Conducting Training and Simulation Research: A Primer for Practitioners

Philip Temby	**Susannah J. Whitney**
Defence Science and Technology*	**Defence Science and Technology**
Edinburgh, South Australia	**Edinburgh, South Australia**
philip.temby@dst.defence.gov.au	susannah.whitney@dst.defence.gov.au

ABSTRACT

Training is a fundamental input to defense capability and each year military units invest significant resources on training service personnel. To ensure this training is effective, it is important that the design, implementation and evaluation processes are based on best practices. Much has been written on the science of training, including considerations for conducting training research in organizational settings. Despite this literature, our experience has highlighted the need for ongoing education of practitioners in the training and simulation community. The aim of this paper is to provide practical guidance for professionals involved in designing, implementing and evaluating training in military settings. While there are many excellent handbooks and articles available on training evaluation and simulation-based training, they are usually quite detailed or focused on specific aspects of training and simulation, and not well suited to professionals who may be new to the area and seeking a quick introduction to key issues. To address this gap, this paper outlines key issues in the form of Frequently Asked Questions (FAQs) within each of the five stages of the commonly used ADDIE model of training. The contents of the paper are drawn from the published literature, as well as the authors' combined experience of working in training and simulation research for 30 years in military settings. We hope this 'primer' will become a practical resource for training designers, developers, instructors, device manufacturers, and researchers – particularly for newcomers to the area. The paper concludes with some suggestions for future research within the training and simulation community. This paper is important to the community because it provides an overview of key issues involved in conducting training and simulation research, and reinforces best practices associated with the design, implementation and evaluation of training systems; all of which is presented as a series of FAQs for a broad audience.

ABOUT THE AUTHORS

Philip Temby is a senior research scientist within Australia's Defence Science and Technology where he holds the position of S&T Discipline Leader Training and Simulation. He has worked in defence training and simulation for over 18 years, and his research interests include training technologies, cognition, resilience, and training effectiveness. He is a Registered Psychologist with a Master of Psychology degree from the University of South Australia and a Bachelor of Science with First Class Honors from the University of Adelaide. He also holds a Graduate Certificate in Scientific Leadership from the University of Melbourne.

Susannah Whitney is a senior human scientist within Australia's Defence Science and Technology. She has worked in simulation and training for over 12 years. Her research interests include training systems design, simulation, and training evaluation. She holds a PhD in Psychology from the University of Queensland and a BA (Psychology) from the University of Newcastle.

*Defence Science and Technology (DST) is part of the Australian Department of Defence. DST is the Australian government's lead agency responsible for applying science and technology to safeguard Australia and its national interests. See https://www.dst.defence.gov.au/ for more information.

Conducting Training and Simulation Research: A Primer for Practitioners

Philip Temby
Defence Science and Technology
Edinburgh, South Australia
philip.temby@dst.defence.gov.au

Susannah J. Whitney
Defence Science and Technology
Edinburgh, South Australia
susannah.whitney@dst.defence.gov.au

INTRODUCTION

Each year defense forces invest significant resources into training service personnel. This training is designed to equip these personnel with the relevant knowledge, skills and abilities (KSAs) they need for specific job roles. Given the critical importance of training in preparing service personnel to carry out their duties, it is vital that training is effective and efficient. While military organizations have been conducting training for centuries, in recent decades, considerable research has been conducted into the science of training (Bell et al, 2017; Salas et al, 2012; Salas & Cannon-Bowers, 2001) and identified practical guidance for the design, conduct, and evaluation of training.

Several excellent handbooks and articles have been written on best-practice training design and delivery (e.g., Best et al, 2013; Phillips, 2012; Salas et al, 2006; Salas & Stagl, 2009); however, these publications are quite detailed, or focus on specific aspects of training and simulation. These qualities can make them less appealing for professionals who may want quick and easy answers or are seeking to become familiar with key issues. While shorter papers outlining training principles and practical lessons for researchers have been published (e.g., Siebert et al, 2010; Whitney, 2019), we saw a need to supplement this literature with a primer in the form of frequently asked questions (FAQs) as a way to (further) educate practitioners. To our knowledge no such paper has been written for the training and simulation community.

This paper outlines key considerations for the conduct of training and simulation research in military settings, and takes its inspiration from other primers (e.g., NTSA, 2011). In doing so, we aim to provide a useful resource for training and simulation professionals, especially for those who may be new to this area. While simulation is used for training, we have deliberately used the terms 'training and simulation' in the title in the hope of capturing the educational interest of professionals in both communities. We have written this paper based on published literature and our thirty years of experience conducting training research in support of the Australian Defence Force. This research has included lab studies (e.g., Hibbard et al, 2016), field trials (e.g., Temby & Stephens, 2014), case studies (e.g., Temby & Stephens, 2015; Whitney & Stephens, 2014), literature reviews (e.g., Whitney, Temby, & Stephens, 2014), and advice on training policy (e.g., Whitney, Temby, & Hoggan, 2018).

The paper is organized into five main sections that correspond with the stages of the ADDIE model of training. We start by briefly introducing the systems approach to training and the ADDIE model. In subsequent sections we list FAQs, provide brief answers to them, and highlight additional references for readers. The FAQs are not an exhaustive list but cover the types of questions we have been asked over the years, and we believe are fundamental to answer to educate those working in the area. We conclude the paper with specific advice for researchers working in military settings, and by outlining suggestions for future training research. Whether you are an instructor, researcher, or industry professional involved with military training, we hope this primer provides a useful resource of key issues to consider before, during, and after training delivery.

A SYSTEMS APPROACH TO TRAINING

Training is a systematic approach to the delivery of KSAs for individuals and teams. This training takes place within a larger socio-technical system comprised of people, instructional methods, and the work environment. A systems approach to training aims to identify interactions between these elements and their impact on training effectiveness (Department of Defense, 2016). While various conceptual models have been developed to describe training effectiveness (e.g., Alvarez, Salas, & Garofano, 2004; Salas et al, 2012), common elements in these models include trainee characteristics, training design/content, and the organizational context in which training is conducted.

THE ADDIE MODEL

A common framework used to describe training design and development is the ADDIE model which derives from instructional systems design (Branch 2009). The ADDIE model considers training in terms of five stages, namely: Analysis, Design, Development, Implementation, and Evaluation. A brief description of each of the five stages is outlined below.

