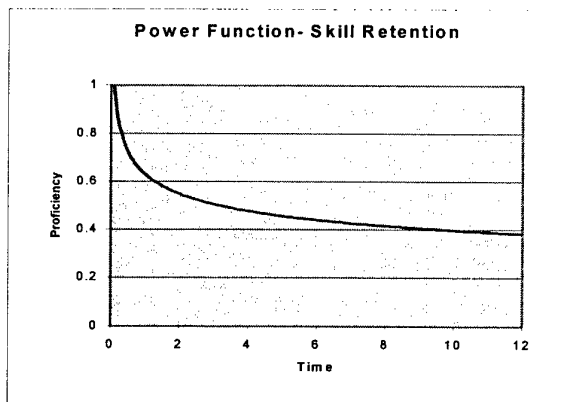
Figure 3-1: Example of Power Function for Skill Retention

It was found (Rubin and Wenzel, 1996; Wicklegren,1972; Wixted & Ebbesen, 1991) that these types of curves best fitted experimental data. Similarly for learning the data was best fitted by functions of the form $1-\alpha t^{-\beta}$ (Anderson, 1991) (see Figure 3-2 below). Basically, learning is characterised by an initial steep learning period that asymptotes to a maximum proficiency. Forgetting is characterised by an initial steep drop in proficiency, which then levels off, dropping more slowly as time goes on.

Power Function - Skill Acquisition
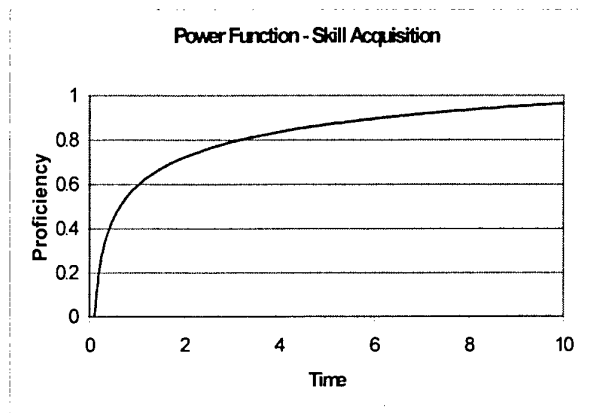
Figure 3-2: Example of Power Function for Learning

These finding are in qualitative agreement with the intuitions of experienced Army instructors [Stothard & Nicholson, 2001]. Where, the largest decay in skill occurs in the first few months/weeks after training, with smaller differences as time goes on. For example, instructors reported that they believed the drop in proficiency, for a variety of skills, was greater between time zero and two months, than between the sixth and eighth month.

It should be noted at this point that one major gap in the literature identified by the survey, is the effect of relearning. That is, most of the studies concentrate on one isolated learning event, rather than one skill learned and then re-learned. The small amount of work that has been done in this area suggests that relearning changes both subsequent learning and forgetting. It may well be that the power function does not describe the Skill Retention of an expert. This area requires more study. Early work by DSTO in new model development, discussed later, shows some promise in addressing relearning. However, such a model would require testing by experiment. It is also worth mentioning at this point that there are some theoretical concerns with the power law description that complicate its use in a simulation such as the STM above. These will be discussed in more detail in section 6.3.

The basis of an U.S. Army approach for estimating/predicting the skill retention of various tasks will also be presented.

3.2 Introduction to Literature Survey

The rest of this section of the report aims to establish the psychological and behavioural principles that underlie training, in more detail. In particular, we will map out the factors found to influence skill acquisition and retention; reviewing and critiquing the approaches described in the literature. The utility and drawbacks of the currently available models of

skill learning and retention are discussed and some directions for future research and development are presented.

While the principles of learning, and relearning have been investigated, from random word lists to military procedural tasks, it is important to keep the overall perspective in mind: that any framework for examining skill acquisition and retention needs to be able to take skill aggregation into account, from the simple to the complex and the individual to the team. For example, shooting is a fundamental skill for any infantry soldier, however, in itself, shooting alone does not make a good infantryman. For a methodology to be of use in optimising training programmes, it should be applicable to more than a simple and isolated individual skill.

3.3 Definitions of Skill

Skilled performance encompasses a huge range of human activities (Allard & Starkes, 1991): sports, recreational and military. In this report, we discuss both cognitive and motor skills. Essentially, the aim of the report is to establish a method to evaluate skill retention that can be applied to any real-world task in the military. The literature discusses both cognitive and motor skills, which may (or may not) generalise directly to an applied activity. A critical question for understanding motor expertise is the link between knowledge (cognitive skill) and performance (motor) skills.

Wisher, Sabol & Ellis (1999) categorise military tasks into three components: knowledge, decision and execution. These categories are based on different mechanisms; "knowledge" is based on recall of domain specific information, "decision" depends on cognitive processing of the domain-specific information, and "execution" refers to the perceptual motor requirements of a task, for example, target acquisition and tracking.

4. Skill Acquisition

4.1 Skill learning: from doing to thinking

Essentially, skill acquisition is based on learning, and skill decay (or retention) is based on the principle of forgetting. There are other factors that come into play when we learn skills (for example practice, aptitude and task difficulty) but to examine skill acquisition, it is necessary to understand how we learn. There are three broad stages of skill development (or skill learning). Sufficient learning leads to expertise. An expert organises their knowledge, and skills, in qualitatively different ways compared with novices.

4.1.1 Three stages of skill acquisition

It is typical to distinguish three stages of skill development (Anderson, 1991; Fitts & Psoner, 1967): first, the cognitive stage, when a learner commits a set of facts to memory by rehearsal; second stage, the associative stage, when links are made between facts and the third and final stage; the autonomous stage, when the links become smooth and continuous.

4.1.1.1 First phase: Cognitive learning (thinking)

The information on what to do next is available, however the execution is very slow. In this cognitive stage, the learner has to explicitly retrieve specific facts and interpret them to address the next stage. For example, when learning to drive a manual car, Anderson (1991) memorised the location of the gears, (e.g. "up, left") and the correct sequence of engaging the clutch and moving the stick shift. He rehearsed this information as he performed the skill. He had to explicitly recall the next step of the procedure while performing one step.

This phase relies on Wisher et al's (1999) definition of "knowledge", that is, domain specific information. In order to learn, it is necessary to have the basic facts to build upon.

4.1.1.2 Second phase: Associative learning (linking)

The second stage of skill acquisition is called the associative stage. In this stage, errors in the initial understanding are gradually detected and eliminated and connections between the stages of the skill are strengthened. So Anderson (1991) found that he could coordinate the release of the clutch in first gear with the application of gas, to prevent the car stalling. Also, he didn't have to wait to recall what to do next, shifting through his explicit memories. The process of the skill has been learnt. He learnt the specific procedures for each scenario in learning the skill, for example:

"*If* the goal is to reverse,

Then the sub goals are:

    A. To disengage clutch
    B. Then move gear stick to the upper left
    C. Then to engage the clutch
    D. Then to push down on the gas." (Anderson, 1991, pg 274)

4.1.1.3 Third phase: Autonomous phase (doing)

The third stage in the procedure (the autonomous stage) involves skills becoming increasingly automatic and rapid. The learner needs less conscious attention to perform the skill, allowing for other activities to be conducted at the same time. For example, now Anderson (1991) can drive using a manual car while carrying on a conversation and

monitoring traffic. He no longer had to think about the gear changes; rather, the process had become automatic, and incorporated into the overall process of driving. The markers of the autonomous stage of learning are speed and accuracy. As the skill becomes more practiced, it becomes more accurate and faster, which is known as "tuning"(Anderson, 1982).

The categories of "execution" and "decision" defined by Wisher, Sabol & Ellis (1999) belong in this phase. In order to learn and use a skill, or information, a learner needs to understand how it occurs and when to apply the skill. These only come with practice and experience; as the learner becomes an expert, they can apply their understanding to other situations, while also evaluating its appropriateness.

There is evidence that the stages of learning, from cognitive to autonomous, have clear physiological markers, for example electroencephalogram (EEG) traces (Clarke et al, 2001). This is one area that shows promise in providing objective and quantitative measures of proficiency.

4.1.2 Expertise

Experts are different from novices, obviously. There are significant differences in how novices and experts organise their knowledge of the domain, how they access the information and how they interpret what is happening in the domain. Experts are an example of the very proficient, competent level of learning.

The nature of expertise can be characterised, in part, by the autonomous stage of learning described above. A skill becomes automatic, faster, more accurate and requires less cognitive effort to execute[2]. Other changes also characterise development of expertise. Tactical learning occurs, which refers to generalising a learnt skill to a new problem. Therefore parts of a skill can be drawn upon and applied to different problems, or can be used to build upon when learning a similar skill. Strategic learning involves organising problems that are optimally suited to the particular domain; an expert organises and represents problems to allow more effective solutions. Experts have a broader and deeper understanding of a domain than the beginner. They recognise solutions, and how best to apply these solutions in any given circumstances.

