ACT ARC Recommendation 21-11

eLearning Guidance and Systematic Design Methodologies

I. Submission

The recommendations below were submitted by the Effectiveness of Knowledge Training Workgroup (EKT WG) for consideration by the Air Carrier Training Aviation Rulemaking Committee (ACT ARC) Steering Committee at its April 28, 2021, meeting. The ACT ARC Steering Committee adopted the recommendations, and they are submitted to the Federal Aviation Administration (FAA) as ACT ARC Recommendation 21-11.

II. Statement of the Issue

Recent developments in the technology, application, and implementation of what is commonly known as "electronic learning" (eLearning) have affected the instructional delivery and the effectiveness of courseware and instructional strategies in aviation. Air carriers operating under the provisions of 14 CFR parts 121 and 135 ("part 121 and part 135 air carriers" or "air carriers") and training centers operating under 14 CFR part 142 ("part 142 training centers" or "training centers") are increasingly turning to eLearning approaches to fulfill their training requirements consistent with these parts. The ACT ARC formed the EKT WG to address industry, academic, and FAA interest in examining how to best incorporate eLearning approaches and technologies into current FAA regulatory guidance and how to ensure the effectiveness of training curricula that employ eLearning. The EKT WG's specific focus is non-instructor led training delivered through electronic means, which is one component of eLearning currently used in air carrier training.[1]

One way to help ensure effectiveness of eLearning curricula is the use of a systematic design methodology. A systematic design methodology consists of a documented process used to develop training and to measure its effectiveness. Systematic design methodologies may vary significantly, but generally include elements or characteristics such as analysis of learning objectives and use of tools to measure effectiveness. An acceptable systematic design methodology should generally include some key elements:

- analysis of training needs and objectives;
- design and development of training to achieve those objectives;
- a structured, planned implementation of the program; and
- a methodical evaluation of its effectiveness to allow for a closed feedback loop leading to continuous improvement of the training.[2]

[1] eLearning refers to the use of computer, internet, web-based, and mobile technologies to deliver learning solutions. eLearning is not, in general usage, synonymous with distance learning, and may refer to instructor-led learning or self-study. For the purposes of this Proposed Recommendation, however, eLearning refers only to non-instructor led training delivered through electronic means, unless context indicates otherwise.

[2] The term "ADDIE" is often used to refer to these elements (Analysis, Design, Development, Implementation, and Evaluation) collectively. "ADDIE" may also refer to methodologies incorporating these or similar elements.

In the absence of a systematic design methodology, the selection of instructional methods and development and evaluation of the effectiveness of training will yield inconsistent results. While use of a systematic design methodology will not guarantee success, research indicates it will increase the likelihood of success.

The EKT WG has identified the following issues with respect to the selection of instructional methods, and the development and effectiveness measurement of eLearning:

Air carriers and training centers do not universally employ systematic design methodologies to select, develop, or evaluate the effectiveness of instructional methods, including eLearning. This reduces the likelihood of selecting an appropriate instructional method, developing effective training, and may allow suboptimal training practices to continue. The availability of FAA guidance material facilitating the application of systematic design methodologies, and of an inspector workforce trained to recognize and evaluate such methodologies, would assist air carriers and training centers in the use of such methodologies to select, develop, and measure the effectiveness of instructional methods, including eLearning. It would also assist FAA inspectors in evaluating and, where appropriate, accepting or approving certificate holders' use of such methodologies to develop training content, including eLearning.

Existing FAA guidance is also not necessarily complete or reflective of modern eLearning methods or systematic design methodologies. In the absence of clear guidance, FAA inspectors may evaluate such methods or methodologies without the benefit of consistent standards, potentially leading to the acceptance of suboptimal eLearning approaches.

For example, over the past year, partly due to the novel coronavirus (COVID–19) pandemic, many air carriers have rapidly increased their use of eLearning. In some cases, classroom material such as PowerPoints presented by instructors have simply been electronically disseminated for self-study, without the benefit of an instructor to teach the material or any speaker notes to provide context for the material. This is reminiscent of early eLearning implementations. "In order to avoid the rapid but poor-quality development that occurred with early e-learning, aviation training professionals need to recognize that sound instructional design is more important than the technology being used." (Kearns, 2013, p. 75). In this context, "training professionals" should include the FAA inspectors who have oversight authority over whether to accept or approve such training.

Finally, many air carriers choose to contract with part 142 training centers to train their personnel. While this delegation of responsibilities can provide an efficient and effective way for air carriers to fulfill their training responsibilities prescribed by regulation, such training centers and air carriers are not similarly situated with respect to access to post-training data and information. Training centers administering training on behalf of air carriers often do not have access to post-training operational safety and performance data from those air carriers, that would allow them to facilitate evaluation and determination of the effectiveness of the training they administer. Safety management systems (SMS) and other safety programs at part 121 and 135 air carriers[3] provide a wealth of operational data relative to both safety and pilot performance, but part 142 training centers currently do not have a mechanism by which they are

[3] SMS is currently more deeply integrated at part 121 air carriers. As SMS expands more broadly to part 135 and other operators, greater capture and use of relevant sources of data will occur.

assured access to such data from the operators for whom they conduct training to support continuous improvement of that training.

III. Recommendations

The ACT ARC recommends the FAA consider the following actions:

1. FAA guidance should encourage and facilitate the use of systematic design methodologies, to include, but not limited to, those found in instructional systems design (ISD) models, by part 121 and 135 air carriers and part 142 training centers to:
    a. Analyze training needs and requirements to select instructional methods, and specifically the selection of eLearning as a method.
    b. Design and develop eLearning.
    c. Implement the eLearning.
    d. Evaluate the effectiveness of eLearning.
        i. Establish certificate holder data collection methodologies as part of eLearning effectiveness evaluations to ensure relevant data is available to training providers at 121 and 135 air carriers and 142 training centers. Such data could include those listed in Appendix 1.
2. FAA guidance on delivery of knowledge training should be updated to reflect the wide availability of new training methods and tools, including eLearning, including—
    a. The promulgation of a glossary setting out a clear definition of terms used in eLearning.[4]
    b. FAA Order 8900.1[5] should be updated (or other guidance developed) to provide inspectors guidance on what to look for when evaluating a certificate holder's systematic design methodologies.
    c. An Advisory Circular containing guidance to certificate holders on possible approaches and best practices in developing and using a systematic design methodology for training content.
3. FAA training of inspectors should educate them on principles of systematic training design and assist them in evaluating certificate holders' use of systematic design methodologies in the selection of instructional methods and the development and effectiveness assessment of eLearning.
4. FAA guidance to encourage the sharing of operational data from air carrier programs relative to both safety and pilot performance between air carriers and operators and the training centers with which they work.