- *Analysis*: the training need is identified, training goals are developed, and a plan is created for designing, implementing, and evaluating the training;
- *Design*: the training and learning objectives, performance measures, and instructional methods are specified;
- *Development*: the content of the training program is developed, trialled, and refined;
- *Implementation*: the training package or program is delivered; and
- *Evaluation*: the training package or program is evaluated against defined criteria.

While the ADDIE model is often depicted as a sequential and cyclic process (see figure above, Source: https://en.wikiversity.org/wiki/Instructional_Design), the research does not have to start at the Analysis phase, and some stages may be conducted concurrently.

1. TRAINING ANALYSIS

The purpose of the Analysis phase is to identify the training need and create a plan for addressing this need. From a systems perspective, this includes understanding the important elements and interactions within the system, and the drivers for seeking to introduce change. In the following sections we outline key considerations during this phase.

How can I identify the need for training? What are some key considerations?
Within the military, the need for training may be identified at the strategic level in response to a perceived performance or training gap (e.g., improving baseline levels of physical fitness). Alternatively, a training need might be linked to organizational imperatives aimed at modernizing current training methods (e.g., trialing a blended learning method). Whatever the rationale, it is important that training researchers have a good understanding of the current state of the system. This includes knowledge of how tasks are conducted, the KSAs required by trainees and instructors, and the way that training is currently delivered. From there, the next step is to conduct a training needs analysis. This is done to determine where training is needed, what needs to be taught, and who needs to be trained (Salas & Cannon-Bowers, 2001). Typical methods employed to address these issues include task analyses, cognitive task analyses, workshops, and interviews with subject matter experts.

Is training the 'right' solution?
Training may not always be the solution to a perceived performance gap or workplace need. Even the best training cannot overcome human capacity limitations (physical and cognitive), poor equipment, organizational policies, or sub-optimal workspace designs (Whitney, 2019). It is important to remember these constraints (when assessing the need for a training solution) and to consider non-training solutions as part of the overall system. Taking a broader human-systems integration (HSI) perspective may help identify non-training solutions and other fundamental inputs to capability. For more details on HSI domains, we recommend reading Booher (2003) and Salvendy (2012).

What if there is no 'need' for training?
While an identified performance gap is one reason for introducing a training intervention, researchers and practitioners may also seek to explore training interventions in the absence of a specific performance gap. For instance, the emergence of virtual, mixed, and augmented reality head-mounted displays has generated considerable research interest in their use for training (e.g., Hoover et al, 2018; Whitney, Fidock, & Hibbard, 2016). In addition, where new training interventions are being proposed, either by researchers or military stakeholders, there may be limited evidence for their efficacy. In such cases, it is important for training researchers and military stakeholders to discuss the potential risks and benefits of the new training solution, and how the outcomes might impact on the current training system, in terms of time, cost, and other resources.

2. TRAINING DESIGN

During the Design phase, the focus is on developing training objectives, considering individual and organizational factors that may influence training, selecting the training method(s), and specifying the training content.

What individual characteristics and organizational factors can influence training outcomes?
Individual characteristics that have been linked to training outcomes include cognitive ability, self-efficacy, motivation, goal orientation, prior knowledge and experience, and attitudes/expectations (Salas et al, 2012). When designing training, a key consideration is how to support the learning of low-ability trainees and those with negative training experiences; this is where treating individual characteristics as key variables may be useful. Organizational characteristics, such as unit climate and tempo, policies and procedures, and availability of resources, can also influence training outcomes. It is recommended that practitioners engage with military stakeholders early to understand how such factors may impact on the feasibility of a training solution and its likely success.

Does the current generation of military personnel prefer using technology for training?
Some authors believe the current generation of military personnel are "digital natives" who have grown up using digital devices, playing videogames and have a strong preference for using technology for learning (e.g., Prensky, 2001). These views are, however, not supported by empirical research, which has examined attitudes towards and use of technology among young military personnel (e.g., Bollard, Kerry, Whitney, & Fidock, 2015; Orvis et al, 2010). As suggested by Christensen and Tremblay (2013), rather than focusing on whether different generations will learn more from specific instructional methods, it is better for training designers to focus on the learning objectives, and to use instructional methods that are tailored to individual learning needs.

What are some key considerations when selecting a training method?
There is no single training method for every situation. The selection of training method will need to consider the number of trainees, availability of instructors (and associated staff), time available, availability of training areas and equipment, and the cost involved (Australian Army, 2017). These issues are invariably a trade-space and researchers need to consider how their research objectives can be achieved with available resources; taking a programmatic approach over a longer timeframe is likely to be helpful. Where new training programs are being considered and published information is unavailable, pilot studies are recommended to estimate the time needed to deliver training and achieve desired learning outcomes.

What training methods or instructional approaches are effective?
The design of training includes the selection of instructional strategies that are appropriate for trainees and organizations, and the KSAs to be trained. Numerous instructional approaches have been found to be effective for individual and team training, including classroom training, distance learning, and technology-based methods. Table 1 provides a brief description of some common instructional strategies. It is beyond the scope of this paper to cover them in depth, and we refer readers to Salas et al (2012) for details about these and other approaches.

Table 1. Common instructional principles, methods, and approaches applied in training research

Principle/Method/Approach	Description
Blended learning	Blended learning is an approach that combines online educational content and interaction methods with traditional classroom methods.
Cognitive training	Cognitive training refers to interventions that aim to target and improve core cognitive functions, such as attention, perception, and working memory. Examples include brain training, neurofeedback, and noninvasive brain stimulation.
Collaborative (peer) learning	This involves situations where trainees are trained in groups and learn vicariously via observation and peer interaction.
Computer-based training	Computer-based training is any course of instruction that is delivered on a computer; the course may be installed as software on a single computer or delivered over the internet as web-based training.
Crawl-walk-run	This philosophy espouses that following a graduated training sequence from easy to more complex tasks is most effective.
Debriefs	The use of debriefs (or after action reviews) is common within the military to reinforce learning points and identify areas for improvement after training exercises or activities (Salas et al, 2008).