Experts have an increased ability to recognise chunks in patterns, which are patterns of elements that repeat over problems (Anderson, 1991). So experts recognise situations that can be addressed using parts of solutions from other experiences (Hutton & Klein, 1998; Klein, 1998). Experts, in time pressured, dangerous, complex environments, use pattern recognition strategies to solve problems. Due to extreme time pressure experts tend to draw on the first workable solution available to them rather than generating a group of options and weighing up the benefits and costs before making an optimal decision. (Klien,

---

[2] It is this change in cognitive activity that can be measured by EEG; as people master tasks their EEG patterns change, with less activity measured in areas of the brain that are associated with higher functions.

1998) Therefore a cognitive tactic used by experts in these situations is to apply "what if" questions to the candidate solution. This projects the possible action into the future, and allows judgements to be made about its appropriateness based on this. This allows solutions to be developed within the time frame needed.

4.2 Rate of learning: "the power law"

The rate of learning is dependent on many different factors, however, everyone must pass through each of the above three stages to gain proficiency. There is evidence that the rate of learning follows a power function, (see Figure 3-2) the "power law of learning" (Anderson, 1991). Many studies using many different tasks have found that the overall shape of learning is best fitted by a power function (Newell & Rosenbloom, 1981; Logan, 1988; Neves & Anderson, 1981). As discussed above, this generally means a steep learning phase, which plateaus towards a maximum proficiency.

There is very little on relearning, however the work that has been carried out suggests that the rate of learning changes as skills are re-learned (Anderson et al, 1999). This is a characteristic that any sophisticated model of learning needs to account for.

5. Skill retention

The bulk of the training literature is concerned with learning and training, rather than directly discussing skill retention. Training literature includes looking at improving the delivery of training and maximising skill acquisition, Lance et al (1998) discovered. For example, learning and learning strategies (Ackerman, 1987; Ackerman & Humphries, 1991; Gallagher, 1994; Hintzman, 1990; Weiss, 1991), training program design, training techniques and training evaluation methodologies (Goldstein, 1991; Kragier, Ford & Salas, 1993; Snow & Swanston, 1992; Tannenbaum & Yukl, 1992) and the transfer of training (Baldwin & Ford, 1988; Michalak, 1981; Royer, 1979; Schmidt & Bjork, 1992).

There are two broad areas of study into retention. The first is examining the factors that moderate retention and is based solidly in the applied social sciences and aims to improve performance (and therefore retention). The second area looks at the rate of learning and retention, and aims mathematically describe learning and retention.

When discussing retention, it is important to note that there are two ways of measuring retention: recall and recognition. Recognition relies on cues being present, for example, answering a multiple choice question (cues available), compared with recall which would require a single written answer (no cue available). So if the task involved remembering a set of twenty random words, recognition would rely on ticking words recognised in another set of forty random words while recall would require listing all twenty words. It has been recognised that these two processes appear to be independent of each other

(Flexser & Tulving, 1978) and that recall is consistently poorer when compared to recognition (Semb & Ellis, 1994). Therefore any task that requires recall, rather than recognising the next step, will be forgotten much easier.

5.1 Factors affecting skill retention

The following section summarises the factors affecting skill retention and presents some examples of retention effects for a range of skills. The identification of such factors may be used to predict/estimate skill retention for particular skills. One attempt to do this is represented by the U.S. Army Research Institute's (ARI) series of studies that produced the User Decision Aid (UDA) tool (Rose et al, 1985). A promising area for future research is to develop and update this approach to one that utilises modern data-mining techniques. This may provide tools that allow the economical and accurate estimation of skill retention and acquisition for a variety of tasks.

The factors thought to affect retention have been categorised into four broad areas (Bryant and Angel, 2000):

1. Task
2. Training
3. Retention interval
4. Individual

5.1.1 Task factors

5.1.1.1 Cognitive tasks[3]

There are conflicting results when considering the rate of retention of cognitive skills. Some studies have found that cognitive skills are less prone to decay (Arthur et al, 1998) however this contrasts to Driskell et al (1992) who found that cognitive skills deteriorate quicker than motor skills. Wisher et al (1999, pg 41) stated that cognitive skills "tend to be stable for long periods over time however people do exhibit forgetting".

Sabol and Wisher (2001) reviewed retention of military decision skills (cognition based) and found that there was a moderate rate of decay (Cooke, Durso & Schvaneldt, 1994). One study examined the application of oceanography principles immediately, and after four weeks interval, and found a 21% drop in scores (Wetzel, Konoske & Montague, 1983). Another study found a 16% loss of basic problem solving after 8 weeks (Austin & Gilbert, 1973).

---

[3] Cognitive tasks are those that do not have a significant motor skill dimension. They include tasks such as, decision making, planning, and memory tests.

5.1.1.2 Procedural tasks[4]

Hurlock and Montague (1982) reported better retention of continuous tasks. Tasks that have a meaningful organisation or coherence of steps tend to be remembered better (Hurlock & Montague, 1982). A well-organised task may include cues for the next step, allowing for recognition of the next step. Shields, Goldberg and Dressel (1979, cited in Hagman & Rose, 1983) found that soldiers tended to forget the steps in a procedure that were not cued by the previous step, for example, forgetting safety steps not intrinsic to the process. Another study has found this, with uncued steps at the beginning and end of a process as well as those addressing safety and those judged to be "difficult", being the least likely to be recalled (Osborne, et al 1979, cited in Hagman & Rose, 1983).

While complex, procedural tasks have been found in general to be more fragile, this is not necessarily so. The importance of intrinsic cues, in overcoming this problem, is illustrated by Shields, Goldberg and Dressel (1979) and Osborne, et al (1979, cited in Hagman & Rose, 1983). Healy et al (1992) reviewed studies that found both good and bad retention of procedural skills. They put forward the proposal of *procedural reinstatement*. For example, "if training employed a job aid or a checklist to aid learners in sequencing steps that aid will be an important cue needed to reinstate the skill at a later time" (Bryant & Angel, 2000, pg 29).

Procedural reinstatement (Healy et al, 1992) contributes to the recall of complex tasks. Marmie and Healy (1995) examined learning and retention of tank gunner skills, using an US Army simulator. They found that trainees retained significant amounts of skill using the simulator, after up to 22 months. This suggests that there is not a direct relationship between task complexity and retention; the task may be complex, but if it provides sufficient cues and was well learnt initially (i.e. procedural reinstatement), then retention improves.

Task factors affecting retention have been characterised in an extensive series of studies by the US Army Research Institute for Behavioural and Social Sciences (ARI) (Wisher et al, 1999), as:

1. Complexity (e.g. number of steps),
2. Significance of ordering (e.g. does the order of steps matter),
3. The nature of steps (e.g. do the steps require cognitive or motor skills),
4. Feedback (presence of feedback)
5. Presence of job or memory aids

These factors were explored by several projects of the ARI (Wisher et al 1999). These studies found that one of the major predictors of skill loss is task complexity. This is made up of three parts: number of steps in a task, whether the steps must be performed in a set sequence and whether there is any built in feedback that indicates the correct performance of steps. Other findings included retention decreasing as the number of steps increases,

---

[4] Procedural tasks are those that involve number of coherent steps that may include any combination of cognitive and motor skills.

with the ordering of the task modifying this effect, and a single sequence for a task makes remembering easier, and presents clearer cues. Memory or job aides also help perform and retain a skill while time limits can constrain the recall of a skill. Aspects of a recalled skill, that may provide metrics, are the slowing of performance and reduction in accuracy (Arthur et al 1998).

These findings are in qualitative agreement with the opinions of Army Subject Matter Experts (SME) related in informal discussions. For example, the SME reported making judgements such as "they can shoot but they can't shoot safely". These judgements are supported by the findings that un-cued steps, such as safety procedures, were easily forgotten.

5.1.2 Training factors

5.1.2.1 Overtraining

One of the most important factors determining retention is the initial level of learning. An individual's level of initial proficiency has a direct relationship with the level of skill retention. The literature uses the concept of over learning, which is defined as learning beyond one successful performance of the skill (Schendel & Hagman, 1980, cited in Hagman & Rose, 1983). Basically, the trainee continues exercises beyond a single complete routine. This continued training is thought to encourage automacity or proceduralisation (storing the skill into the long term memory) therefore reducing cognitive demands and allowing better long term functioning (Anderson, 1983). By definition, over learning means more practice, and so provides more opportunity to assimilate feedback (Arthur et al, 1998; Lance et al, 1998). So the more automatic the skill has become, the greater the retention.