IV. Rationale and Discussion

The goal of training is to prepare learners to solve real problems and complete real tasks. Knowledge training provides a vital building block to achieve this goal. Knowledge training, combined with other activities that place the knowledge gained via training in context, provides a foundation for the practical application of that knowledge. Effective delivery of knowledge is key to the learner's deeper understanding of the material, reinforced through application, and the

[4] The ACT ARC has already recommended such a glossary in ACT ARC Recommendation 20-11, which includes a suggested list of terms and definitions.

[5] In particular, updates of 8900.1, Volume 3, Chapter 54, Section 6, and Volume 3, Chapter 19, Section 5 would be beneficial.

ability to adapt and apply that knowledge to tasks and events that differ from those seen in training.

As stated above, air carriers and training centers are increasingly turning to eLearning approaches to provide training to their personnel and fulfill their training requirements. Evidence from research and practice has shown that these modern training practices can be equivalent to traditional training practices in terms of acceptance by trainees and in their effectiveness (see Sonnenfeld et al., 2021). Also, eLearning methods may, in fact, provide notable advantages over conventional training practices in some respects, especially regarding their scalability, cost per trainee, and optimality. That said, there are a number of concerns specific to eLearning that make a systematic design methodology and effectiveness evaluation of this type of training even more important than for traditional, instructor led, in-person training:

- Unlike in instructor-led, in-person learning, the instructor and student(s) are not physically present together. Consequently, neither direct observation of the learner by the instructor, nor the social or modeling aspect of learning, whereby students actively learn from the presence or absence of action, inaction, silence, and discussion, can occur.
- Non-instructor led eLearning does not lend itself as well to having small groups of students perform tasks in face-to-face classroom instruction. These activities can lead to exchanges between the students that lead to increased levels of learning beyond that of rote memorization or understanding, fostering higher levels of learning such as application or correlation. eLearning inhibits the students and instructors from observing the social context in which training material is both communicated and received by all involved. Rather, each student becomes an individual learner and the immediate flow of information moves only in a single direction without real-time feedback.
- When implementing eLearning, operators and students also relinquish control of the standardized and controlled learning environment in which each student would otherwise learn material via traditional classroom or training center attendance. Unless the opportunity for such variance is otherwise mitigated, the requirement to secure an environment both free of distraction and conducive to learning is left to each individual student. The ability of a student to ensure an adequate learning environment, as well as the assessment by each student in determining what constitutes an appropriate environment, may vary significantly due to student location as well as social, cultural, and economic factors.
- eLearning does not have the same opportunity for ad hoc learning that occurs when students travel to the training center, are away from home and any associated family responsibilities, and, as a result, often gather outside of classroom sessions in study groups, mealtime discussions, or other interactions. Even if virtual study groups or other methods of student-to-student contact are made available, because the students are trying to learn from home or an uncontrolled environment, they may be pressured to meet their household family responsibilities and are more likely to try to complete only the required training as quickly as possible, and not take the added time to make use of such opportunities.

- Internet connection speeds available vary by location. The necessary online training accommodation of slower internet speeds to ensure the training runs with minimal interruption can have a significant impact on what material and media (*e.g.*, video) can be used in online training.
- Those developing online courses must have training on the challenges of non-instructor led eLearning so that their courses have maximum learning and retention effectiveness. For example, reducing the hours of learning in a given day, as compared to in person instruction, may be necessary when learning on an electronic device to avoid fatigue and burnout, which could result in reduced learning and retention effectiveness (European Union Aviation Safety Agency 2020). This is an example of why it is inappropriate to simply capture recordings of lessons given in a classroom by an instructor and place them online without modification to the material and schedule.
- FAA inspectors require specialized training to evaluate air carrier and training center non-instructor led online training courses, to ensure they have the ability to recognize ineffective training and require corrective action before it is implemented.

Use of Systematic Design Methodologies

Modern training is best when systematic design methodologies are employed for training needs analysis, selection of instructional methods, the design of the training, and the evaluation of its effectiveness. In the absence of a systematic design methodology, the selection of instructional methods and development and evaluation of the effectiveness of training will yield inconsistent results. While use of a systematic design methodology will not guarantee success, research indicates it will increase the likelihood of success (Salas & Stagl, 2014). However, a systematic design methodology will only be helpful in solving the challenges of eLearning if it embodies a thorough, comprehensive process and is inclusive of all stakeholders. An effective systematic design methodology process will incorporate all stakeholder viewpoints through compromise to reach the end product. Systematic design methodologies and their subsidiary processes should be thoroughly documented, to ensure their consistent and complete application. This documentation should be made available to FAA inspectors evaluating eLearning developed using the documented methodologies. In addition to facilitating improved oversight, thorough documentation of systematic design methodologies ensures that institutional knowledge is not lost because of turnover of management and other personnel involved in the development, implementation, and continuous improvement of eLearning content.

When training program developers properly use systematic design methodologies to develop training, the need to evaluate training content on a slide-by-slide or word-by-word basis is diminished to the point where inspectors may shift their emphasis from evaluation of content to evaluation of the design methodology itself, permitting a more consistent application of guidance and greater efficiencies to both the FAA and the training program developers.[6]

[6] The extent to which inspectors are able to place such emphasis on design methodology is dependent on the complexity and characteristics of the design methodology itself. For example, many certificate holders use commercial-off-the-shelf (COTS) training content not specifically tailored to their operations or training needs for