Distributed learning	Distributed learning is an instructional approach that allows trainees, instructors, and training content to be in different locations so learning can occur anywhere and anytime.
Event-based approach to training	This approach involves scripting specific events to trigger desired behaviors during a training activity (Fowlkes et al, 2009).
Identical elements theory	This theory espouses that the more similar elements of a training activity are to elements in the actual performance environment, the greater the odds of transfer.
Intelligent tutoring	Intelligent tutoring systems are computer-based systems that can deliver training content and provide customized instruction and feedback to trainees usually without the need for a human instructor.
Live Virtual and Constructive (LVC) Simulation	Live simulation involves real people operating real equipment and systems (e.g., field exercises with soldiers operating real vehicles). Virtual simulation involves real people operating simulated systems (e.g., real pilot in flight simulator). Constructive simulation involves computer generated forces which may involve a human-in-the-loop or be fully autonomous (e.g., simulation for wargaming).
Modeling	Modeling is an instructional strategy in which the teacher demonstrates a new concept or approach to learning and students learn by observing.
Simulation-based training and games	Simulation-based training includes the use of LVC simulation methods (see above), immersive technologies, and role-playing / rehearsals. Games are a form of simulation usually typified by competitive elements, scenarios, as well as defined rules, goals, and end-states.
Team training	This includes methods such as cross-training, coordination training, error management training, and crew resource management for teams (see Hoggan & Temby, 2017; Lacerenza et al, 2018 for reviews).
Train as you fight	This philosophy espouses that training conditions should be as realistic as possible to maximize learning and training transfer.

What methods can be used to enhance cognitive skills?
Cognitive skills include paying attention, remembering, processing, analyzing, comprehending, evaluating, problem-solving, reasoning, and decision-making. Such skills are essential for military personnel in various roles. There is evidence that cognitive skills can be developed using several training and education approaches, although the effectiveness of different approaches varies. Some common approaches include: memory aids (or mnemonics), performance feedback and debriefs, critical thinking and reflection training, meditation/mindfulness, cognitive walkthroughs and visualization, deliberate practice, and simulations and games. Additional methods for enhancing cognitive performance include cognitive training, neurofeedback, supplements and ergogenic aids (see Sala and Gobet (2018) and Strobach and Karbach (2016) for an overview of cognitive enhancement research).

3. TRAINING DEVELOPMENT

The Development phase involves developing training scenarios, defining and developing performance measures, and piloting or trialing the training solution. In this section we discuss some of the key issues in this phase.

What are important considerations when developing a training program?
Once the instructional approach has been determined, there are a couple of key considerations in the development of a training program. The first is the development of training scenarios. These scenarios should be realistic and engaging for trainees, and of varying difficulty to support skill progression to desired levels. The second key consideration is the development of performance measures, including the construction of performance measurement tools (e.g., observation checklists), and the development of success criteria. In our experience, bespoke performance measures often need to be developed to ensure they are sufficiently sensitive to performance differences (e.g., Likert-type rating scales). The development of bespoke measures is best done in consultation with military subject matter experts in the early stages of training design. During this stage, it is also important for researchers to clarify expectations about what is meant by a "trial". In our experience it can vary from a 'demonstration' to 'proof of concept', to 'full-scale evaluation' or a 'controlled experiment'.

What is simulator fidelity and how much do I need?
Fidelity refers to the extent to which a training environment, such as a simulator, represents different aspects of the criterion or real-world environment (Hays, 1980). There are three main types of simulator fidelity: physical, functional, and psychological. Physical fidelity refers to the degree to which a simulator 'looks, sounds, and feels'

like the real device. Functionality fidelity refers to the degree to which the behavior and response of the simulator to the actions of the operator match the real device. Psychological fidelity is the degree to which simulator tasks can trigger the same behaviors that are required on-the-job (Galanis, Stephens, & Temby, 2013). While it is commonly believed that high-fidelity simulations lead to better training, there is some evidence that lower-fidelity simulators can have training benefits (e.g., Stephens et al, 2011; Whitney, Fidock, & Ferguson, 2012). However, the relationship between simulator fidelity and training outcomes is complex; for example, researchers have found mixed results for the benefits of simulated motion (a type of physical fidelity) on task performance and training outcomes (e.g., Brunner, Fidock, Temby, & Whitney, 2015; Hibbard, Fidock, & Whitney, 2015). There is no magic formula for simulator fidelity, however, more physical fidelity often equates to greater financial cost. Further guidance on determining simulator fidelity requirements can be found in Estock et al (2006) and Stone (2012).

4. TRAINING IMPLEMENTATION

During the Implementation phase, the training program and instructional methods are delivered. In this section we outline key issues for consideration when delivering training, including factors that can influence skill acquisition during training, as well as skill retention following training.

What are some key principles to consider when delivering training?
There is no single method of training that works in every situation. However, several best-practice principles have been identified for training delivery (e.g., Riddle et al, 2008; Salas & Cannon-Bowers, 2001; Siebert et al, 2010; Wampler et al, 2006). These include ensuring: (1) relevant information or concepts to be learned are presented, (2) training follows a proven sequence, (3) the KSAs to be learned are demonstrated, (4) pre-requisite knowledge and skill levels are assessed before commencing new training, (5) clear, measurable, and achievable objectives are established for the KSAs to be trained, (6) there are opportunities for trainees to practice the skills, (7) a learner-centric, problem-solving approach is employed, (8) feedback is provided to trainees during and after practice, (9) a mix of instructional methods and techniques are employed, and (10) the KSAs/tasks are trained in the context within which they will be executed. For some lessons learned from delivering training in military settings, we recommend the papers by Siebert et al (2010) and Wampler et al (2006).

What are important considerations when delivering simulation-based training?
There are several issues to consider when implementing simulation-based training (SBT). Salas and Burke (2002) suggest that for SBT to be effective: (1) instructional features should be embedded within the simulation, (2) scenarios should be well-designed and contain opportunities for performance measurement and diagnostic feedback, (3) it should provide guided learning experiences, and (4) simulation fidelity should be matched to training requirements. In addition, some other considerations include: (a) using valid and reliable performance measures, (b) using story-boarding and prototype testing, (c) assessing trainee and instructor expectancies of SBT, (d) measuring participants' prior experience with simulation, (e) minimizing sensory overload in virtual environments, and (f) managing the risk of simulator sickness. By addressing these issues, SBT will likely be more effective. For more information on best practice performance measurement in SBT, see Salas et al (2009).