Driskell et al (1992) conducted a meta-analysis of the over learning and overtraining literature. There was a threshold for the benefits of over learning, that is, to demonstrate improvement at least 50% overtraining was needed. At least half of the number of trials needed to reach criterion were provided again, for the trainees to show benefits, if 10 trials were needed to pass the test, then the trainees would go through another five trials after the successful test. There was a "dose response" with an increase in retention with more overtraining. For example, with 100% and 150% overtraining there were respectively larger increases in retention.

5.1.2.2 Method of training

The content and quality of training affects learning. Hagman and Rose (1983) reviewed the early literature and found that retention was enhanced by introducing testing during training, and by spacing the training sessions. Studies (Hagman 1980a, 1980b, cited in Hagman & Rose, 1983; Hurlock & Montague, cited in Bryant & Angel, 2000) found that trainees who were tested during training performed better over time, compared to those who received presentations only.

Spacing repetitions and practice sessions have been found to enhance retention, provided the interval is not too long (Schendel and Hagman, 1991, cited in Bryant & Angel, 2000). This was also found by Shute and Gawlick (1995), who contrasted abbreviated and extended practice sessions and found that performance was better, over time, for those taking part in the abbreviated sessions. Subjects who had the mixed sessions (both abbreviated and extended practices) performed the best on tests for retention. Shute and Gawlick (1995) found that those in the abbreviated sessions were more self-directed, having accessed the self-help facility on the simulation more frequently. The use of the self-help was thought to produce greater retention by involving the subjects in more meta-cognition, that is, understanding what they did and did not know. This is thought to lead to greater cognitive effort, greater understanding and therefore, better retention.

Extra training enhances retention regardless of whether it is during initial training or conducted as a refresher course afterwards (Schendel & Hagman, 1990, cited in Hagman & Rose, 1993). Other research suggests that refresher training is beneficial as long as over learning is achieved (Driskell et al 1992).

5.1.2.3 Feedback

Feedback affects the quality of learning, which, in turn, affects retention. Feedback can vary in the type and quality. We learn better with feedback, or "knowledge of results" (Hurlock & Montague, 1982). Feedback allows the trainee to correct errors, observe and use cues associated with task performance and generate effective procedures. Schmidt (1997) commented that feedback could have opposing effects on performance. Feedback can aid learning, leading to better retention but it can also build up a reliance on the feedback. This reliance can reduce performance once the feedback is not available. The aim of training should be to provide sufficient feedback to improve performance without producing dependency.

5.1.2.4 Transfer of training

Skill retention can be thought of as the transfer of skills from training to test. Skill transfer is taken to refer to skills applied to tasks and situations not explicitly anticipated during training (Bryant & Angel, 2000). Any trained skill can be applied in situations that differ from the training environment, so we can generalise using the skill beyond the original training environment. Training aims to maximise performance from practice to the work environment (Kraiger et al, 1993, cited in Bryant & Angel, 2000). There is a variation in people's ability to extrapolate and use their skills beyond the training environment.

The encoding specificity principle states that cues for retrieval will be effective if and only if encoded at the time of learning (Tulving, 1983). This can be applied to the retention of skills; perceptual and cognitive cues are needed to retrieve and perform learned skills (Bryant & Angel, 2000). If such cues are not present during recall, then performance will suffer. This is reflected in the increased retention of procedural tasks, with better retention of tasks with intrinsic cues and specific ordering (Hagman & Rose, 1983). Essentially, the

training environment needs to resemble the operational environment, if skills are to be transferred to the new situation.

5.1.3 Retention interval factors

5.1.3.1 Time interval

One of the main factors, whether direct or an intervening variable, is the time interval between training and performance. So the longer the time between practice and performance, the greater the skill loss. Studies have consistently found skill loss over time; performance decreases rapidly soon after training then occurring at an increasingly slower rate (Arthur et al, 1998; Wixted & Ebbesen, 1991). This pattern appears to be consistent across a variety of skills and tasks, humans and animals (Driskell et al. 1992; Wixted & Ebbesen, 1991).

5.1.3.2 Opportunity to practice

The opportunity to practice the skill would, obviously, reduce the rate of skill loss over time. Practice, not necessarily specific training for the skill, includes cognitive rehearsal (Farr, 1987; cited in Arthur, 1998) and using the skills on similar tasks (transferring and generalising the skill, Lance et al, 1998). Lance et al (1998) examined task retention of US Air Force recruits. Using scores from workplace assessments, the study aimed to examine the effects of task complexity and individual ability on the performance of tasks over time. While they found a negative relationship between the time interval and performance, similar to (mostly) laboratory studies, in this study the time interval only accounted for 10% of the variance in performance. Ability and task complexity did not significantly contribute to the variation of airmen's performance in the workplace. Lance et al (1998) discussed these findings, and suggested that performance in the workplace was determined by the ability to transfer skills from one task to another.

5.1.4 Individual factors

5.1.4.1 Ability

Researchers have focussed on a few individual characteristics that are hypothesised to affect retention. A common assumption is that lower aptitude trainees forget more rapidly, and this has been found when training time has been fixed (Hurlock & Montague 1982). However, when both groups are trained to the same level of performance this effect is lost (Hurlock & Montague, 1982). When "fast" and "slow" learners were taught to the same standard, their forgetting curves are parallel (Gentile et al. 1982). Schendel and Hagman (1991) demonstrated that while high and low ability soldiers differed in initial performance, they did not differ in the *rate* of skill loss. The difference in skill level after an interval was attributed to the degree of initial learning.

It appears that there is no direct relationship between ability and task retention. Lance et al (1998) found that aptitude moderated retention for some tasks, but not others. Aptitude may be an intervening factor that interacts with other task and training factors to affect retention. Ability may interact with the training, affecting learning levels and so affect the overall skill performance and retention. This appears to occur in the "real world". Lance et al (1998) used workplace assessments, job training and results of the aptitude testing at recruitment to examine the relationships between work performance and aptitude. They found that retention did not vary with aptitude, and attributed this to the training system. The recruits had to repeat sessions or task training until they satisfied the criteria. There were no specific time limits on training; they were provided with sufficient training to complete the task however long that took. This reflects Schendel and Hagman's (1991) findings, that there was no difference in the rate of skill loss. So once all soldiers were trained to a specific level, retention did not vary according to individual ability.

5.1.4.2 Motivation

Motivation has been found to affect learning and is, therefore, thought to affect retention (Hurlock & Montague, 1982). Related emotional states such as self-efficacy (self-belief) have been found to increase learning. Low self-efficacy has been found to impair performance. Consequently, reducing a person's self-belief over a retention interval may contribute to performance loss.

5.2 Summary of Factors Affecting Skill retention

Many factors affect skill retention, as discussed above. What is needed is a method of assessing which factors have the greatest impact. If these can be identified then more work can be carried out into designing training programmes that lever off this knowledge. One approach to investigating the relative importance of various factors could use data-mining techniques, to assist in untangling underlying factors that have the biggest impact.

One example that is of interest is the effect of initial over training on long term retention, and subsequent time to train back up. In short, there may be significant advantage in initially training individuals beyond their required level of proficiency, in order to reduce future training requirements. To validate this hypothesis would require longitudinal studies involving at least two groups with different training regimes.

6. Models of skill retention

This work aimed to address the issues raised by the sensitivity analysis, described in Section 2.2 above. The rate of skill retention, within the field of psychology, is currently accepted to be best fit by a power function. However, there are several significant

difficulties encountered when using the power function in a dynamic model. In our work, we have suggested an alternative model of retention, using a System Dynamic (SD) framework that is suitable for dynamic modelling. This model requires more development and validation to assess its robustness.

In this section we will present the current thinking on models of skill retention from the psychological literature, including the criteria of assessing these models. We will then discuss some alternatives including the SD approach mentioned above.

6.1 Form of memory retention

Not surprisingly, one of the most robust findings, in both the laboratory and field, is the correlation between time since last practice and the drop in task performance (Arthur, et al., 1998; Schendel & Hagman, 1992; Farr, 1987; Mengelkoch, Adams & Gainer, 1971). Essentially, without practice, skills degrade. Mathematical models are used, within the literature to describe the rate of forgetting. While there have been studies looking at mathematical models of the long-term retention of skills, and short term forgetting, both cognitive (Bahrick, 1984; Bahrick, 1983) and complex multi-task skills (Sauer, Hockey & Wastell, 2000; Arthur et al, 1997), only one applied, predictive tool has been published (Hagman et al., 1980); the ARI Users Decision Aid (UDA). This tool will be discussed in a following section. Initially, the form of retention will be discussed, as this underlies any modelling effort of real-world skills.

