A Theoretical Framework for Multimodal Learner Modeling and Performance Analysis in Experiential Learning Environments

Caleb Vatral [1], Gautam Biswas[1] and Benjamin Goldberg [2]

[1] Institute for Software Integrated Systems, Vanderbilt University, Nashville, TN USA
[2] U.S. Army DEVCOM Soldier Center, Orlando, FL USA

Abstract

This paper presents a theoretical framework for multimodal learner modeling, performance analysis, and feedback generation in personalized experiential learning environments. Effective learning and training programs are crucial in modern workplaces that require complex cognitive and psychomotor skills in individual and team settings. Experiential learning environments have gained popularity for teaching these skills in realistic yet safe contexts. Our framework for assessing learner performance in these environments integrates cognitive task analysis, distributed cognition, and multimodal learning analytics. Drawing on competency-based education theories, the framework introduces a hierarchical learner competency model and Bayesian inferencing techniques to assess individual and team performance. This model provides insights into learner performance at different cognitive levels and tracks performance and behaviors over time. The information derived from multilevel learner modeling is then used to generate explainable and actionable feedback for learners and instructors. This feedback aims to support personalized needs and facilitate improvement.

Keywords

Multimodal, Human Performance, Automated Assessment, Feedback, After Action Review, Experiential Learning

1. Introduction

Success in modern work environments requires the effective interplay of complex cognitive and psychomotor skills in individual and team settings for competent performance outcomes. Because of this increasing complexity and demand, effective learning and training programs have to teach these skills in realistic but safe environments. One of the most common paradigms for this sort of learning and training are personalized experiential learning environments, which train learners in realistic scenarios through direct experiences with task material. The increasing proliferation and demand for experiential learning environments has led to the need for effective modeling techniques that can measure and respond to the personalized learning performance and behavior of individuals and teams of learners. However, an effective learning model in such complex environments comes with a set of unique challenges that include assessment of learner performance and learning behaviors.

In this paper, we present a theoretical framework for multimodal learner modeling, performance analysis, and feedback generation to support debriefing and after-action reviews. Grounded in the theories of competency-based education and experiential learning, our framework combines cognitive task analysis, distributed cognition, and multimodal learning analytics to create a systematic method for analysis of learner behaviors and performance, and generation of effective feedback interventions. Our framework begins with a cognitive task analysis (CTA) of the training environment, generating a detailed hierarchical model of the cognitive, behavioral, and psychomotor tasks and strategies employed

Workshop on Artificial Intelligence in Support of Guided Experiential Learning, Held in conjunction with the International Conference on Artificial Intelligence in Education (AIED) 07 July 2023, Tokyo, Japan
EMAIL: caleb.m.vatral@vanderbilt.edu (A. 1)
ORCID: 0000-0001-5710-2911 (A. 1); 0000-0002-2752-3878 (A. 2); 0000-0002-1396-9887 (A. 3)

by the learners. This CTA model is then combined with a qualitative distributed cognition analysis of the captured learner data to generate needs analysis and feature engineering to support multimodal learning analytics (MMLA) that can effectively structure and interpret the raw learner data. Then, building on the theories of competency-based education, this MMLA is combined with a newly introduced hierarchical learner competency model to generate performance assessments. The hierarchical learner competency model utilizes Bayesian inferencing techniques to generate insights about learner's cognition at multiple levels of abstraction and map them on to performance metrics that can be evaluated across time. Finally, the insights generated by this multilevel learner modeling can be used to generate explainable and actionable feedback, which is presented back to the learners and their instructors.

2. Case-Study: Army Battle Drill Training

In the paper, we will demonstrate each component of our learner modeling frame-work using examples of soldiers training on dismounted battle drills in mixed reality environments. We will then present lessons learned from this implementation and how these techniques might be applicable to a larger class of learning and training environments through the lens of this case study.

Figure 1. The two dismounted battle drills run on the SAM-T used for this study. Left: Break Contact; Right: Enter and Clear a Room

In Fall 2021, several infantry fire teams participated in a study at the Fort Camp-bell US Army Installation where they performed training on two dismounted battle drills: Enter and Clear a Room (ECR) and Break Contact (BC) using the Squad Advanced Marksmanship Trainer (SAM-T) system. In the ECR drill, each team is tasked with entering a room that they may not have experienced before, neutralizing all enemy combatants, securing all civilians, and then safely exiting the room. The team must follow proper protocols and best practices while entering and operating in the room to minimize the risks of casualties to the team and civilians who may be present in the room. In the BC drill, each team is tasked with exploring a region where enemy forces may be present, and upon making contact with the enemy forces, breaking the contact and retreating to a safe distance. While breaking contact, the team must follow proper protocols and best practices to avoid casualties from enemy or friendly fire.