Why should we use simulation for training? Isn't live training always better?
In general, where trainees can apply the required knowledge and skills in the actual training environment using the actual equipment (i.e., live training) this is preferable. However, synthetic training environments can provide military personnel with training experiences that may be too unsafe or costly to undertake using live training. Synthetic training can sometimes generate higher levels of complexity and realism than may be possible in live training, as in the case for close air support team training (see for example, Stephens et al, 2011), distributed mission training with joint and coalition forces, and training for multi-domain operations. Using simulation can also allow trainees to rehearse KSAs more often, replay scenarios, gain exposure to a broad range of training experiences, and develop competency to high levels before conducting live training, thereby maximizing training resources.

What is the optimal mix of live and simulation training?
This question is often asked by practitioners seeking to optimize training schedules and allocate a given percentage of time to using simulation. The optimal mix will be on a case by case basis and depends largely on (i) the tasks to be trained, (ii) the cost of live training, and (iii) the training objectives. For example, using live rounds for tank gunnery training is significantly more expensive than using smaller caliber live rounds for rifle marksmanship training. As a result, it makes more economical sense to conduct a higher proportion of tank gunnery training (say

95%) in the simulator than rifle marksmanship training (say 50%). However, if the training audience is an elite infantry forces then the requirement for realistic training may trump the costs involved. In this latter case, marksmanship tasks may be conducted almost exclusively using live training. Where the economic benefits for using simulation are less obvious, the optimal mix of simulation and live training may be explored by trialing different combinations of the two training methods and analyzing which ratio achieves better training outcomes on specific criteria (e.g., time to achieve competency, resources required). Furthermore, simulation can provide better training where opportunities for live training may not be available; one example of this would be training for future operational scenarios involving the use of capabilities (such as autonomous system) that have yet to be fielded, or where the opportunity to train in the real environment is limited by geography (e.g., coalition mission training). In these cases the cost involved may be secondary to the training requirement to prepare forces for missions.

What can be done to ensure skill learning and acquisition during training?
Delivering training does not guarantee learning will occur. Research has found that skill learning and acquisition is influenced by several key factors (e.g. Arthur et al, 1998) including: trainee ability, the amount of training time available, the frequency of training (spaced practice is generally better than massed practice), the complexity of the task(s), trainee goal orientation (performance vs mastery focus), trainee motivation and engagement levels, the type and frequency of feedback provided, and the quality of instruction. Consequently, when designing and implementing training programs, it is important to take these factors into account to maximize skill learning and acquisition.

How can skill retention be facilitated?
The retention and maintenance of skills at desired levels is important for military readiness. Skill retention is affected by several factors (Arthur et al, 1998), including task factors (e.g., task complexity, type of skill), training factors (e.g., retention interval, amount of training), measurement factors (e.g., assessment format), and trainee factors (e.g., cognitive ability, personality). Research has shown that skill retention is enhanced by increasing the amount of task repetition, testing during training, use of job aids, spaced practice, and using recognition (versus recall) measures of performance. In contrast, skill loss is more likely with lack of practice, especially for more complex skills. More complex tasks that involve a greater number of steps are also more susceptible to skill fade. See Bryant and Angel (2000) for an overview of military skill retention.

How often should training be delivered to maintain skills?
The frequency of training will depend on the perishability of the skills involved, the cost of training, the criticality of keeping skills at certain levels of readiness, and the opportunity to train. For safety critical and complex tasks (e.g., medical procedures) more frequent refresher training (perhaps using simulation) may be essential to prevent skill decay. Some guidance about refresher or maintenance schedules may be gleaned from training records and published studies. In addition, tools such as the User's Decision Aid (Rose et al, 1985) can provide estimates of skill fade for groups of individuals. This tool makes a prediction of skill loss based on the amount of time since the last task performance, task type and difficulty, and the quality of any job aids.

5. TRAINING EVALUATION AND TRANSFER

Training evaluation is important for assessing how well training achieved desired learning outcomes, including the relevant KSAs. In this section we outline some key principles, methods, and measures for evaluating training.

What are some key principles to consider when evaluating training?
Training evaluation is designed to measure learning outcomes and identify what worked. There are several key principles to consider when evaluating training. These include: (a) start the process of evaluation early, (b) treat every evaluation as unique, (c) use multiple methods to assess outcomes, (d) adopt a comprehensive and rigorous approach, and (e) treat evaluation as a continuous and iterative process (Riddle et al, 2008). Implementing these principles may not always be achievable in military settings; compromises may need to be made to training evaluations to give way to practical constraints (e.g., working around training schedules, only being able to access personnel for short periods).

What frameworks and tools can I use to evaluate training outcomes?
One of the most commonly used training evaluation frameworks is the Kirkpatrick model (Kirkpatrick & Kirkpatrick, 2006). The model consists of four levels: (1) Reactions, (2) Learning, (3) Behavior, and (4) Results. Reactions are assessed by asking trainees and instructors what they thought about the training, usually with self-

report surveys. Learning is measured by examining the extent to which trainees have acquired the relevant KSAs; this is typically measured by knowledge and skill-based tests (e.g., written exam, time to achieve competency). Behavior is measured by assessing trainees' application of learned competencies to the job environment (i.e., transfer of training); this is often measured by direct observation or with competency-based assessments. Results are the organizational impact of training outcomes, such as reduced costs, improved quality and efficiency; such data might be derived from cost-benefit analyses, and longitudinal analyses of training records. In addition to the Kirkpatrick model, Kraiger, Ford, and Salas (1993) suggest that trainee learning is multi-dimensional and should be measured in terms of cognitive (e.g., declarative knowledge, reactions), skill-based (e.g., proficiency level) and affective outcomes (e.g., post-training attitudes and motivation). The MST-READI methodology (Riddle et al, 2008) and website (http://www.cws-i.com/MST-READI/site-map.htm) contains useful information and tools to support training evaluations.

How should I design a training evaluation trial in a military setting?