The forgetting function has been characterised as a power function since Wicklegren (1972; Wixted & Ebbesen, 1991). Rubin and Wenzel (1996) reviewed the literature and found that the power function fitted the data (using accuracy of fit as a measure of performance), when compared to other models used, in the literature, to describe forgetting, including logarithmic, hyperbolic, and exponential. Anderson (1991) reviewed the skill retention literature, and also found that skill decay follows a power function. He called this the "power law of retention". J.R. Anderson and Schooler (1991) and Schooler and J.R. Anderson (1997) showed that the power function does extend to response time as a measure of performance. These forms, and their usefulness, have been evaluated using criteria beyond best fit by Lee (2001, in preparation). Roberts and Pashler (2000) have also critiqued the approach of only using "best fit" as the criteria for models. Basically, we should be cautious about the appropriateness of power function models, in the context of data which is generally noisy (i.e. large variance/error bars). This is discussed in detail in section 6.2.

Most of the mathematical descriptions of retention have used the results of cognitive tests on memory (Rubin and Wenzel, 1996) with the underlying assumption is that these results can be applied to other, more complex, tasks.

6.1.1 Mathematical model of learning and retention

Anderson, Fincham & Douglass (1999) describe a mathematical model, based on the power function, that aims to integrate practice and retention. Using a "slowed psychological clock" to explain the reduced rate of forgetting over time, Anderson et al (1999) suggest that the total strength of a memory is the sum of its decay (memory strength relies of the accumulation of repeated practice and retention). Their work directly relates learning to retention using a mathematical approach.

However, it does appear to have some shortcomings, namely, reference to the use of a "slowed psychological clock". This concept can be summed up as: forgetting slows down as clock-time increases; psychological time changes (slows) over time. This is not a universally accepted approach to explaining cognitive functioning, however it does demonstrate that the rate of forgetting changes over time (but not apparently as a direct function of clock time). For instance, we do not forget at the same rate at all times; time asleep reduces the amount of forgetting while other intervening activities increases the rate of forgetting.

Anderson, Fincham & Douglass (1999) used the power function to underpin their model, while acknowledging the continuing debate over its applicability. The power function does imply several theoretical concerns (see following section) researchers have not addressed directly. So their work, while positing a direct relationship between learning and forgetting, does not produce a robust model of retention that is amenable to dynamic modelling.

6.2 Criteria of "best fit"

The use of the power function to describe forgetting and retention is justified by it providing the "best fit" of the experimental data (Wixted & Ebbesen, 1997). However, there are concerns that using "best fit" alone as a measure of a model's suitability is limited. There are other factors, both theoretical and computational, that need to be taken into account when comparing models. Myung and Pitt (1997) examined one aspect of power functions difficulties, by showing that just measuring best fit fails to account for differences in the complexity of competing models. Lee (2001, in preparation) points out that a good psychological model should balance the competing demands of fit and complexity, providing the simplest possible account of the data.

There are several papers critiquing the apparent "best fit" of the power function to cognitive function. R.B. Anderson and Tweeney (1997) and Myung, Kim and Pitt (2000) discuss the circumstances under which the power function would artificially "best fit" averaged data. Myung, Kim and Pitt (2000) described the three conditions that are needed for the power law artefact to apply: "arithmetic averaging of data that are generated from a non-linear model in the presence of individual differences" (pg 832). Wixted and Ebbesen (1997) and R.B. Anderson and Tweney (1997) reanalysed data and found that the power function still achieved "best fit" once the conditions of the power law artefact were

addressed. So there are conditions under which the best fit of the power function is artificial but this does not eliminate the apparent best fit of the power function in all data.

Roberts and Pashler (2000) outlined a set of criteria for evaluating competing psychological models, discussing the shortcomings of "best fit" as proof of the theory. Roberts and Pashler (2000) state that ...

> "showing a theory fits data is not enough. By itself, it is nearly meaningless. Because of the flexibility of many theories, the variability of measurements and the simplicity of most psychological data functions, it is quite possible that the theory could fit any plausible outcome to within the precision of the data." (pg 361)

The authors continue to discuss why the use of "best fit" as the sole criteria for model use in psychology has remained. Roberts and Pashler (2000) compare psychology with other domains, such as statistics and computer sciences, which have been familiar with the problem of *"overfitting"*, and have accounted for it in various models (Anscombe, 1967; Leahy, 1994; Schaffer, 1993). Overfitting occurs when the model is too flexible; any data would support the model. So a model should not be judged only by how well it fits a data set, its flexibility needs to be assessed (and possibly penalised) (Roberts and Pashler, 2000).

Lee (in preparation) discusses the shortcomings of using "best fit" to compare different mathematical models, and compares both fit and model complexity to assess the models. Complexity aims to evaluate any model's robustness, and the noisiness of data, into account. Lee re-analysed 205 data sets collated by Rubin and Wenzel (1996), using a Bayesian network to compare both fit and complexity of the different models. Lee found that for data-fit, the power, hyperbolic and logarithmic functions were all found to fit the data equally well. For the measure of complexity, the hyperbolic function was found to be simpler, and therefore, more robust when using imprecise data (such as psychological data). When taken together, using best fit and complexity as criteria for comparison, the hyperbolic has a broader application to data than the other functions.

6.3 Theoretical concerns

The power function has two problems, aside from not offering any physical insights into skill retention or memory. First it is undefined at zero. It is possible to argue that this is a simple mapping problem and is fixed by simply avoiding the singularity. This ensures that the power function is valid in the timeframe of the experiment. However, this argument denies the continuity of learning then forgetting. That is, to isolate the singularity one must ignore very small times; this requires one to deny a smooth transition from learning to forgetting, which contrasts with all our experience of learning and forgetting.

Secondly the fact that the rate of change (1st order differential) is time dependent is also a problem. Each state of the system, at each time-step, in System Dynamics should ideally be

dependent only on the state of the system in the previous time-step. A rate of change dependent on time requires knowledge of the entire history of the system, i.e. its age. This suggests a notion of absolute time, which is a stronger requirement than in an exponential model that, for example, requires only the notion of relative time.

6.4 Alternative mathematical models of retention

An approach to modelling retention is needed that addresses the problems associated with the power function, yet is still able to provide the fit that power function supplies. Using a System Dynamics approach eliminates several of the theoretical concerns of a power function; it is defined at time zero, and the current state determines the future state. We have developed a simple System Dynamics model to illustrate the interaction between memory reinforcement (learning) and memory decay (forgetting).

6.5 Theoretical Basis of System Dynamic approach

The Influence Diagram (Figure 6-1) represents a simple System Dynamics (Coyle, 1996) model of retention. Simply, it assumes at each time-step an exponential decay that is dependent on another parameter, which we have called memory strength (M_S). Essentially as the memories are reinforced, the average memory strength increases, and hence the probability of decaying is reduced.This model has the added benefit that it has some parallels with the current thinking about the neurobiological basis of memory, which includes the concept of the strengthening of synaptic connections (Korn, 2001). It also addresses the issue of relearning. It does this because as the memory-set (M_R) is learnt-retained-and-relearned its average memory strength increases, which changes the nature of the retention curve. The model predicts a slower decay and higher plateau after relearning than in an initial learning event. This fits with the intuition of experienced trainers, and there is some evidence of this occurring in the limited experimental data.

This model directly relates learning (strength of memory reinforcement) with forgetting (strength of memory decay). This is similar to Anderson, Fincham & Douglass's (1999) power function model, which directly relates memory strength to retention, yet it does not have the same theoretical and computational shortcomings. The System Dynamic model is defined at zero, and relies on the interaction between memory strength and its likelihood of retention over time. This contrasts with Anderson, Fincham & Douglass's (1999) that is based on the power functions of practice and retention, and a concept of "slowed psychological time".

