IEEE Learning Standards Special Track: Recommended Practices for the Evaluation of Adaptive Instructional Systems (AISs) Special Event: Evaluate and critique Descriptors that are needed to Inform AIS Buyers about different AIS Capabilities

Natalie B. Steinhauser
Naval Air Warfare Center Training Systems Division*
Orlando, FL
Natalie.Steinhauser@navy.mil

Adaptive Instructional Systems (AISs), or Adaptive Training Systems (ATS), are instructional systems that adapt to the needs of the learner in real-time or after a training event has been completed (Landsberg, Van Buskirk, Astwood, Mercado, & Aakre, 2011). Research has shown that AIS/ATS can improve training effectiveness (Landsberg, Mercado, Van Buskirk, Lineberry, & Steinhauser, 2012; Graesser, Conley, & Olney, 2012; Park & Lee, 2003; Dzikovska, Steinhauser, Farrow, Moore, & Campbell, 2014). Thus, the demand for these systems has increased.

However, as the demand for AIS/ATS increases, it is important to ensure the systems are properly evaluated before adoption. There are many aspects to consider when evaluating an AIS/ATS. These include evaluations of (a) system capabilities, (b) training effectiveness, and (c) transfer of training to on-job performance.

First, it is important to evaluate the capabilities of the AIS/ATS. Campbell (2014) identified three functional attributes that make up all AIS/ATSs: Observe, Assess, and Respond. That is, all adaptive systems must observe student behavior and performance, assess that performance, and provide a response to the student. The sophistication of each of these system features can vary widely. Therefore, in evaluating an adaptive system, it is important to assess the system capabilities for each attribute considering the questions posed below:

- **Observe** (i.e., tracking student behavior and performance)
    - What type of data does the system capture?
        - Completion time?
        - Correctness on questions?
    - At what granularity does it collect data? Overall performance vs. part-task performance vs. skill-based assessment?
    - How does the system store performance data and how does it generate the data (i.e., what is the form of the data output and what data is included in the output)?
- **Assess**
    - How does the system assess student performance?
        - Time-based performance assessment?
        - Performance-based assessment?
            - Assessment of correct/incorrect responses or more analysis?
        - Any other type of assessment?
            - e.g., Physiological measures?
- **Respond**
    - How does the system respond to student behaviors?
        - How does the system adapt to student inputs and behaviors (Landsberg, Astwood, Van Buskirk, Townsend, Steinhauser, & Mercado, 2012)?
            - Micro adaptions
                - Adapts during training and uses on-task performance to drive adaptations of feedback, scenario difficulty, etc.
            - Macro adaptions

** The views expressed in this paper are solely those of the authors, and do not necessarily reflect the opinions of the Naval Air Warfare Center Training Systems Division, or any other Department of Defense agency, unless stated in official directives.*

        - Adapts based on previous performance (usually on a test or assessment) to guide further adaptations (typically in the form of additional instructional content) in the areas of performance that fell below the desired level of proficiency
    - Aptitude-Treatment Interaction
        - Adapting the instructional techniques or content based on a learner's aptitudes or abilities (e.g. tutoring students by using different levels of instructor involvement, dependent upon the students' overall SAT scores [Shute, Lajoie, & Gluck, 2000])
    - Two-step Approach
        - Combines the Micro and the Aptitude-Treatment Interaction approaches
        - Bases the initial training approach on aptitude, then adapts based on on-task performance during training
- What does the system adapt?
    - Feedback
        - What type of feedback does the system provide?
        - When does the system provide feedback?
    - Difficulty
        - Does the training get easier or harder based on performance?
    - Additional Content or Practice
        - Does the system provide additional content or opportunities to practice if the student is struggling?
        - Does the system let the student end early if they master the learning objectives?

After evaluating system capabilities, it is important to assess training efficacy. A widely used approach for conducting training evaluations is Kirkpatrick's Four Levels of Training Evaluation Model (Kirkpatrick & Kirkpatrick, 2006). The objective of this approach is to evaluate the full business value of training. The Four Levels are described below; the ideal training effectiveness evaluation for an AIS/ATS should seek to evaluate all levels.