There are several important issues to consider when designing a training evaluation trial. These are briefly summarized in Table 2 and cover issues ranging from trial participants, evaluation design, logistics, data analysis, and reporting of outcomes. We recommend that practitioners allow plenty of time to address these issues in the planning phases and to engage key stakeholders throughout the entire process. For an excellent article on the challenges of conducting training evaluation in military settings, see Salas, Milham, and Bowers (2003).

Table 2. Key considerations when designing a trial to evaluate training in military settings

Issue	Description
Personnel	Who will take part in the trial? Trainees, instructors, and support staff? Are the participants representative of the target population?
Sample size	How many participants will be needed to generate meaningful results?
Duration	How much time will be needed / available to conduct the evaluation?
Resources	What equipment, resources and facilities will be required to conduct the trial? Will some of this equipment have to be borrowed? Is there a reliance on a third party to provide training tools? Will internet connectivity be required/available?
Trial Design	How will the trial be designed? Is the design appropriate for the research objectives? What will the participants be required to do? Will the design involve treatment and control groups? Will all participants receive the training if it is effective? If a control group is used, what will they do? How will the 'new' training intervention be included in the trial? Will trainees only use the new training method, or will there be a blend of old and new modes of instruction? What is the minimal viable design that will produce useful insights/outcomes?
Data Collection & Measurement	What data will be collected to assess performance and evaluate training/learning? How many data collection points will be involved? Will a follow-up assessment of KSAs be included? How easy will it be to follow-up with participants?
Data Analysis	How will the data be analyzed? Does the evaluation team have the skills to conduct appropriate (statistical) analyses and interpret the results?
Interpretation & Reporting	What is the possible range of outcomes from the trial? How will null results be interpreted and discussed with stakeholders? Will there be opportunities to conduct follow-on trials as part of a longer-term research effort? What timeframe will be negotiated to report the research outcomes to military clients?

Which evaluation design should I use to evaluate training?

There are several evaluations designs that are often used in training research, which are briefly described below, including their pros and cons for training evaluations.

- Observational/Non-experimental Designs: This type of design is useful for initial scoping studies and in situations where allocating military personnel to different groups may not be possible. It may also be used in qualitative research to study the influence of key variables (e.g., quality of instruction) and generate hypotheses. A limitation of this design is that it cannot be used to establish causal relationships about training factors.
- Post-Test Only No Control Group Design: This type of design is useful for determining whether trainees have reached a target level of performance at the end of training. It does not allow any assessment of the amount of change in performance from baseline levels to be determined, or the specific contribution of training factors.
- Pre-and Post-Test No Control Group Design: This type of design can identify whether a change in performance has occurred, but any observed changes cannot be solely attributed to training.

- Pre-and Post-Test with Control Group Design: This type of design is considered a gold standard for training evaluation, particularly when it also includes randomization of subjects to different groups. This type of design provides information about changes in trainee levels of performance, as well as the efficacy of different training methods, including no training and training as usual. This design may not always be possible in field settings for practical (e.g., number of personnel available) and ethical reasons (e.g., exposing untrained personnel to unacceptable risk). The use of unequal sample sizes across groups, and proxy measures of performance are ways to address these issues (see Salas, Milham, & Bowers, 2003). For more guidance on conducting defense experiments, readers are referred to the Guide for Understanding and Implementing Defense Experimentation (GUIDEx) produced by The Technical Cooperation Program (TTCP, 2006).

What is transfer of training and what can be done to ensure it occurs?

Transfer of training refers to the degree to which the knowledge and skills acquired in training are applied, generalized, and maintained in the workplace environment (Galanis, Stephens, & Temby, 2013). Several factors influence training transfer, including trainee characteristics (e.g., ability, personality, motivation, pre-training attitudes), training design (e.g., training content, sequencing, quality of instruction), and the work environment (e.g., opportunities for practice, supervisor support) (Blume et al, 2010). The degree of overlap between task elements in the training and job environment can also contribute to training transfer, as suggested by identical elements theory. The post-training environment is also critical for training transfer. It is important that supervisors provide opportunities for trainees to apply new knowledge and skills in the workplace, and to provide trainees with job aids and reinforcement. For a recent overview of transfer of training research see Ford, Baldwin, and Prasad (2018).

How will I know if training has been effective?

The short answer is whether the training (research) achieved its stated objectives. Using the Kirkpatrick and Kraiger frameworks can help researchers define success. For example, did trainees like the training (*reactions*), did trainees acquire new KSAs (*learning*), were these KSAs applied to the work environment (*behavior*), and did the training result in benefits at the organizational level (*results*). It should be noted, however, that positive reactions to training may not always be the goal, and that measuring organizational results may not be possible in the short-term. There is a need for more cost-effectiveness studies to inform organizational results and return on investment (e.g., Stephens, Temby, McLean, & Hansen, 2013). Further guidance on return on investment considerations can be found in Cohn and Fletcher (2010) and Whitney (2019). Finally, it is important to highlight that training is only effective if there are no negative consequences or side effects.

How should I evaluate simulator effectiveness?

The effectiveness of a simulator as a training aid should be assessed against training requirements. Any simulator evaluation should also consider the broader context in which the simulator will be used. This includes the training curriculum, the role of the instructor, facility and maintenance costs, and the extent to which the simulator can support the rehearsal of KSAs. Researchers and practitioners should also be wary of what Rolfe and Caro (1982) refer to as deceptive measures of simulator effectiveness; namely (1) user opinion ('if trainees like it, it must be a good training device'), (2) realism ('if it is like the real thing, it must be effective'), and (3) utilization ('if it used a lot, it must be a good training device'). These statements are based on assumptions and stakeholders should give priority to more objective data when assessing the effectiveness of a simulator. This includes objective data of the underlying models in a simulator, as might be derived from validation, verification, and accreditation (VV&A) testing. Practitioners should not assume that these models are correct, and always consider the possibility that negative training may be occurring.

How can researchers evaluate manufacturers' claims about training devices?

There are many devices on the market which claim to be effective for training. Such devices include virtual and augmented reality systems, cognitive training apps, and intelligent tutoring systems. Manufacturers may highlight device features such as the level of fidelity and range of scenarios available; however, these features will not necessarily lead to better training outcomes. Manufacturers may also highlight user feedback, usage rates, and device realism as measures of effectiveness (Rolfe & Caro, 1982). When assessing manufacturers claims, researchers should look for evidence of: (1) manufacturers engaging with training specialists and end-users during the development process, (2) relevant performance and feedback measures embedded in the device, (3) correlations between trainee performance on the device and real-world performance (see Temby & Stephens, 2015 for example), (4) device evaluations being conducted with end-users under realistic conditions, and (5) independent evaluations of

these devices being conducted and the findings being reported in peer-reviewed publications. Where such evidence is not found or made available, researchers should remain skeptical of any effectiveness claims.