Essentially, both the Anderson, Fincham & Douglass (1999) model and a System Dynamic model aim to show that at any time, a memory strength is the sum of the previous learning (or memory strengths). There is a direct relationship between learning, forgetting and relearning.
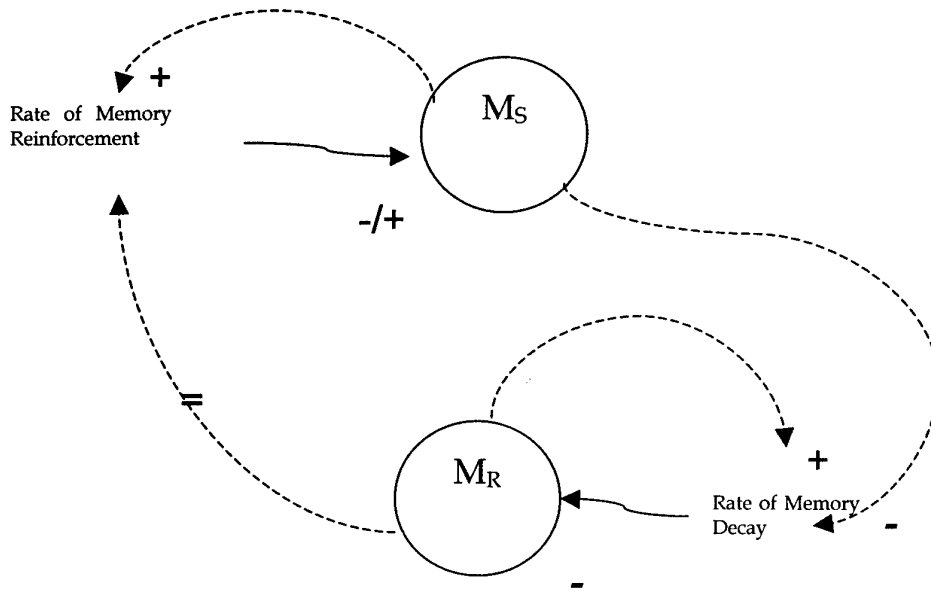
Figure 6-1: Influence Diagram of Learning and Retention

6.6 Comparison of System Dynamic fit to experimental data

The performance of the model was tested by fitting experimental data from Wixted and Ebbesen (1991; Wixted and Ebbesen, 1997). It compares favourably with the fit achieved by the power function for the same data-set (Table 6.1). The chi-squared for Wixted and Ebbesen (1991;1997) is compared to the chi-squared for the System Dynamic model. When allowing for the number of free parameters, using an adjusted chi-squared, the System Dynamic model has a better fit than the power function.

Table 6.1 Comparison of fit for power function versus System Dynamic model of Wixted and Ebbesen (1991) data set.

Model function	Sum of squares	Chi squared Test	(Reduced) Chi squared
W&E Power function	3.36×10^{-4}	99.85	1.12×10^{-4}
System Dynamic	3.38×10^{-6}	99.98	1.69×10^{-6}

*Note the data used was generated from the charts in Wixted and Ebbesen 1991 paper, and hence are only reliable to the second decimal place.
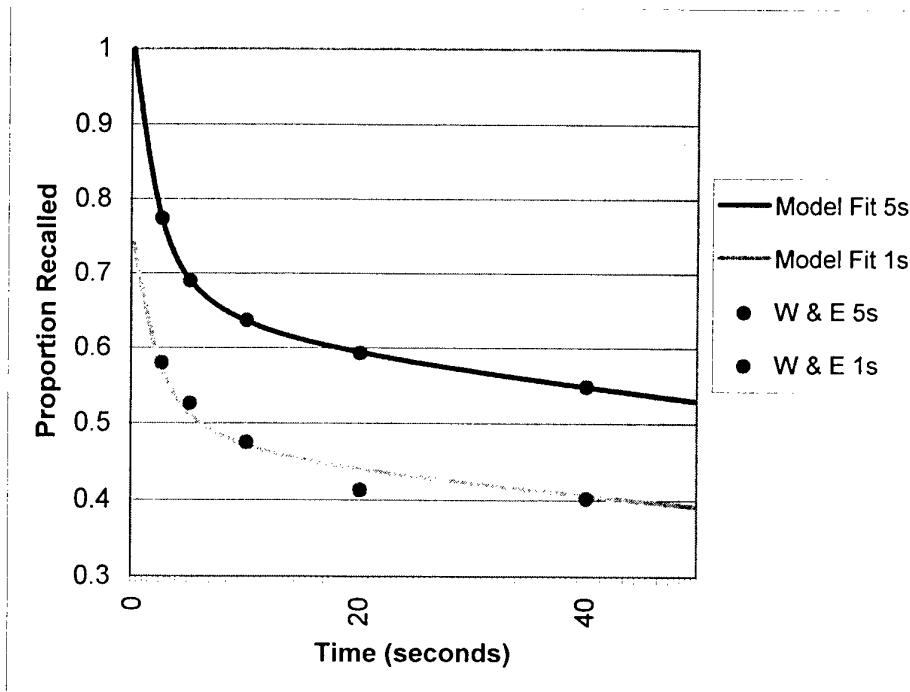
Figure 6-2 Re-analysis of Wixted and Ebbesen (1991) using System Dynamic model

6.6.1 Modelling Relearning

Another important advantage the System Dynamics approach has over a simple power function model is that, potentially, it can describe the effects of relearning. This is demonstrated in the two figures below. The power function produces a "static" learning and relearning relationship (left figure) compared with the SD model which shows a changing rate of relearning over time (right figure). There is some evidence in the literature that this occurs. However, this area has not been explored in great detail and is one that would require further study.
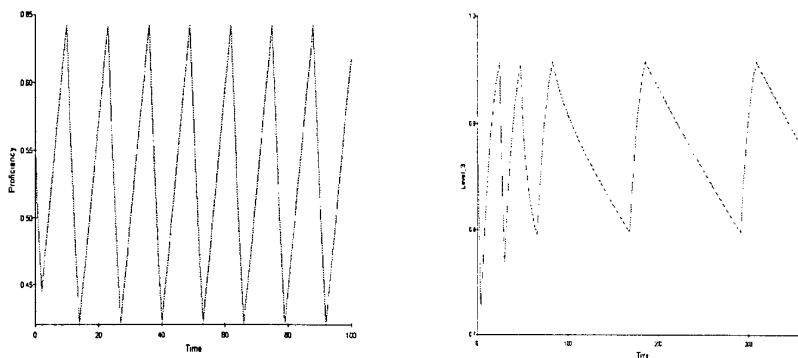

Figure 6-3 Relearning modelled by (left) power function and (right) SD model

6.6.2 Summary

The System Dynamic model provides a very close fit to Wixted and Ebbesen (1991) data, and produces a better fit using an adjusted chi-squared. So this initial application of the System Dynamic model to experimental data shows that it can provide the same if not better, precision as fitting a power function to the data. Further exploration of the model is being undertaken, in order to assess its usefulness compared to the current models. This includes evaluating and comparing its precision, robustness to parameter changes, and numbers of parameters to the current standard models (power function). These results will be circulated once completed.

The System Dynamic model is defined at time zero and so provides smooth transition from "learning" to "forgetting". The use of feedback loops also provides an analogous understanding of the physiological mechanisms that may underlie memory.

Directly linking learning and forgetting through the feedback loops provides a theoretical basis for further work offering a method of establishing simulation models and predictive models. The additional advantage of using this method is that the training models produced would be easily integrated into the existing System Dynamics readiness models discussed earlier.

23

6.6.3 Future directions for System Dynamic approach

Using System Dynamics as a method of modelling retention and learning is new; we are developing methods of evaluation and comparison to establish the utility and validity of such an approach. Adelaide University, in conjunction with Land Operations Division (LOD), is mapping a method of evaluation to assess the System Dynamic approach against the current standard models used in the domain.

7. Development of future directions

7.1 USA Army Research Institute (ARI) model of skill retention: the Users Decision Aid (UDA)

This model was developed as a convenient and practical method for unit commanders and training managers to allocate resources and training schedules (Rose et al 1985b). The tasks used to inform the UDA were set procedural tasks ("Common Soldier Tasks in MOS 11B" etc; Rose et al, 1985).

The ARI sought to identify the factors that affected retention that were available at the time of the study. They then used these to develop theoretical curves of retention to predict skill loss. ARI studies investigated the following variables, all of which were found to affect skill retention:

1. Overtraining
2. Practice schedules
3. Training techniques
4. Task difficulty
5. Number of steps required
6. Inter-step cuing
7. Step relevance

Selection of factors into the model was problematic. ARI focused on the task factors, such as complexity, as there is a greater consistency in the task factors compared to the training factors (Bryant & Angel, 2000).

The ARI did not aim to generate the retention curves from experiential data. The UDA was developed by: one, identifying task dimensions most likely to be related to retention; two, convert task dimensions into rating scales, develop anchor points, and analytically assign weights to each point on the scales: three, assess each scale's reliability and validity: four, examine inter-rater agreement and correlation between task ratings and actual retention data: five, iterate these steps until completed (Bryant & Angel, 2000).

The algorithm that weights and summarises the scores was initially developed from the available empirical literature. The ARI performed regression analysis on studies that they had conducted. The regression found eight task variables accounted for significant variation in retention levels at two months (79% percent of variance) and at four months, although less at the four-month interval. The results of the UDA provide a percentage of the group who would be able to perform the skill to standard.

Bryant and Angel (2000) point out that only one study (Rose et al 1985b) compared the predicted retention levels with the actual performance levels. Rose et al (1985b) found that the UDA provided a lower estimation of retention than was actually found; the longer the retention interval, the larger the difference. So more skill was retained than predicted. However, there was a significant correlation between the UDA scores and actual scores.