Training Needs Analysis

Systematic training needs analysis (TNA) is a fundamental step in systematic development of training interventions, systems, and materials. TNA describes the practice of defining training requirements (Salas *et al.*, 2012; see also Iqbal & Khan, 2011). Its importance, basic methods, and guidance are extensively covered in the literature (Salas *et al.*, 2012), and TNA is used for the development of training in aviation (Baily & Shaw, 1996). Training requirements as derived from a systematic TNA typically describe: (a) *who* needs to be trained, (b) *what* needs to be trained, (c) the *context* in which training is to occur, and often (d) *how* training should occur (Ferreira *et al.*, 2015; Salas *et al.*, 2012). Despite variations in terminology and methodology in both research and practice (Ferreira & Abbad, 2013; Triner *et al.*, 1996; Watkins & Kavale, 2014), the overall purpose of TNA appears consistent—it is used to identify the competencies required for effective task performance and to prescribe performance interventions appropriate for supporting acquisition and subsequent performance of those competencies (Ferreria *et al.*, 2015; McCelland, 2002; *cf.* CAE, 2017). Although empirical findings are limited (Ferreria *et al.*, 2015; Garavan *et al.*, 2020), absence of TNA in training development may lead to ineffective training and wasted financial resources (Denby, 2010; Iqbal & Khan, 2011), while systematic use of TNA ensures alignment between identified training needs and selected training methods (Salas *et al.*, 2012), and may improve post-training outcomes including training transfer (Ludwikowska, 2018; see also Kodwarni & Prashar, 2019; Lacerenza *et al.*, 2017; *cf.* Garavan *et al.*, 2020).

Training Design

After systematic training needs analysis, the next step in the systematic development, administration, and evaluation of training is systematic training design. This includes the selection of training topics, selection of training modalities, selection and development of training contents, and the formative evaluation of the created training materials and processes.

When an air carrier or training center considers utilizing eLearning, and when an FAA inspector evaluates it, many considerations should be well thought out, such as whether the characteristics of the material to be taught make interaction with others necessary or advantageous; or whether the training in question requires certain equipment available only at the training center. When making such decisions, certificate holders and inspectors should be mindful of research indicating some classroom instruction, combined with eLearning, results in better learning than eLearning alone; “For job-relevant knowledge and skills, blended learning that includes both eLearning and classroom instruction has been found to be more effective than either approach alone” (Kearns, 2018; citing Sitzmann *et al.*, 2006; Zhao *et al.*, 2005).

portions of their training. While providers of such COTS materials may represent the use of a systematic design methodology in its development, the degree to which it does or does not take into account the specific training needs of the certificate holder will be a factor in determining an inspector’s focus in evaluating it. The more general the audience a product is developed for, the more scrutiny of the content itself may be warranted.

As with other parts of the training development process, the application of systematic methodologies for training design increases the likelihood that effective and efficient training is being developed (see Hampton *et al.*, 2018, Salas & Stagl, 2014).

The FAA has encouraged the use of systematic methods for training design in a number of guidance documents, including Advisory Circular (AC) 120–54A (Advanced Qualification Program), AC 120–92B (Safety Management Systems), AC 120–72A (Maintenance Human Factors Training), and AC 120–35C (Line Operational Simulation). One of these methodologies, Instructional Systems Design (ISD), is described in detail in AC 120–54A, but other systematic design methodologies (*e.g.*, ADDIE), are available and may be suitable for the development of training using eLearning.

Similarly, both the International Civil Aviation Organization (ICAO) and the International Air Transport Association (IATA) have encouraged the use of systematic training design methodologies. For example, ICAO has promulgated *Guidelines for the Development of Online Courses* (ICAO, 2019). In it, ICAO references the two fundamental questions for any training design, but specifically applicable to training using eLearning: (1) Is eLearning the appropriate and suitable training delivery medium? and (2) What specific considerations should the training design consider with respect to the curriculum development under eLearning? In response to the first question, ICAO stated that "Not all performance problems can be addressed through eLearning. An analysis of whether eLearning is appropriate to effectively achieve the learning outcomes (knowledge, skills, attitudes) for the target audience is an important part of" systematic training design (ICAO, 2019, p. 8). In response to the second question, ICAO provided detailed guidance on an approach that systematically develops the curriculum. Guidelines suggested include items such as "design elements need to adequately deliver the content through the selected method" (page 25), "eLearning content must be accurately prepared and presented in order to be effective" (page 30), and "Prior to release, an online course should undergo quality assurance testing" (page 33). Note that, while ICAO evokes ADDIE as a design methodology in other documents (*e.g.*, the TRAINAIR PLUS Training Development Guide (TDG) Competency-based Training Methodology), no specific framework is prescribed or recommended in the *Guidelines for the Development of Online Courses*.

In sum, training design should, like Training Needs Analysis before it and Training Effectiveness Evaluation after, use and apply a systematic design methodology. A number of frameworks are available to training designers, and training organizations should be able to select an appropriate systematic design methodology based on their specific needs and considerations.

Training Effectiveness Evaluation

Knowledge testing and knowledge training[7] have both been required for decades, and the need to evaluate their effectiveness is not new. Effectiveness evaluation ensures evaluation criteria are being met, encourages excellence, and supports continuous improvement and corrective actions to address inadequacies. However, such evaluation has not always been systematic and data-driven. Without an analysis of relevant data, it is not possible to determine, for

[7] The scope of eLearning includes far more than knowledge training, but the EKT WG's tasking and this discussion are limited to knowledge training and testing.

example, whether a student passed an evaluation because of the training or in spite of it. Examination of pass or fail rates on the basis of a traditional knowledge test in multiple choice format is not always the best way of determining effectiveness. In the same way as today's eLearning can do more than merely convey static information, it can also support creative ways to assess knowledge, skills, and competencies. For examples of ways to evaluate effectiveness of knowledge training using eLearning, see Appendix 1.

Training, and the systematic validation of that training, is one of the core foundations of aviation safety (IATA, 2018); yet, in both theory and practice, training evaluation is a complex subject. Defined as a systematic procedure for determining the effectiveness of training using standardized criteria (FAA, 2014), training evaluation involves systematic data collection to validate whether learning objectives were achieved as evidenced by enhanced operational performance (Salas *et al.*, 2012). Training evaluation, similar to other modern aviation training practices, is shaped by a diverse array of theories, methods, and technologies. Review of the literature (Sonnenfeld *et al.*, 2020) suggests that there are four primary forms of training evaluation conducted in commercial aviation:

1. Curriculum-based evaluation, as outlined in 14 CFR parts 121 and 135 (FAA, 2020b);
2. Development-based evaluation, as exemplified in the Advanced Qualification Program (AQP; FAA, 2015b; 2017a; 2017b);
3. Systems-based evaluation, as illustrated in 14 CFR part 60 on flight simulation training device qualification and use (*e.g.*, FAA, 2020a; see also Goodwin *et al.*, 2018); and
4. Effectiveness-based evaluation.