What should I do if my training intervention doesn't produce the desired outcome?
While it is generally accepted that 'training works' (Salas & Cannon-Bowers, 2001), not every training intervention will produce the desired outcome of improving KSAs, and there may even be unintended negative consequences. In such cases, researchers should take the time to understand any confounds and limiting factors that may have contributed to the outcomes. While there are many reasons why training may not be effective, it can be helpful to categorize them into trainee characteristics, training design/content (including technology), and organizational factors. Once the reason(s) have been identified, and there may be more than one, it is recommended that they be discussed with key stakeholders and ways to address them be explored. If a good relationship has been established with military stakeholders and expectations managed about possible training outcomes, there should be no real surprises and null or negative results may be treated as 'lessons learnt' rather than failure.

WORKING IN TRAINING AND SIMULATION RESEARCH

During our career we have acquired a range of experience and learned valuable lessons from working with military personnel. In this section we highlight key skills that we believe are important for practitioners, offer advice for researchers working with military personnel, and suggest some areas for future training and simulation research.

What attributes, skills, and qualifications are needed to work in training and simulation research?
We believe training researchers need to possess several key attributes and skills including: knowledge of training and simulation, knowledge of human factors, good knowledge of research methods and experimental designs, a good understanding of psychometrics and data analysis techniques, good communication skills including the ability to sell science to different stakeholders, willingness to be flexible, interest in technology, an analytical mindset, and a curiosity for how things work. The ability to remain impartial and report unbiased results are also key attributes; these are particularly important given the significant investments made by defense forces into training and simulation, and the fact that researchers may be asked to provide advice to inform capability acquisition decisions.

There is no single career path to becoming a training researcher or practitioner. Such personnel have a variety of academic backgrounds including human factors/ergonomics, cognitive science, cognitive psychology, experimental psychology, industrial and organizational psychology, systems engineering, education, and instructional design. In the military, training staff usually gain on-the-job experience and obtain qualifications from internal courses and tertiary education. For an interesting read on a career in training research within defense, see Andrews (2017).

What specific advice would you give to training researchers working in military settings?
As noted by Siebert et al (2010), conducting training research 'in the wild' can be challenging. Drawing on our years of experience, we offer the following specific advice for practitioners when working in military settings.
- Adopt a systems perspective of training problems and seek the views of military personnel at different levels
- Spend time talking with military personnel and instructional staff about their training needs and issues
- Try to use military terminology and avoid jargon when communicating with military personnel
- Be clear about the goals of your research and how the outcomes may benefit military commanders
- Use concrete examples from your research experience when communicating with military personnel; this may help to establish rapport and build their confidence in your abilities
- Discuss research ethics issues such as informed consent with military commanders early in research planning
- Identify a military champion for your research and develop a working relationship with them
- Obtain commitments and approvals from commanders in writing (where possible); this will help facilitate hand-overs when there are personnel changes during a project
- Be flexible and willing to adapt your research plans; be clear about what is negotiable and non-negotiable
- Explore possible outcomes of training evaluations with commanders in the planning stages; this may help to facilitate shared ownership of the research questions and outcomes
- Spend time reading military doctrine and other training documentation
- Invest time in understanding how current training is delivered before discussing new training solutions
- Conduct task analyses and develop performance measures in conjunction with military training staff; where possible use existing measures if they are suitable for research purposes

- Spend time educating military stakeholders about the merits of different training evaluation designs
- Consider how your research will potentially save time, improve performance, and reduce training costs
- Ask questions and show genuine interest in how military training is conducted and factors impacting on it
- If training evaluations yield 'null results' try to diagnose the reasons and develop solutions to address them
- Remember that military personnel have competing priorities outside your research; be prepared to chase down information and have back-up plans to keep your project on track
- Develop a schedule for data collection with military units including post-training measures of effectiveness
- Try to gain support from commanders for an iterative approach to training evaluations (versus one-off studies)
- Have a defined end-state for your research; be willing to cease your research if necessary
- Secure funding for travel to cover visits and data collection activities on military bases
- Develop an agreed plan with military stakeholders about when and how you will report your research findings

What are some important areas for future training and simulation research?

As the nature of military operations change and technology continues to advance, defense forces will seek to continually adapt their training methods to keep pace. There are several enduring and emerging areas that we believe are important for future research to address in support of military training. These include:

1. *Mobile technologies*: Mobile technologies, such as intelligent tutoring systems (e.g., GIFT), apps and wearables, and immersive simulation systems (such as virtual and augmented reality) will likely continue to be used for training. Understanding the use cases and efficacy of these technologies will be important for deciding when, how, and where they might be best employed.
2. *Personalized training*: There is an increasing focus on tailored training for defense personnel. Understanding the specific methods that work for individuals will be important to inform the development of personalized training protocols. Wearable sensors, embedded metrics, and data analytics will also be important in this area to guide individualized performance assessment, monitoring and feedback.
3. *Multi-modal training*: There is growing interest in exploring the potential benefits of multi-modal training methods (where different training interventions are used in combination; such as physical and cognitive training) for enhancing human performance. Further research is needed to examine the feasibility of such approaches in military settings and the additional performance gains that might be achieved by using them.
4. *Longitudinal field studies*: There is a need for longitudinal field studies to monitor changes in the KSAs and performance of military personnel over time. While such studies are logistically challenging, they will help us better understand training effectiveness, skill retention rates, and inform refresher training schedules.
5. *Human-autonomy team training*: The use of autonomous systems in military operations is expected to increase in future and operators will need to be trained in how to employ them effectively. Research examining training requirements for these systems will help to inform the development of appropriate training programs. We envisage that simulation will play a key role in these training solutions.
6. *Cost-effectiveness*: As defense budgets become tighter, there will be a greater demand for more economical training. Studies examining the cost-effectiveness of different training solutions are needed to help inform decisions regarding the best use of limited resources.
7. *Cognitive enhancement*: Enhancing the cognitive performance of warfighters for mission-specific tasks is of ongoing interest to defense forces (as evidenced by this year's conference theme). Cognitive enhancement methods such as neurofeedback, cognitive training, and brain stimulation show promise for military application. However, research examining the short- and long-term benefits and risks and ethical implications of these interventions is needed to better inform their use with military populations.