- Level 1: Reaction
    - Evaluation of student subjective reactions to the training
    - How much they like or dislike using the system to learn the given content
- Level 2: Learning
    - Assessment of how well the students learn the instructional content
    - To determine learning gain on the training content, a pretest (i.e., prior to training) and posttest (i.e., after training) on the content are usually given.
        - The effect size of the learning gain can be calculated using Cohen's *d* ([post-test score minus pre-test score] divided by [total questions minus pre-test score]).
- Level 3: Behavior
    - Evaluation to determine how much of what was learned can be applied on the job.
    - This is a measure of "near-transfer" of learning and should look at how much job performance/behaviors changed due to the training.
- Level 4: Results
    - Assessment of the overall impact the training had on the business or organization.
    - These results should be based on an overall objective that drove the training (e.g., a new training for teachers should improve student performance in the classroom).

Evaluating the system capabilities, learning gains, and transfer of training offers a comprehensive analysis of the benefits of a specific AIS/ATS. Potential purchasers of AIS/ATS should also consider the costs of system development

** The views expressed in this paper are solely those of the authors, and do not necessarily reflect the opinions of the Naval Air Warfare Center Training Systems Division, or any other Department of Defense agency, unless stated in official directives.*

and modification (i.e., to update content or the system adaptations) when making decisions on purchasing these types of systems. Adaptive systems have the potential to aid in improving performance, decreasing training time, and reducing training cost; however, effective evaluation of these intended features is critical to ensure the systems actually perform as promised.

REFERENCES

Campbell, G. E. (2014). Adaptive, intelligent training systems: Just how "smart" are they? In G.E. Campbell (Chair), *Adaptive Training Systems*. Symposium conducted at the Naval Air Systems Command Fellows Lecture Series, Orlando, FL.

Dzikovska, M., Steinhauser, N., Farrow, E., Moore, J., & Campbell, G. (2014). BEETLE II: Deep natural language understanding and automatic feedback generation for intelligent tutoring in basic electricity and electronics. *International Journal of Artificial Intelligence in Education, 24 (3),* 284-332. doi: 10.1007/s40593-014-0017-9

Graesser, A. C., Conley, M. W., & Olney, A. M. (2012). Intelligent tutoring systems. In K. R. Harris, S. Graham, & T. Urdan (Eds.), *APA Educational Psychology Handbook* (pp. 451–473). Washington, DC: American Psychological Association.

Kirkpatrick, D. L., & Kirkpatrick, J. D. (2006). *Evaluating Training Programs* (3rd ed.). San Francisco, CA: Berrett-Koehler Publishers.

Landsberg, C. R., Astwood, R. S., Van Buskirk, W. L., Townsend, L. N., Steinhauser, N. B., & Mercado, A. D. (2012). Review of adaptive training system techniques. *Military Psychology, 24 (2)*, 96-113.

Landsberg, C. R., Mercado, A. D., Van Buskirk, W. L., Lineberry, M., & Steinhauser, N. (2012). Evaluation of an adaptive training system for submarine periscope operations. *Proceedings of the Human Factors and Ergonomics Society, 56*(1), 2422-2426.

Landsberg, C. R., Van Buskirk, W. L., Astwood, R. S., Mercado, A. D., & Aakre, A. J. (2011) *Adaptive training considerations for simulation-based training systems* (Special report 2010-001). Defense Technical Information Center (DTIC).

Park, O., & Lee, J. (2003). Adaptive instructional systems. In Jonassen, D. (Ed.). *Handbook of Research for Educational Communications and Technology* (pp. 651–684). New York, NY: MacMillan Publishers.

Shute, V. J., Lajoie, S. P., & Gluck, K. A. (2000). Individualized and group approaches to training. In S. Tobias & J. D. Fletcher (Eds.), *Training and Retraining: A Handbook for Business, Industry, Government, and the Military* (pp. 171-207). New York, NY: Macmillan.

** The views expressed in this paper are solely those of the authors, and do not necessarily reflect the opinions of the Naval Air Warfare Center Training Systems Division, or any other Department of Defense agency, unless stated in official directives.*