The first three of these evaluations, *i.e.*, curriculum-, development-, and systems-based evaluations, only indirectly address training effectiveness—for example, through audits of individual performance/proficiency data (Air Transport Association, 1998) and instructor/evaluator assessments. While these other forms of evaluation each provide a unique perspective on training quality and shape modern practices in air carrier training, Recommendation 1.d specifically addresses the use of systematic frameworks for training effectiveness evaluations, with a specific focus on the fact that instructor-less eLearning training may have profound implications for the application of these frameworks for training effectiveness evaluation.

Training effectiveness evaluation (TEE) is defined as a determination of the quality of training provided, the comprehension of a subject by a learner, or the ability of a learner to acquire a trained method, skill, or technique (FAA, 2010). TEE is conducted for multiple purposes—to quantify that a training intervention is achieving intended objectives and goals, to quantify resultant differences in performance gaps, to quantify the cost/benefit ratio of training, to verify and validate instruction, and to justify changes in training interventions (Goodwin *et al.*, 2018). Further, there is an expectation in effectiveness-based evaluation approaches that there is a systematic evaluation of the effectiveness and efficiency of instruction—with, at a minimum, a summative evaluation of reactions, learning, performance, transfer, and instructional techniques used (FAA, 2014).

Regardless of the topics, tasks, or KSAs (knowledge, skills, and attitudes) being trained, there are underlying principles and practices that guide the training effectiveness evaluation process. Evaluation principles and practices are often implemented within frameworks or models (*e.g.*, Stufflebeam, 2001). Many of the earlier models were based on Kirkpatrick's taxonomy

(Kirkpatrick, 1996; see also Kirkpatrick, 1956; Thalheimer, 2018), which probably has been the most influential effectiveness evaluation framework in the training domain (*e.g.*, Passmore & Velez, 2012; Sitzmann & Weinhardt, 2019), and is used across industries (*e.g.*, Birdi, 2020; Botek, 2018; Moreau, 2017), including in aviation (Salas *et al.*, 2001). Most other TEE frameworks allude to Kirkpatrick's taxonomy (Botek, 2018; Giangreco, 2008), further developing upon its strengths or addressing its shortcomings (*e.g.*, Holton, 1996a; 1996b; Sitzmann & Weinhardt, 2019).

Common to all these frameworks are a number of characteristics that have particular applicability to the evaluation of training using eLearning methods:

- Collection of multiple indicators of training effectiveness. All systematic frameworks recommend that organizations collect data at multiple levels of aggregation and analysis. This includes obtaining feedback from the students, the formal testing of knowledge and skill (including pass-fail rates, remediation rates, amount of remediation required), but also on-the-job performance measurement (*e.g.*, Line Operating Experience (LOE), Line Check, Line Operations Safety Audit (LOSA)), and the analysis of business results (*e.g.*, safety metrics – Aviation Safety Action Program (ASAP), Flight Operational Quality Assurance (FOQA), LOSA, Safety Management Systems (SMS)).
- Collection of indicators of training effectiveness at multiple points. All systematic frameworks recommend that data are collected not just once, but at multiple points in time within the training and on-the-job performance periods.
- Collection of data from multiple sources. All frameworks suggest data collection from multiple sources (*e.g.*, trainees, instructors, training managers, and end users).

Validation of eLearning consisting in large part of open book tests completed by flightcrew members at the end of their self-study has long been subject to criticism. One example of a way to help overcome this suspect validation and improve the effectiveness of eLearning training evaluation would be to conduct proctored, non-jeopardy closed book exams when flightcrew members first participate in on-site training following their eLearning, to determine their extent of learning and retention and the effectiveness of the training material.[8, 9]

Based on the available literature, recent research evaluated over a dozen TEE frameworks using criteria that included the utility, feasibility, propriety, accuracy, and suitability (Sonnenfeld *et al.*, 2020). Overall, this research found that each TEE framework has unique advantages and limitations, and that the selection of a framework, in itself, should be a systematic process based on comparison of the unique advantages of each against training needs and organizational requirements.

Given that no single framework is clearly preferred over others, an opportunity exists for the FAA to allow air carriers and training centers to adapt frameworks to suit their needs, given that systematic evaluation still occurs consistent with contemporary training science—involving

[8] Such exams would be analogous to "First Look" assessments used under AQP.

[9] This concern is largely limited to validation of eLearning related to aircraft-specific certification or qualification of flightcrew members. For course material that is delivered via eLearning that is not aircraft-specific or tailored to aircraft-specific flightcrew member certification, a non-proctored online examination may be sufficient.

assessment training input factors, system affordances, and multilevel measurement of learning, performance, and transfer.

In any case, however, an air carrier or training provider seeking to assess the effectiveness of knowledge training should go beyond an associated demonstration or evaluation that is part of a training program itself. Operational performance is a necessary part of evaluating the effectiveness of eLearning. The safe and proficient demonstration of skills or knowledge in an operational environment away from the training environment is the intended outcome of eLearning. As a result, the determination of effectiveness requires that performance data extend beyond what was assessed within the confines of the training program.

Data Collection as a Function of Evaluating eLearning Effectiveness

Air carriers and training providers may be best equipped to evaluate eLearning effectiveness by utilizing a diversified portfolio of validation tools, including data collection and dissemination to provide continuous feedback on the effectiveness of their training. Training programs should incorporate data collection methodologies diversified enough to capture a measure of effectiveness at various time intervals after training takes place, account for various learning styles, and provide quality feedback on the effectiveness of eLearning training.

An example using student surveys as a data collection method:

- Provide surveys during the early, middle, and late stages of training that focus on both the perception of the training program and the learning experience of the student.
- Survey the student after completion of the training program and again after a predetermined time of operational experience, to provide further perspective with respect to the training received, as well as the ability of the E&DL training to prepare the pilot for other aspects of training and subsequent flight operations.

Qualitative surveys are, however, by themselves an insufficient means of evaluating the effectiveness of a modern training program (*e.g.*, Holton, 1996A; Sitzmann & Weinhardt, 2019). A list of examples of data collection methods is provided in Appendix 1. This list is intended to provide examples of data collection methods and is not meant to be comprehensive or all encompassing.