CONCLUSION

Training is a fundamental input to capability for defense forces. With significant investments made into training each year, it is important that training is well-designed, conducted and evaluated to achieve desired outcomes. This paper has outlined key issues to consider when undertaking training and simulation research in military settings. In addition, we have highlighted skills we believe are needed by training researchers, provided advice for working with military personnel, and outlined some directions for future training research. While much of the contents of this paper may be familiar to those who work in the field, we hope that this primer provides a useful resource for those who may be new to the area, and anyone interested in maximizing the effectiveness of training.

ACKNOWLEDGEMENTS

REFERENCES

[1]   Alvarez, K., Salas, E., & Garofano, C.M. (2004). An integrated model of training evaluation and effectiveness. *Human Resource Development Review*, *3*(4), 385-416.

[2]   Andrews, D. H. (2017). Training research and development: Retrospective of a career in the Defense Department. *Education Review*, *24*.

[3]   Arthur Jr, W., Bennett Jr, 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.

[4]   Australian Army (2017). *Land Warfare Procedures - General 7-1-2 The Instructor's Handbook 2017*. Canberra, ACT.

[5]   Bell, B.S., Tannebaum, S.I., Ford, J.K., Noe, R.A., & Kraiger, K. (2017). 100 years of training and development research: What we know and where we should go. *Journal of Applied Psychology, 102*(3), 305-323.

[6]   Best, C., Galanis, G., Kerry, J., & Sottilare, R. (2013). *Fundamental Issues in Defense Training and Simulation*. Ashgate Publishing.

[7]   Blume, B.D., Ford, J.K., Baldwin, T.T., & Huang, J.L. (2010). Transfer of training: A meta-analytic review. *Journal of Management*, *36*(4), 1065-1105.

[8]   Bollard, L. M., Kerry, J.T., Whitney, S. J., & Fidock, J. J. T. (2015). Digital literacy in the Australian and New Zealand Defence Forces: Current levels and implications. *In 2015 Fall Simulation Interoperability Workshop*.

[9]   Booher, H.R. (2003). *Handbook of human systems integration* (Vol. 23). John Wiley & Sons.

[10]   Branch (2009). Branch, R. M. (2009). Instructional design: The ADDIE approach (Vol. 722). Springer Science & Business Media.

[11]   Brunner, C., Fidock, J.T., Temby, P., & Whitney, S.J. (2015). *Transfer of Training to the Real-World: A Study of Simulation-Based Training for a mobile ICT Device*. Proceedings of the Simulation, Technology and Training Conference. Adelaide, SA.

[12]   Bryant, D.J., & Angel, H. (2000). *Retention and fading of military skills: Literature review*. Humansystems Inc, Guelph Ontario.

[13]   Christensen, L. B., & Tremblay, L.R. (2013). Generational Learning Differences: Myth or Reality. In *Fundamental Issues in Defense Training and Simulation* (pp. 21-30). Ashgate.

[14]   Cohn, J., & Fletcher, J. (2010). What is a Pound of Training Worth? Frameworks and Pracical Examples for Assessing Return on Investment in Training. In I/ITSEC; Orlando: FL.

[15]   Department of Defence (2016). *The Systems Approach to Defence Learning*. Canberra, ACT.

[16]   Estock, J.L., Alexander, A.L., Gildea, K.M., Nash, M., & Blueggel, B. (2006). A Model-Based Approach to Simulator Fidelity and Training Effectiveness. *In I/ITSEC*, Orlando: FL.

[17]   Ford, J.K., Baldwin, T.T., & Prasad, J. (2018). Transfer of training: The known and the unknown. *Annual Review of Organizational Psychology and Organizational Behavior*, *5*, 201-225.

[18]   Fowlkes, J., Dwyer, D.J., Oser, R.L., & Salas, E. (1998). Event-based approach to training (EBAT). *The International Journal of Aviation Psychology*, *8*(3), 209-221.

[19]   Galanis, G., Stephens, A., & Temby, P. (2013). *What is Transfer of Training? And what does it have to do with simulators?* In C Best (Eds). Fundamental Issues in Defense Training and Simulation. Ashgate Publishing.

[20]   Hays, R. T. (1980). *Simulator fidelity: A concept paper* (No. ARI-TR-490). Army Research Institute for the Behavioral and Social Sciences: Alexandria, VA.

[21]   Hibbard, S. J., Whitney, S. J., Carter, L., Fidock, J.J., Temby, P., & Thiele, L. (2016). Making Virtual Sense: Display Type and Narrative Medium Influence Sensemaking in Virtual Environments. In *Intersections in Simulation and Gaming* (pp. 222-236). Springer, Cham.

[22]   Hoggan, B., & Temby, P. (2017). *Enhancing collective training in the Australian Army: Insights from the team development literature*. Paper presented at Defence Human Sciences Symposium, 6-7 November 2017. Adelaide, South Australia.

[23]   Hoover, M., Miller, J., Gilbert, S., Winer, E., & Davies, P. (2018). Evaluating Augmented Reality Assembly Instructions Delivered via Microsoft HoloLens. In: *I/ITSEC,* Orlando, FL

[24]   Kirkpatrick, D.L., & Kirkpatrick, J. D. (2006). *Evaluating training programs: The four levels (3rd edition).* San Francisco, Berrett-Koehler.

[25]   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*(2), 311.

[26]   Lacerenza, C. N., Marlow, S. L., Tannenbaum, S. I., & Salas, E. (2018). Team development interventions: Evidence-based approaches for improving teamwork. *American Psychologist*, *73*(4), 517.