The output of the UDA is a table of theoretical curves that are selected on the basis of a set of task-specific questions. Essentially, any procedural task has characteristics that influence its retention over time, so by ranking task difficulty, the UDA provides pre-set likelihoods (percentages able to complete task to standard) eg. for a task with a score of 120, at two weeks after last practice, 45% would complete the task correctly.

7.1.1 Critique of ARI UDA model

While the paucity of other useable models prevents any comparison, there are several concerns regarding the assumptions that underpin the UDA. The generation of the theoretical curves, and that lack of validation is a concern. This can be overcome, with time and effort allocated to conduct longitudinal studies following up a cohort's skill set.

An important factor that the UDA failed to take into account is the level of initial training (or retraining). The level of training on retention has a direct effect on the rate of retention and relearning. This is a factor that the UDA needs to account for, if it is to be useful in scheduling training.

The form of the model is very user friendly; by answering a simple questionnaire, some idea of the skill's fragility is generated. This makes it very easy to administer and apply. If the format of the UDA could be kept, and the theoretical retention curves developed and validated, it would be a very useful tool aid in the optimisation of training schedules.

Further, it should be mentioned that the UDA not only predates much of the research on skill retention, but also the advent of neural nets and other data-mining tools that could efficiently and effectively map the correlation between the factors affecting retention of a skill and the parameters of the model that describes the data associated with that skill. Using such tools could allow the retention curve for a particular skill to be directly estimated, rather than selected from a pre-set family as in the current UDA. This represents another area where Science and Technology (S&T) could value add to training development. How the UDA could be used as a basis for developing a tool that

economically and accurately estimates skill retention curves is discussed in more detail in the following section.

7.2 Developing the ARI's UDA

The UDA could provide an easy to use method of estimating retention, however it has two major drawbacks: lack of validation and no consideration of training levels. These drawbacks make it unreliable. It is possible to make the UDA a basis for further development by providing a valid retention model and applying (and testing) the predicted outcomes. This would produce a useable tool for policymakers to evaluate task-training needs.

By using an approach like that of the US ARI's UDA model it is possible to develop it further, by using modern data-mining techniques (neural networks etc.), and questionnaires to estimate the parameters of a retention model.

7.2.1 Proposed method of further developing the UDA:

The process would be as follows:

1. Factors that may affect retention are compiled.
2. Questionnaire asking experts to assess, using numerical ratings, each of these factors for a certain skill, is then developed.
3. Taking a series of experiments for which data is available, the numbers from the questionnaire, with regards to these experiments, are fed into a neural network (or some other system identification tool) and these answers are associated with the parameters of the model that best fits the available data. This may be carried out on a large range of tasks. This process can also be used to refine the questionnaire.
4. The neural network can then be validated using further experimental data. By administering the questionnaire for these further data sets and testing the models generated by the neural net by how well they fit the data. If this is successful, then we are in a position to generate model parameters by simply administering questionnaires.
5. This will then generate estimates for the parameters, which will define the retention model for that particular skill.

Once we have the retention model an optimal training programme can be developed for that skill.

The advantage of this approach if successful would be that, once the neural network was trained, the development of a good, if not optimal, training programme would only require the completion of the questionnaire by training experts.
Future research should focus on finding more objective measures than a questionnaire. This might include physiologically based measures of cognitive activity such as the aforementioned EEG. For example, (albeit a speculative one) the level of cognitive activity

associated with a particular EEG, measured early in the training of certain tasks, might be used as a predictor for task complexity and hence retention. It is important to note that with the use of data-mining techniques it is not necessary to understand the cause and effects of the correlated measures and outcomes. Hence, objective measures may be identified that can be used as proxies for the factors affecting skill, and used without the requirement of knowing why or how they affect skill retention.

7.3 Summary of Skill Retention

There is a significant amount of information on the qualitative factors affecting retention; eg, overtraining, task characteristics, and generalisation of training. These factors can inform training design to maximise retention, and produce better task performance over time.

The standard quantitative description of retention is commonly accepted to be a power function. This form best fits a variety of experimental data. However, there are recent critiques of a power function describing a natural dynamic phenomenon such as task performance and retention. The criticism of the power function has been based on both theoretical and pragmatic grounds. We have developed an alternative mathematical model that addresses the shortcomings associated with a power function, when it is part of a dynamic system. The alternative model uses a System Dynamics approach to construct a simple theoretical model, which fits a sample data set as well as the power function.

The System Dynamic model will be developed, tested and validated using a wider variety of data sets utilising a variety of tasks including cognitive and motor tasks.

There is only one useable, predictive model of retention available, the ARI UDA. The UDA has several shortcomings. Developing the ARI UDA as a framework for further work to include other factors and establish the validity of the retention curves may produce a useful tool to inform policymakers on training schedules.

8. Conclusion and Recommendations

The work we have carried out so far has provided qualitative support for the judgements made by Army trainers. To extend this to quantitative support requires more work, in the development of experiments and measures for proficiency, and the development of models that predict proficiency of certain skills. The main aim would be to augment the qualitative and subjective measures of the Army Subject Matter Experts with quantitative, objective measures. As well as providing Army training planners with information on how individuals and teams learn and forget, to allow them to plan more effective training regimes.

There are four areas where high pay-off research can be applied to the Army training problem.

- The development of theoretical models that better describe Skill Retention and Acquisition, across a variety of skills and tasks and for populations of varying experience and training history.
- The development of tools, such as data-mining techniques, to efficiently predict the Skill Retention and Acquisition for the large number of individual and collective skills.
- Research into the relationship between individual and collective skills. As well as the development of measures of collective proficiency.
- The development of objective, quantitative measures, such as physiological measures, for proficiency.

The above list is by no means exhaustive, but still represents a large Research and Development effort. Any such programme would have a mixture of short and long term goals. The short term goals of this research would include:

- Investigate the training regimes of ammunition intensive training such as Artillery, and tank training. To investigate the possibility of collecting performance data in these areas, and to use this data to inform simple models of skill degradation in these areas.
- Ensure that data for small arms firing is being collected in a form that can inform the development of theoretical models of skill degradation. This data should be collected both in the Weapons Training Simulation System (WTSS) facilities and Live Firing Ranges across the country.
- Investigate the possibility of applying the US Army Research Institute, UDA model, to Australian Army Skill sets. As well as endeavouring to improve this approach.

Longer term goals would include:

- The development of sophisticated models of skill degradation and acquisition that account for the effects of relearning.
- Establishing Research Agreements with universities to develop objective measures for proficiency, such as interpreted electroencephalograms (EEG), collected by field deployable equipment.
- To develop a framework that economically and accurately predicts skill retention and acquisition, as encapsulated in the models described above. This would incorporate experimental data sets such as collections of the objective measures mentioned above, and may use data-mining techniques to achieve the predictions.
- Develop an understanding of the relationship between individual and collective skills and develop appropriate measures for collective skills.

Related research would include:

- An investigation as to the effectiveness of simulation in a variety of training environments and for a variety of desired outcomes.
- The development of improved training techniques and the use of technology in training. For example, the effect of different types of feedback within training sessions.

The above represents a number of significant R&D programmes that, to be properly supported, would require significant investment of time and money by Army, DSTO, and other research institutes. However, the pay-offs of such research have the potential to be very high. These pay-offs would be beyond the confines of the current Army Ammunition Study, and into the wider arena of Army training. The ultimate goal of such R&D would be to augment and complement the methods currently used by Army Subject Matter Experts to produce more effective, auditable, and hence efficient training systems.

9. References

Ackerman, P.L. (1987) Individual differences in skill learning: AN integration of psychometric and information processing perspectives. *Psychological Bulletin, 102*(1), 3-27.

Ackerman, P.L. (1992) Predicting individual differences in complex skill acquisition: Dynamics of ability determinants. *Journal of Applied Psychology. 77,* 598-614.

Ackerman, P.L., & Humphreys, L.G. (1991) Individual difference theory in industrial and organisational psychology. In M.D. Dunnette & L.M. Hough (Eds.) *Handbook of industrial and organisational psychology* (2nd ed., Vol 1, pp.223-282). Palo Alto, CA: Consulting Psychologists Press.

Allard, F & Starkes, J.L. (1991) Motor skill experts in sports, dance and other domains. In K. Anders Ericsson and J. Smith (Eds) *Towards a general theory of expertise.* Cambridge: Cambridge University Press.

Anderson, J.R. (Ed) (1991) *Cognitive Skills and their acquisition* Hillsdale, NJ: Erlbaum.