Updated Guidance

The world of technology and innovation continues to change and evolve. Guidance with respect to training methods should also be updated to reflect the current ways in which the aviation industry delivers training content. To ensure an industry common understanding, the ACT ARC has recommended the establishment of a glossary of accepted terms and definitions and proposed that the FAA continue to maintain and use it to create such guidance. This glossary should be a living document, continually updated and adjusted as technology and methods change, and new terms and usages are developed. The terms used in the glossary should be harmonized across other guidance documents to the best extent possible.

FAA Order 8900.1 does not currently provide inspectors a sufficient knowledge base or guidance to evaluate and differentiate between what is an acceptable method of instructional design and what would be considered inadequate or unacceptable. Clear terms and definitions along with basic guidance established in Order 8900.1 would provide examples and a better understanding of what basic elements make up a good systematic design methodology. Access to such information would help to improve the quality and consistency of FAA approval and review processes and provide air carriers and training centers consistent information about FAA review processes and principles.

We also recommend the development of an advisory circular to inform operators and inspectors on possible approaches and best practices in developing and using a systematic design methodology for training content.

Training of Inspectors

Training and education of the FAA inspectors tasked with assessing and approving certificate holders' systematic design methodologies is of critical importance. In the absence of clear guidance on how to evaluate methodologies and training, inspectors will be without the benefit of standards, and may, consequently, focus on details such as spelling errors or graphics quality, while missing key attributes such as alignment with learning objectives, process-driven media selection, or learning effectiveness.

Training for inspectors should provide them with the knowledge and tools necessary to transition from the paradigm of approval of individual training programs and content to approval of both a certificate holder's systematic design methodology and the proposed use of eLearning. (See Use of Systematic Design Methodologies, above.) This requires more than training to execute evaluation processes; instead, inspectors must receive education sufficient to provide a meaningful, broad-based understanding of the necessary elements of a systematic design methodology, and of how to ensure those elements are in place and functioning as intended.

Inspector instruction on systematic methodologies should be incorporated into their initial and recurrent training curricula, and should encompass all aspects of such methodologies, including assessment of training needs, development and design of content and systems, implementation and administration of training, and evaluation of effectiveness. The training should also include, reference, and/or review some of the most commonly used models (*i.e.*, ADDIE), instruction, and materials should examine issues experienced by field personnel when applying rules, standards, and guidance to evaluate the effectiveness of certificate holder training, including "pitfalls" that may result in incomplete or inaccurate evaluation.

Sharing of Data Between Air Carriers and Training Centers

To ensure the continued effectiveness of any eLearning training program, a closed feedback loop to allow for continuous improvement is critical. The ACT ARC recommends the FAA encourage the sharing of deidentified aggregate operational data between certificate holders and their contract training providers, as well as training data from training providers to certificate holders. Where a Safety Management System (SMS) is either required or voluntarily utilized,

there is a tool already in place to address the process and guidelines for this two-way feedback. Based on current regulations, not every certificate holder is required to have an SMS. In such circumstances, the FAA should establish guidelines relative to the practice of sharing such reciprocal information. This could be a recommended strategy for risk mitigation for the certificate holder as well as the contract training provider.

Conclusion

The ACT ARC recommends the FAA take broad-based steps, including, but not limited to development of new guidance, to both encourage and facilitate the use of systematic design methodologies by certificate holders and to educate and equip the aviation safety inspector workforce to meaningfully evaluate the effectiveness of such methodologies and the training content developed using them. Both the systematic design methodologies used by certificate holders and the effectiveness evaluation processes used by inspectors should rely on a variety of training and operational data sources to ensure the ultimate objectives of enhanced safety of operations is achieved.

V. Background Information

Recommendation 21-11 addresses Item 3 in the EKT WG Scope of Work and ACT ARC Initiative #44 (see below):

EKT WG Scope of Work:

3. Develop recommendations for FAA guidance to certificate holders and aviation safety inspectors on systematic training development methodologies, to include deployment and effectiveness assessment, for electronically delivered, non-instructor led knowledge training.

ACT ARC Initiatives:

- Initiative #44: Recommend guidance for the development and approval of knowledge training with a focus on maximizing training effectiveness.

VI. References

Air Transport Association. (1998, May 12). Advanced Qualification Program Data Management Guide. *Data Management Focus Group.* Retrieved from https://www.faa.gov/training_testing/training/aqp/library/media/dmg.pdf

Bailey, L. L., & Shaw, R. V. (1996). Flight Inspection Crew Resource Management Training Needs Analysis (No. DOT/FAA/AM-96/24). *Federal Aviation Administration Office of Aviation Medicine.* Retrieved from https://apps.dtic.mil/sti/pdfs/ADA318104.pdf

Birdi, K. (2020). Insights on impact from the development, delivery, and evaluation of the clear ideas innovation training model. *European Journal of Work and Organizational Psychology,* 1-15. https://doi.org/10.1080/1359432X.2020.1770854

Botek, M. (2018). Comparison of education evaluation models: Acces la success. *Calitatea,* 19(165), 63-67. Retrieved from https://search.proquest.com/scholarly-journals/comparison-education-evaluation-models/docview/2089263278/se-2?

CAE. (2017, June 7). Collective Training Needs Analysis Review. *Defence Research and Development Canada.* Retrieved from https://cradpdf.drdc-rddc.gc.ca/PDFS/unc334/p805422_A1b.pdf

Denby, S. (2010). The importance of training needs analysis. *Industrial and Commercial Training*, 42(3), 147-150. https://doi.org/10.1108/00197851011038132

European Union Aviation Safety Agency. (2020, August 18). Guidance for allowing virtual classroom instruction and distance learning; *Guidelines in relation to the COVID–19 pandemic*. Retrieved from https://www.easa.europa.eu/sites/default/files/dfu/easa_guidelines_-_virtual_classroom_instruction_and_distance_learning_is.pdf

Federal Aviation Administration. (2010, February 26). *Web Service Taxonomies.* FAA-STD-066. U.S. Department of Transportation. Retrieved from http://www.tc.faa.gov/its/worldpac/standards/faa-std-066.pdf

Federal Aviation Administration. (2014, May 15). *Air Carrier Training: Check Airman and Instructors*. Retrieved from https://www.faa.gov/pilots/training/air_carrier/ac_chk_airman_instructor/