[27]   National Training and Simulation Association (2011). *A Primer on Modeling and Simulation*. Publisher: NTSA. Arlington, VA. Retrieved on May 31, 2019 from http://www.trainingsystems.org/-/media/sites/ntsa/homepage/miscellaneous/ms-primer.ashx?la=en

[28]   Orvis, K.A., Moore, J.C., Belanich, J., Murphy, J.S., & Horn, D.B. (2010). Are soldiers gamers? Videogame usage among soldiers and implications for the effective use of serious videogames for military training. *Military psychology*, *22*(2), 143-157.

[29]   Phillips, J.J. (2012). *Handbook of training evaluation and measurement methods*. Routledge.

[30]   Prensky, M. (2001). Digital natives, digital immigrants part 1. *On the horizon*, *9*(5), 1-6.

[31]   Riddle, D.L., Chapman, R.J., Pike, W.Y., & Norfleet, J. (2008). MST-READI: Practical Guidance for Military Medical Simulation Training Evaluation Research. *Journal of Interactive Instruction Development*, *20*(4), 21-26.

[32]   Rolfe, J.M., & Caro, P.W. (1982). Determining the training effectiveness of flight simulators: Some basic issues and practical developments. *Applied ergonomics*, *13*(4), 243-250.

[33] Rose, A., Radtke, P., Shettel, H. & Hagman, J. (1985). User manual for predicting military task retention. (AIR FR 37800) U.S. Army Research Institute for Behavioral & Social Science: Alexandria, VA.

[34] Sala, G., & Gobet, F. (2018). Cognitive training does not enhance general cognition. *Trends in cognitive sciences (in press)*.

[35] Salas, E., & Burke, C.S. (2002). Simulation for training is effective when. *BMJ Quality & Safety, 11*(2), 119–120.

[36] Salas, E., & Cannon-Bowers, J.A. (2001). The science of training: A decade of progress. *Annual Review of Psychology*, *52*(1), 471-499.

[37] Salas, E., DiazGranados, D., Klein, C., Burke, C.S., Stagl, K.C., Goodwin, G.F., & Halpin, S.M. (2008). Does team training improve team performance? A meta-analysis. *Human Factors*, *50*(6), 903-933.

[38] Salas, E., Milham, L.M., & Bowers, C.A. (2003). Training evaluation in the military: Misconceptions, opportunities, and challenges. *Military Psychology*, *15*(1), 3-16.

[39] Salas, E., Rosen, M.A., Held, J.D., & Weissmuller, J.J. (2009). Performance measurement in simulation-based training: A review and best practices. *Simulation & Gaming*, *40*(3), 328-376.

[40] Salas, E., & Stagl, K.C. (2009). Design training systematically and follow the science of training. *Handbook of principles of organizational behavior: Indispensible knowledge for evidence-based management*, 59-84.

[41] Salas, E., Tannenbaum, S.I., Kraiger, K., & Smith-Jentsch, K.A. (2012). The science of training and development in organizations: What matters in practice. *Psychological Science in the Public Interest*, *13*(2), 74-101.

[42] Salas, E., Wilson, K.A., Priest, H.A., & Guthrie, J.W. (2006). Design, delivery, and evaluation of training systems. *Handbook of Human Factors and Ergonomics*, *3*, 472-512.

[43] Salvendy, G. (2012). *Handbook of human factors and ergonomics*. John Wiley & Sons.

[44] Siebert, M., Diedrich, F., MacMillan, J., & Riccio, G. (2010). *Training research in the wild*. In I/ITSEC, Orlando: FL.

[45] Stephens, A., Crone, D., Temby, P., Best, C., & Simpkin, G. (2011). *Using Synthetic Environments to Enhance Close Air Support Training*. Proceedings of the 2011 Simulation, Technology and Training Conference. Melbourne, Victoria.

[46] Stephens, A., Temby, P., McLean, & Hansen, J. (2013). *Simulation in Training Workshop 2012*. DSTO-GD-0754. Defence Science and Technology Organization. Edinburgh, South Australia.

[47] Stone, R. J. (2012). *Human Factors Guidelines for Interactive 3D and Games-Based Training Systems Design*. 2nd ed. UK, Bimingham.

[48] Strobach, T., & Karbach, J. (Eds.). (2016). *Cognitive training: An overview of features and applications*. Springer.

[49] Temby, P., & Stephens, A. (2014). *Development and Evaluation of a Novel Method for Basic Marksmanship Training on an Australian Army Course. DSTO-TR-2951*. Defence Science and Technology Organization. Edinburgh, South Australia.

[50] Temby, P., & Stephens, A. (2015). *A Case Study of Construction Equipment Training: From the Simulator to the Real-World*. Proceedings of the 2015 Simulation, Technology and Training Conference. Adelaide, South Australia.

[51] TTCP (2006). Guide for Understanding and Implementing Defense Experimentation (GUIDEx).

[52] Wampler, R. L., Dyer, J. L., Livingston, S. C., Blankenbeckler, P. N., Centric, J. H., & Dlubac, M. D. (2006). *Training lessons learned and confirmed from military training research*. Northrop Grumman Mission Systems, Columbus GA.

[53] Whitney, S.J. (2019). *Key considerations in conducting Army training systems research*. DST-Group-GD-1035. Defence Science and Technology Group. Edinburgh, South Australia.

[54] Whitney, S. J., Fidock, J.J., & Ferguson, N. (2012). Assessing the effectiveness of simulation-based counter-IED training. *Journal of Battlefield Technology*, *15*(1), 57.

[55] Whitney, S.J., Fidock, J.J.T., & Hibbard, S. (2016). *Virtual and mixed reality head mounted display systems: Implications for future land force training.* Proceedings of the Future Land Force Conference, 5-7 September 2016. Adelaide, Australia.

[56] Whitney, S.J., & Stephens, A. K. (2014). *Use of Simulation to Improve the Effectiveness of Army Welding Training* (No. DSTO-TR-2997). Defence Science and Technology Organisation. Fisherman's Bend, Victoria.

[57] Whitney, S., Temby, P., & Hoggan, B. (2018). *Driving Under Load: Measuring Instructor Workload on a Military Heavy Vehicle Training Course*. Proceedings of the 2018 Human Factors and Ergonomics Society Association (HFESA) Conference. Perth, Western Australia.

[58] Whitney, S.J., Temby, P., & Stephens, A. (2014). A review of the effectiveness of game-based training for dismounted soldiers. *The Journal of Defense Modeling and Simulation*, *11*(4), 319-328.