Anderson, J.R., Fincham, J.M. and Douglass, S. (1999) Practice and Retention: A unifying analysis. *Journal of Experimental Psychology: Learning, Memory and Cognition, 25*(5), 1120-1136.

Anderson, J.R., & Schooler, L.J. (1991) Reflections of the environment in memory. *Psychological Science, 2,*396-408.

Anderson, R.B., & Tweney, R.D., (1997) Artifactual power curves in forgetting. *Memory & Cognition, 25,* 724-730.

Anscombe, F.J. (1967) Topics in the investigation of linear relations fitted by the method of least squares (with discussion). *Journal of the Royal Statistical Society, Series B, 90,* 1-52.

Arthur, W., Bennett, W., Stanush, P.L. & McNelly, T.L. (1998) Factors that influence skill decay and retention: A quantitative review and analysis. *Human Performance, 11*(1), 57-101.

Austin, S., and Gilbert, K. (1973) Student performance in a Keller-Plan Course in Introductory Electricity and Magnetism. *American Journal of Physics, 41,* 82-87.

Australian National Audit Office (2001) Australian Defence Force Reserves: Department of Defence. (Audit Report No. 33 2000-2001) Performance Audit. Australian National Audit Office: Canberra.

Bahrick, H.P.(1983) The cognitive map of a city: fifty years of learning and memory. *Psychology of Learning and Motivation, 17,*125-163.

Bahrick, H.P. (1984) Semantic memory content in permastore: Fifty years of memory for Spanish learned in school. *Journal of Experimental Psychology: General, 113,*1-29.

Baldwin, T.T., & Ford, J.K. (1988) Transfer of training: A review and directions for future research. *Personnel Psychology, 41,* 63-105.

Bryant, D.J., & Angel, H. (2000) *Retention and fading of Military Skills: Literature Review.* (PWGSC Contract No W7711-9-7539/001/TOR) Guelph, Ontario: Human*systems* Incorporated.

Clarke, C.R. Private communication of unpublished work, Centre for Cognitive and Neuropsychological Research, Flinders University of South Australia.

Coyle , R.G. *System Dynamics Modelling.* (1996) Chapman and Hall, London UK.

Cooke, N., Durso, F., and Schvaneveldt, R. (1994) Retention of skilled search after nine years. *Human Factors,* 36, 4, 597-605.

Driskell, J.E., Willis, R.P. & Copper, C. (1992) Effect of overlearning on retention. *Journal of Applied Psychology, 77,*615-622.

Flexser, A.J., & Tulving, E. (1978) Retrieval independence in recognition and recall. *Psychological Review, 85,* 153-171.

Gallagher, J.J. (1994) Teaching and learning: New models. *Annual review of Psychology, 57,*171-195.

Gentile, J.R., Monaco, N., Iheozor-Ejiofor, I.E., Ndu, A.N. & Ogbonaya, P. (1982) Retention by "fast" and "slow" learners. *Intelligence, 6,* 125-138.

Goldstein, I.L., (1991) Training in work organisations. In M.D. Dunnette & L.M. Hough (Eds) *Handbook of Industrial and Organisational Psychology*(2nd ed, Vol 2., pp507-620). Palo Alto, CA: Consulting Psychologists Press

Hagman, J., & Rose, A. (1983) Retention of military tasks: A review. *Human Factors, 25*(2), 199-213.

Healy, A.F., Fendrich, D.W., Crutcher, R.J., Wittman, W.T., Gesi, A.T., Ericsson, K.A., & Bourne, L.E. Jr. (1992) The long term retention of skills. In A.F. Healy & S. Kosslyn (Eds) *Essays in honour of William Estes: From learning processes to cognitive processes* (Vol 2, pp 87-118). Hillsdale, NJ: Lawrence Erlbaum Associates.

Hintzman, D.L. (1990) Human learning and memory: Connections and disassociations. *Annual Review of Psychology, 41,*109-140.

Hutton, R.J.B. & Klein, G. (1999) Expert decision making. *Inc Syst Engineering, 2,* 32-45.

Hurlock, R.E., & Montague, W.E. (1982) *Skill retention and its implications for nay tasks: AN analytic review* (NPRDC Special Rep. No. 82-21) San Diego, CA: Navy Personnel Research and Development Centre.

Klein, G. (1998) *Sources of power: How people make decisions.* MIT Press: Cambridge, MA.

Korn, M.L., 2001, New Findings in Molecular Biology and Memory, Amer. Psychiatric Assoc. 2001 Annual Meeting.

Kraiger, K., Ford, J.K., & Salas, E. (1993) Application of cognitive, skill-based and affective theories of learning outcomes to new methods of training evaluation. *Journal of Applied Psychology, 78,* 311-328.

Lance, C.E., Parisi, A.G., Bennett, W.R., Teachout, M.S., Harville, D.L. & Welles, M.L. (1998) Moderators of skill retention interval/performance decrement relationship in eight US Air Forced enlisted specialties. *Human Performance, 11*(1), 103-123.

Leahy, K. (1994) The overfitting problem in perspective . *AI Expert, 9*(10), 35-36.

Lee, M.D. (in preparation) A Bayesian Analysis of Retention Functions. Unpublished manuscript, Adelaide University.

Logan, G.D., (1988) Toward an instance theory of automatization. *Psychological Review, 95,* 495-527.

Marmie, W.R.& Healy, A.F. (1995) The long-term retention of a complex skill. In A.F. Healy & L.E. Bourne (Eds) *Learning and memory of knowledge and skills: Durability and specificity* (pp 30-65). Thousand Oaks, CA: Sage Publications.

Mengelkoch, R.F., Adams, J.A., & Gainer, C.A. (1960) *The forgetting of instrument flying skills as a function of the initial level of proficiency* (Report No NAVTRADEVCEN 71-16-18). US Naval Training Device Centre: Port Washington, NY.

Michalak, D.F. (1981) The neglected half of training. *Training and Development Journal, 35*(5), 22-28.

Myung, I.J., Kim, C. & Pitt, M.A. (2000) Toward an explanation of the power law articfact: Insights from response surface analysis. *Memory and Cognition, 28*(5), 832-840.

Myung, I.J. & Pitt, M.A. (1997) Applying Occam's razor in modelling cognition: A Bayesian approach. *Psychonomic Bulletin & Review, 4(1),* 79-95.

Neves, D.M. & Anderson, J.R. (1981) Knowledge compilation: Mechanisms for the automatisation of cognitive skills. In J.R. Anderson (Ed) *Cognitive Skills and their acquisition* (pp 57-84) Hillsdale, NJ: Erlbaum

Newell, A. & Rosenbloom, P.S. (1981) Mechanisms of skill acquisition and the law of practice. In J.R. Anderson (Ed) *Cognitive Skills and their acquisition* (pp 1-55) Hillsdale, NJ: Erlbaum

Prophet, W.W. (1977) *Long term retention of flying skills: A review of the literature.* (AD/A-036-077) Alexandria, VA: Human Resources Research Organisation.

Roberts, S., & Pashler, H. (2000) How persuasive is a good fit? A comment on theory testing. *Psychological Review, 107(2),* 358-367.

Rose, A.M., McLaughlin, D.H., Felker, D.B., Hagman, J.D. (1985) Retention of soldering skills: Review of recent ARI research (Tech Rep 530). Alexandria, VA: US Army Research Institute.

Royer, J.M. (1979) Theories of the transfer of learning. *Educational Psychologist, 14,* 53-69.

Rubin, D.C. & Wenzel, A.E., (1996) One hundred years of forgetting: A quantitative description of forgetting. *Psychological Review, 53,* 734-760.

Sauer, J., Hockey, G.R.J. & Wastell, D.G. (2000) Effect of training on short- and long-term skill retention in a complex multi-task environment. *Ergonomics, 43(12),* 2043-2064.

Schaffer, C. (1993) Overfitting avoidance as bias. *Machine Learning, 10,* 153-178

Schendel, J.D., & Hagman, J.D. (1980) *On sustaining procedural skills over prolonged retention intervals.* (Research Rep 1298) US Army Research Institute for the Behavioural and Social Sciences: Alexandria, VA

Schendel, J.D., & Hagman, J.D., (1982) On sustaining procedural skills over a prolonged retention interval. *Journal of Applied Psychology, 67,* 605-610.

Schooler, L.J., & Anderson, J.R. (1997). The role of process in the rational analysis of memory. *Cognitive Psychology, 32,* 219-250.

Schmidt, R.A., & Bjork, R.A. (1992) New conceptualisations of practice: Common principles in three paradigms suggest new concepts for training. *Psychological Science, 3,* 207-217.