Ferreira, R. R., & Abbad, G. (2013). Training needs assessment: where we are and where we should go. *Brazilian Administration Review*, 10(1), 77-99. https://doi.org/10.1590/S1807-76922013000100006

Garavan, T. N., O'Brien, F., Duggan, J., Gubbins, C., Lai, Y., Carbery, R., Heneghan, S., Lannon, R., Sheehan, & Grant, K. (2020). The current state of research on training effectiveness. *In Learning and Development Effectiveness in Organisations* (pp. 99-152). Palgrave Macmillan, Cham. https://doi.org/10.1007/978-3-030-48900-7_5

Giangreco, A., Sebastiano, A., & Carugati, A. (2008). Are we doing the right thing? Food for thought on training evaluation. *Paper presented at workers' rights protection in a new world of work. The case for a comparative and interdisciplinary approach to Labour Relations 6th International Conference in Memory of Professor Marco Biagi, Modena, Italien*. http://www.fmb.unimore.it/on-line/en/Home/Conferences/articolo4084.html

Goodwin, M. S., Reinerman-Jones, L., Harris, S., & Arrieta, A. (2018). A taxonomy for the evaluation of training simulations and environments. *MODSIM World 2018.* Retrieved from http://www.modsimworld.com/papers/2018/MODSIM_2018_Paper_not_required_for_a_43.pdf

Hampton, S., Neal, J. G., Ramirez, L. A., Talkington, D. R., & JMA Solutions, Inc. (2018). Task 10: Research an Alternative Instructional Design Model. *Center of Excellence for Technical Training and Human Performance (TTHP): Solutions for Operational Aviation Research.* Retrieved from https://commons.erau.edu/publication/1027

Holton III, E. F. (1996a). The flawed four-level evaluation model. *Human Resource Development Quarterly,* 7(1), 5-21. https://doi.org/10.1002/hrdq.3920070103

Holton III, E. F. (1996b). Final word: Response to reaction to Holton article. *Human Resource Development Quarterly,* 7, 27-30. https://doi.org/10.1002/hrdq.3920070105

International Civil Aviation Organization (2019). *Guidelines for the development of online courses.* Retrieved from https://www.icao.int/NACC/Documents/Meetings/2019/CATCWG5/Guidelines%20for%20the%20Development%20of%20Online%20Courses_V1_Feb2019.pdf

Iqbal, M. Z., & Khan, R. A. (2011). The growing concept and uses of training needs assessment. *Journal of European Industrial Training.* 35(5). 439-466. https://doi.org/10.1108/03090591111138017

Kearns, Suzanne K. (2013). Snap-Courses: An instructional design strategy for aviation mobile learning. *Collegiate Aviation Review International*, 31. Retrieved from https://ojs.library.okstate.edu/osu/index.php/CARI/article/download/7474/6876

Kirkpatrick, D. L. (1956). How to start an objective evaluation of your training program. *T+D*, 5, 16. Retrieved from https://www.kirkpatrickpartners.com/Portals/0/Resources/How%20to%20Start%20an%20Objective%20Evaluation%20of%20Your%20Training%20Program.pdf?ver=2018-01-29-083410-360

Kirkpatrick, D. L. (1996). Invited reaction: Reaction to Holton article. *Human Resource Development Quarterly*, 7(1). https://doi.org/10.1002/hrdq.3920070104

Kodwani, A. D., & Prashar, S. (2019). Exploring the influence of pre-training factors on training effectiveness-moderating role of trainees' reaction: A study in the public sector in India. *Human Resource Development International*, 22(3), 283-304. https://doi.org/10.1080/13678868.2019.1596012

Lacerenza, C. N., Reyes, D. L., Marlow, S. L., Joseph, D. L., & Salas, E. (2017). Leadership training design, delivery, and implementation: A meta-analysis. *Journal of Applied Psychology*, 102(12), 1686–1718. https://psycnet.apa.org/doi/10.1037/apl0000241

Ludwikowska, K. (2018). The effectiveness of training needs analysis and its relation to employee efficiency. *Scientific Papers of the Poznań University of Technology: Organization and Management*, 77, 179-193. http://dx.doi.org/10.21008/j.0239-9415.2018.077.11

McClelland, S. D. (2002). *A Training Needs Assessment for the United Way of Dunn County Wisconsin [Master's Thesis]*. University of Wisconsin-Stout. Retrieved from https://minds.wisconsin.edu/bitstream/handle/1793/40545/2002mcclellands.pdf?sequence=1

Moreau, K. A. (2017). Has the new Kirkpatrick generation built a better hammer for our evaluation toolbox? *Medical Teach*, 39(9), 999-1001. https://doi.org/10.1080/0142159x.2017.1337874

Passmore, J., & Velez, M. (2012). SOAP-M: a training evaluation model for HR. *Industrial and Commercial Training,* 44(6), 315-325. https://doi.org/10.1108/00197851211254743

Salas, E., Burke, C. S., Bowers, C. A., & Wilson, K. A. (2001). Team training in the skies: does crew resource management (CRM) training work? *Human Factors,* 43(4), 641-674. https://doi.org/10.1518/001872001775870386

Salas, E., Stagl, K. C. (2009). Design training systematically and follow the science of training. In Locke, E. (Ed.), *Handbook of Principles of Organizational Behavior: Indispensable Knowledge for Evidence-Based Management* (2nd ed., pp. 59–84). Chichester, UK: Wiley. https://doi.org/10.1002/9781119206422.ch4

Sitzmann, T., Kraiger, K., Stewart, D., & Wisher, R. (2006). The comparative effectiveness of web-based and classroom instruction: A meta-analysis. *Personnel Psychology*, 59, 623–664.

Sitzmann, T., & Weinhardt, J. M. (2019). Approaching evaluation from a multilevel perspective: A comprehensive analysis of the indicators of training effectiveness. *Human Resource Management Review*, 29(2), 253-269.

Sonnenfeld, N. A., Boesser, C. T., Nguyen, B., & Jentsch, F. (2021a) Modern training practices: Effectiveness of electronic and distance learning (E&DL) for procedural training in aviation [Manuscript in preparation]. Institute for Simulation and Training, University of Central Florida.

Stufflebeam, D. (2001). Evaluation models. *The American Evaluation Association's Program Evaluator Competencies*, 2001(89), 7-98. https://doi.org/10.1002/ev.3

Thalheimer, W. (2018, January 30). Donald Kirkpatrick was not the originator of the four-level model of learning evaluation. *Work-Learning Research.* Retrieved from https://www.worklearning.com/2018/01/30/donald-kirkpatrick-was-not-the-originator-of-the-four-level-model-of-learning-evaluation/

Triner, D., Greenberry, A., & Watkins, R. (1996). Training needs assessment: A contradiction in terms? *Educational Technology*, 36(6), 51-55. https://www.jstor.org/stable/44428374

Watkins, R., & Kavale, J. (2014). Needs: Defining what you are assessing. In J. W. Altschuld & R. Watkins (Eds.), *Needs Assessment: Trends and a View Toward the Future*, 144, 19–31. https://doi.org/10.1002/ev.20100

Zhao, Y., Lei, J., Yan, B., Lai, C., & Tan, H. S. (2005). What makes the difference? A practical analysis of research on the effectiveness of distance education. *Teachers College Record*, 107, 1836–1884.

VII. Attachment(s)

Appendix 1: Data Collection Examples for Evaluating eLearning
Appendix 2: Effectiveness Evaluation Framework

Appendix 1
Data Collection Examples for Evaluating eLearning

The following is a list of data types and associated collection methods during and post training. It is exemplar in nature, and is not intended to be prescriptive or exhaustive.

Data Type	Description	Potential Method(s)	Qualitative/ Quantitative	Level of Analysis[10]
Student Feedback	Student's opinions of training	Surveys, Focus Groups, Classroom Interactions, Online Forum Interactions	Qualitative	L1 - Reaction
Instructor Feedback	Instructor's opinions of training	Surveys, Focus Groups, Classroom Interactions, Online Forum Interactions	Qualitative	N/A
Evaluator Feedback	Evaluator's opinions of training	Surveys, Focus Groups	Qualitative	N/A
SME Feedback	SME's opinions of training	Surveys, Focus Groups	Qualitative	N/A
Peer Feedback	Crewmember opinions of pilots' performance post-training	Surveys and other feedback	Qualitative	N/A
L&D Department Feedback	Designer's opinions of training	Process, Focus Groups	Qualitative	N/A
Third party feedback	Feedback from part 142 customer/certificate holder	Surveys and other feedback	Qualitative	N/A
Training Metrics	Frequency and time spent with training modules; Module Elapsed Time	Learning Management System	Qualitative/ Quantitative	N/A

[10] Level of Analysis refers to the four levels of the Kirkpatrick Taxonomy for evaluating training effectiveness, and its derivatives: 1) learner reaction, 2) learning, 3) learner behavior, and 4) results. (See Appendix 2, Training Effectiveness Evaluation Frameworks.) The Kirkpatrick Taxonomy is an example of one method of such analysis, but not the only method. Another example would be Sitzmann & Weinhardt's levels, which, along with others, should be considered for use.

Data Type	Description	Potential Method(s)	Qualitative/ Quantitative	Level of Analysis[10]
Knowledge	Assessment of knowledge	Analysis of Computer, Oral, or Written examination scores and Pass/Fail rates.	Quantitative	L2 - Learning
Procedural Skill/ Procedures Validation	Assessment of Normal, Non-normal, Emergency, and Supplemental procedures using SBT, Procedures Trainers, and FSTDs	Analysis of procedures testing task scores and evaluator comments.	Both	L2 - Learning
Maneuver Skill/ Maneuvers Validation	Assessment of manual and automated flying skills using FSTDs	Analysis of maneuver testing task scores and evaluator comments.	Both	L2 - Learning
CRM Skill	Assessment of CRM skills, typically in a real-time scenario using FSTDs	Analysis of CRM testing task scores and evaluator comments.	Both	L2 - Learning
On-line Testing (Line Operational Evaluations (LOE), line checks, etc.)	Validation of pilot's ability to perform in the real-world operational environment. Examples include Line Check, Theater Qualification, SAQ	Analysis of Knowledge, Procedures, Maneuvers, and CRM testing task scores and evaluator comments.	Both	L3 - Behavior
Knowledge Test Content	Quality of test questions, answers, distractors, and graphics.	Analysis of question quality, alignment with learning objectives, alignment with instructional content, frequency of Correct/Incorrect, ordering of Incorrect answer choices, comparison of question topics scoring data.	Quantitative	N/A

Data Type	Description	Potential Method(s)	Qualitative/ Quantitative	Level of Analysis[10]
"Not Recommends"	The number of pilots that were not recommended for a performance assessment by the instructor responsible for the recommendation. Typically, the training module prior to the evaluation.	Analysis of training modules graded Not Recommend, including instructor comments.	Quantitative	N/A
Extra Curriculum Days	The number of extra curriculum days pilots needed to complete a training program. The extra days may be used for training, checking, or both.	Analysis of Extra Curriculum Day (ECD) usage and reasons for ECD assignment.	Quantitative	N/A
OE Extensions	The number of flight hours, segments, or days pilots needed to complete a training program.	Analysis of the number of OE hours, flight segments, and days it takes for pilots to complete the OE phase of a curriculum. This data is typically correlated to Line Check data.	Quantitative	N/A
First Look Data	First Look is specific to AQP programs. As part of recurrent FFS training, pilots are given one or two maneuvers to perform without preparation or briefing.	First Look data is used to determine if the training frequency of specific tasks is appropriate by measuring skill and knowledge retention. It's also used to determine how a crew may have performed on the line prior to training as part of research projects.		