Semb, G., and Ellis, J., (1994) Knowledge taught in school: What is remembered? *Review of Educational Research, 64,* 3, 243-286

Shields, J.L., Goldberg, S.L., & Dressel, J.D. (1979) Retention of basic soldiering skills (research Rep No 1225) Alexandria, VA: US Army Research Institute of Social and Behavioural Sciences.

Shute, V.J. & Gawlick, L.A. (1995) Practice effects on skill acquisition, learning outcome, retention and sensitivity to relearning. *Human Factors, 37*(4). 781-803.

Snow, R.E., & Swanson, J. (1992) Instructional psychology: Aptitude, adaptation and assessment. *Annual Review of Psychology, 43,* 583-626.

Stothard, C.F. & Nicholson, R.J.F. (2001) *Skill retention: Building a foundation.* DSTO-CR-XXXX. In preparation.

Tannenbaum, S.I., & Yulk, G. (1992) Training and development in work organizations. *Annual Review of Psychology, 43,* 399-441.

Wetzel, S., Knoske, P. and Montague, W. (1983) *Estimating skill degradation for aviation anti-submarine warfare operators (AWS): Loss of skill and knowledge following training* (technical Report NPRDC TR 83-31). San Diego, CA: Navy Personnel Research and Development Center.

Weiss, H.M. (1991) Learning theory and industrial and organisational psychology. In M.D. Dunnette & L.M. Hough (Eds.) *Handbook of industrial and organisational psychology* (2nd ed., Vol 1, pp.171-222). Palo Alto, CA: Consulting Psychologists Press.

Wickelgren, W.A. (1972) Trace resistance and the decay of long-term memory. *Journal of Mathematical Psychology, 9,* 418-455.

Wisher, R., Sabol, M.A. & Ellis, J.A. (1999) *Staying Sharp: Retention of military knowledge and skills (ARI Special Report 39).* US Army Research Institute for Social and Behavioural Sciences: Alexandria, VA

Wixted, J.T., & Ebbesen, E.B., (1991) On the form of forgetting. *Psychological Science, 2,* 409-415.

Wixted, J.T., & Ebbesen, E.B., (1997) Genuine power curves in forgetting: A quantitative analysis of individual subject forgetting functions. *Memory & Cognition, 25,* 731-739.

DISTRIBUTION LIST

Skill Acquisition and Retention in Training:
DSTO Support to the Army Ammunition Study

C Stothard and R Nicholson

AUSTRALIA

DEFENCE ORGANISATION

Task Sponsor   **COS LHQ**

S&T Program
  Chief Defence Scientist ⎱
  FAS Science Policy ⎰ shared copy
  AS Science Corporate Management ⎰
  Director General Science Policy Development ⎰
  Counsellor Defence Science, London (Doc Data Sheet)
  Counsellor Defence Science, Washington (Doc Data Sheet)
  Scientific Adviser to MRDC Thailand (Doc Data Sheet)
  Scientific Adviser Joint
  Navy Scientific Adviser (Doc Data Sheet and distribution list only)
  Scientific Adviser - Army
  Air Force Scientific Adviser
  Director Trials

  Aeronautical and Maritime Research Laboratory
  Director

  Electronics and Surveillance Research Laboratory
  Director (Doc Data Sheet and distribution list only)

  Chief of Land Operations Division
  Research Leader Operations and Exercise Analysis
  Research Leader Human Systems Integration
  Mission Head Support Systems Studies
  Mission Head Training Technologies
  HME
  HHSI
  A Vozzo
  P Temby
  R Nicholson
  C Stothard

DSTO Library and Archives
  Library Fishermans Bend (Doc Data Sheet only)
  Library Maribyrnong (Doc Data Sheet only)
  Library Edinburgh (1 copy)

Australian Archives
Library, MOD, Pyrmont (Doc Data sheet only)
US Defense Technical Information Center, 2 copies
UK Defence Research Information Centre, 2 copies
Canada Defence Scientific Information Service, 1 copy
NZ Defence Information Centre, 1 copy
National Library of Australia, 1 copy

Capability Development Division
Director General Maritime Development (Doc Data Sheet only)
Director General Aerospace Development (Doc Data Sheet only)

Knowledge Staff
Director General Command, Control, Communications and Computers (DGC4)
(Doc Data Sheet only)

Army
DGPP-A AHQ
Chairman of Army Ammunition Study (COL Ahmad Mostafa), Victoria Barracks,
Paddinngton NSW 2021
COL DEV LHQ
SO1 Force Structure LHQ
SO1 TRG PLANS LHQ
SO2 Concepts, Development Branch, LHQ
Stuart Schnaars, ABCA Standardisation Officer, Tobruk Barracks, Puckapunyal, 3662
(4 copies)
SO (Science), DJFHQ(L), MILPO Gallipoli Barracks, Enoggera, QLD 4052
SO (Science), LHQ
SO (Science), TC-A HQ, Suakin Drive, Georges Heights NSW 2988

Intelligence Program
DGSTA Defence Intelligence Organisation
Manager, Information Centre, Defence Intelligence Organisation

Corporate Support Program
Library Manager, DLS-Canberra

UNIVERSITIES AND COLLEGES
Australian Defence Force Academy
Library
Head of Aerospace and Mechanical Engineering
Serials Section (M list), Deakin University Library, Geelong 3217
Hargrave Library, Monash University (Doc Data Sheet only)
Librarian, Flinders University

OTHER ORGANISATIONS

NASA (Canberra)
AusInfo
State Library of South Australia

OUTSIDE AUSTRALIA

ABSTRACTING AND INFORMATION ORGANISATIONS
Library, Chemical Abstracts Reference Service
Engineering Societies Library, US
Materials Information, Cambridge Scientific Abstracts, US
Documents Librarian, The Center for Research Libraries, US

INFORMATION EXCHANGE AGREEMENT PARTNERS
Acquisitions Unit, Science Reference and Information Service, UK
Library - Exchange Desk, National Institute of Standards and Technology, US
National Aerospace Laboratory, Japan
National Aerospace Laboratory, Netherlands

SPARES (5 copies)

Total number of copies: **64**

DEFENCE SCIENCE AND TECHNOLOGY ORGANISATION DOCUMENT CONTROL DATA	1. PRIVACY MARKING/CAVEAT (OF DOCUMENT)

2. TITLE Skill Acquisition and Retention in Training: DSTO Support to the Army Ammunition Study	3. SECURITY CLASSIFICATION (FOR UNCLASSIFIED REPORTS THAT ARE LIMITED RELEASE USE (L) NEXT TO DOCUMENT CLASSIFICATION) Document (U) Title (U) Abstract (U)

4. AUTHOR(S) Christina Stothard and Robin Nicholson	5. CORPORATE AUTHOR Electronics and Surveillance Research Laboratory PO Box 1500 Edinburgh SA 5111 Australia

6a. DSTO NUMBER DSTO-CR-0218	6b. AR NUMBER AR-012-050	6c. TYPE OF REPORT Client Report	7. DOCUMENT DATE December 2001

8. FILE NUMBER D9505-21-169	9. TASK NUMBER ARM 01/059	10. TASK SPONSOR DGLD	11. NO. OF PAGES 34	12. NO. OF REFERENCES 63

13. URL on the World Wide Web: http//www.dsto.defence.gov.au/corporate/reports/dsto-cr-0218.pdf	14. RELEASE AUTHORITY Chief, Land Operations Division

15. SECONDARY RELEASE STATEMENT OF THIS DOCUMENT
Approved for public release

OVERSEAS ENQUIRIES OUTSIDE STATED LIMITATIONS SHOULD BE REFERRED THROUGH DOCUMENT EXCHANGE, PO BOX 1500, EDINBURGH, SA 5111

16. DELIBERATE ANNOUNCEMENT No Limitations

17. CASUAL ANNOUNCEMENT	Yes

18. DEFTEST DESCRIPTORS Performance tests Training evaluation Learning Military training Skills acquisition Human factors engineering

19. ABSTRACT
This work was undertaken within the Land Preparedness Studies Task (ARM 01/059) within the Land Operations Division. It is in response to the request for DSTO support contained within the Terms of Reference for the Army Ammunition Study (AAS). This request was for DSTO to assist in developing a better understanding of skill degradation (retention) and acquisition, with the view that this would enable accurate prediction of training requirements and in the longer term the development of tools to predict levels of proficiency provided by different training regimes.

In summary we:
• Took a systems approach to the training problem and assessed the impact of the treatment of skill acquisition and degradation of predictions of training frequency requirements.
• Undertook a literature survey, concentrating on the Cognitive Psychological literature, to ascertain the current thinking on how people learn and forget.
• Assessed the gaps in the literature and scoped the work needed to address these gaps.
Assessed how DSTO could address the issues raised by the AAS in both the short and long terms, through the development of appropriate R&D programmes.