Data Type	Description	Potential Method(s)	Qualitative/ Quantitative	Level of Analysis[10]
Aviation Safety Action Program (ASAP)	ASAP is a robust self-reporting system used as part of a Just Culture. Pilots report items that the company or FAA would not have known about without the ASAP system.	ASAP data is analyzed and reported to inform leadership about a wide variety of issues in line flying operations. Many of the reports are directly related to pilot performance	Qualitative	L4 - Results
Flight Operations Quality Assurance (FOQA)	FOQA uses data recording systems that are integrated with the aircraft's flight data system. Hundreds of parameters are captured and dumped into a FOQA data repository for analysis.	FOQA data is analyzed and reported to inform leadership about a wide variety of issues in line flying operations. The data is directly related to pilot performance.	Quantitative	L4 - Results
Accidents/Incidents	Airlines conduct investigations into accidents and incidents as defined by the NTSB.	While small in number, accident and incidents undergo a thorough investigative process. The findings are captured in detailed reports.	Both	L4 - Results
Operational Performance	Airlines have specific metrics designed to determine safety and risk levels of the operation.	Airline-specific operational performance data is analyzed and reported to leadership to determine the state of the operation.	Both	L4 - Results
Audits	Line Operational Safety Audits (LOSA), IATA Operational Safety Audits (IOSA), and other third-party audit programs provide company-specific performance data at a specific moment-in-time (snapshot).	Audit data is collected, cleansed, analyzed, and reported to leadership.	Both	L4 - Results

Data Type	Description	Potential Method(s)	Qualitative/ Quantitative	Level of Analysis[10]
Benchmarking (CAST, A4A ASIAS)	Airlines benchmark their data against other data sets using a variety of tools and processes. This allows the carrier to compare performance with other airlines.	Airlines compare their operational performance metrics with US carriers by processes managed by A4A and RAA. Comparison with national and international research is done through CAST, ICAO, and other organizations. FOQA data can be compared using ASIAS.	Both	L4 - Results
Business System Analytics	Information Technology systems capture data related to user experience, behaviors, and technology.	Analysis of LMS usage data, Learning records Store (LRS) data, company Intranet data, Google Analytics data.	Quantitative	N/A

Appendix 2
Training Effectiveness Evaluation Frameworks

The following is a list of training effectiveness evaluation (TEE) frameworks derived from a presentation SME Florian Jentsch gave to the EKT WG in December 2020. The following bullets are exemplar in nature, and are not intended to be prescriptive or exhaustive. The EKT WG wishes to acknowledge the contributions of researchers at the University of Central Florida (UCF), without which neither this appendix, nor this recommendation, would have been possible.

- **Kirkpatrick Frameworks**
  - Highlights
    - Emphasized ease of use for practitioners & communication with management; generally directed toward business domain
  - **Kirkpatrick Taxonomy**
    - Structure
      - Level 1 – Reactions
      - Level 2 – Learning
      - Level 3 – Behaviors
      - Level 4 – Results
    - Pros/Cons:
      - Simple; highly feasible to implement; well-recognized in industry.
      - Criticized in the empirical literature as incomplete and qualitative; provides minimal guidance on quantitative measurement
  - **New World Kirkpatrick Model (NWKM)**
    - Improved on Kirkpatrick, expanded four levels
    - Pros/Cons:
      - Improved on Kirkpatrick Taxonomy; provided additional guidance for communicating evaluation plans with management and stakeholders.
      - Shared criticisms with Kirkpatrick taxonomy; and is largely based on internal case studies.
  - **Kirkpatrick + Return on Investment (ROI)**
  - **Kirkpatrick + Society**
- **Factors-based models**
  - **Comprehensive Training Effectiveness Model**
    - Structure
      - Expectation fulfillment
      - Learning
      - Training performance
      - Job performance
      - Results/O.E.
  - **Holton's Human Resource Development (HRD) model**
    - Structure

- Influences (Ability, Environment, Motivation, and Secondary Influences) affect outcomes
  - Learning
  - Individual performance
  - Organizational performance
- Pros/Cons:
  - HRD model was evidence-based and included specific influence factors to measure during evaluations to more accurately determine if training was effective.
  - Appeared to lack further development after 2005. Training design influence factors and variables/measures of learning and performance were not explicitly specified.

- **Integrated Model of Training Evaluation and Effectiveness (IMTEE)**
  - Structure
    - Training content & design
    - Changes in learners
    - Organizational Payoffs
  - Pros/Cons:
    - The IMTEE was evidence-based and integrated previous TEE frameworks. Variables & measures for training factors were specified in the literature on the model.
    - Measures of performance variables were unspecified; also, several variables and measures were still being researched in the training science of the time. No indication of further guidance on IMTEE after 2004.

- **Modern Frameworks**
  - Full-Scope Evaluation Model (FSEM)
    - Pros/Cons:
      - Depicted Formative, Summative, Confirmative, and Meta-Levels of Analysis. While not as specific as other models on specific variables/measures, the inclusion of meta-evaluation was unique and informative.
  - Generalized Intelligent Framework for Tutoring (GIFT)
    - Pros/Cons:
      - GIFT was evidence-based, and specific to modern electronic and distance learning. Variables & measures specified in the supporting literature within the GIFT community. The GIFT model supported the GIFT adaptive intelligent training platform, and has been integrated with modern training tools (simulations, peripheral devices, sensors, etc.). GIFT has extensive community support.
      - GIFT did not appear to be applicable to conventional, non-electronic training. Requires increased front-end development of training, the cost/benefit ratio of use appeared to be lower for non-adaptive training. GIFT may be more difficult to implement without integration or interoperability with the GIFT software platform. Also appeared to require practitioners to have other training science domain expertise to implement successfully. Furthermore, the GIFT model did not include transfer or “results” types of outcomes.

  - Taxonomy for Holistic Evaluation and Training Assessment (THETA)
    - Pros/Cons:
      - THETA provided a framework for integrating simulation training evaluations with evaluations of training systems, based on an assessment of the capability of the training system technologies to meet training requirements. While not explicitly specifying factors which influence training among variables to measure, THETA also included measurement of physiological factors among performance outcomes.
  - Multilevel Training Evaluation Taxonomy (MTET)
    - Pros/Cons:
      - MTET was evidence based, and specified variables & measures of training inputs and outcomes. Many outcomes compatible with data inherent to electronic and distance learning (*e.g.*, training utilization; within-subjects learning and performance).
      - However, some outcomes of the model may require additional organizational support to evaluate (*e.g.,* time and resource investment)

- **Quality Models**
  - Software Quality Requirements and Evaluation (SQuaRE)
    - Pros/Cons:
      - Included to provide a basis for guiding training evaluation of E&DL systems from existing ISO standards.
      - Effectiveness included as a sub-outcome of system quality, also evaluated system outcomes (*e.g.*, usability). Would require substantial organizational support to implement.
- **Other Frameworks**
  - Non-evaluation
  - Success Case Method
  - Context-Input-Process-Product (CIPP)/Context-Input-Process-Output (CIPO)/ Input-Process-Output (IPO)