All the training took place using the SAM-T system, shown in Figure 1, which is a mixed reality battle simulator that projects a Virtual Battle Space 3 (VBS3) simulation onto a set of screens placed in a simulated physical environment. The team moves around in the physical environment while interfacing with the simulation using modified weapons that have a digital interface. In addition, trained instructors can modify the simulation in real-time to help their trainees develop or reinforce skills and become more proficient in the task domain. The SAM-T system records log information related to events that occur in the simulation, and this can include weapon fire messages, entity positions and health states, and so on. In addition, using the data collector in the Generalized Intelligent Framework for Tutoring (GIFT), we have also collected video and audio of the trainees psychomotor, speech, gaze, and posture as they enact tasks in response to the simulation scenario. Data from these sensors are then synchronized with the VBS3 and SAM-T logs automatically. The data recorded in log files can be played back through a GIFT Gamemaster interface to support debriefing and after-action review.

3. Background

Competency-based education has gained significant attention in recent years as a powerful approach to learning and assessment. It is a learner-centered paradigm that focuses on the mastery of specific competencies or skills rather than the completion of a fixed curriculum (Gervais, 2016). By putting emphasis on the demonstration of knowledge, abilities, and behaviors, CBE aligns with the demands of modern work-places, where the focus is on attainment and application of specific skillsets that are relevant to the success of organizations. Experiential learning, on the other hand, is a pedagogical approach that emphasizes learning through direct, typically realistic experiences and reflection on those experiences (Kolb, 2014). Therefore, it involves active engagement with real-world tasks, problems, and challenges, allowing learners to construct their knowledge and develop relevant skills in relevant contexts. Experiential learning is highly valued for its ability to bridge the gap between theory and practice, providing learners with opportunities to apply their knowledge in authentic contexts.

The integration of competency-based education with experiential learning has long been the subject of research in both theory and implementation and is now often referred to as a completely integrated research space known as competency-based experiential learning (CBEL) (Lewis & Williams, 1994; Reynolds, 1981; Hernandez et al., 2022). CBEL recognizes the importance of both competency mastery and the experiential nature of learning and aims to create effective learning environments where learners can acquire and apply competencies through immersive experiences, enabling them to develop a deep understanding of the subject matter and the ability to transfer knowledge to real-world situations. CBEL forms the basis for our theoretical framework, which offers a holistic approach to learner modeling and performance analysis by acknowledging the importance of context and application, as well as the individualized nature of learning.

4. Theoretical Framework

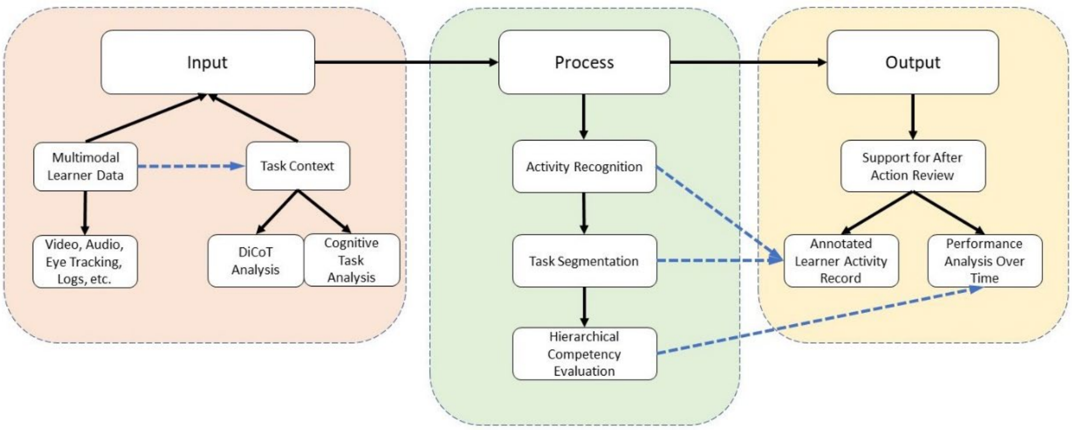

Figure 2. The high-level structure of the theoretical framework for learner modeling, evaluation, and feedback generation

At a high-level, our theoretical framework can be described using the standard Input-Process-Output (IPO) structure as illustrated in Figure 2. As input to the model, we first take multimodal learner data. This data can come in many forms depending on the specific training environment, but often includes things such as video, audio, eye tracking, and system logs. This data is used qualitatively analyzed and used to form the basis of the task context, which is the other input to the system. Since each task has unique attributes that require associated unique analyses, we add the task context to the model as an input. In our model, this task context takes the form of Cognitive Task Analysis and DiCoT Analysis, which will be described in detail in the next sections. Taking in this raw learner data and task context, the model then processes the data using Multimodal Learning Analytics algorithms, which breakdown into three phases. First, we perform action recognition, which takes the raw unstructured learner data and converts it into a structured and interpretable form. Based on these interpretable actions, we then perform as task segmentation, which breaks down the complete training exercise into meaningful smaller chunks which can be analyzed in detail. Each of these chunks is then passed to the hierarchical competency evaluation phase, which takes the learner data and produces a set of assessments based on the expected behaviors defined in the task context. Finally, this processed learner data and associated assessments are passed to the output phase, which formats them into actionable feedback for the end user that can be used to support debriefing and after-action review processes. This learner feedback takes two forms: (1) an annotated learner activity record which allows instructors and trainees to quickly and easily navigate and review a training session and (2) performance analysis over time which allows instructors and trainees to easily monitor learner progress and identify areas for improvement.

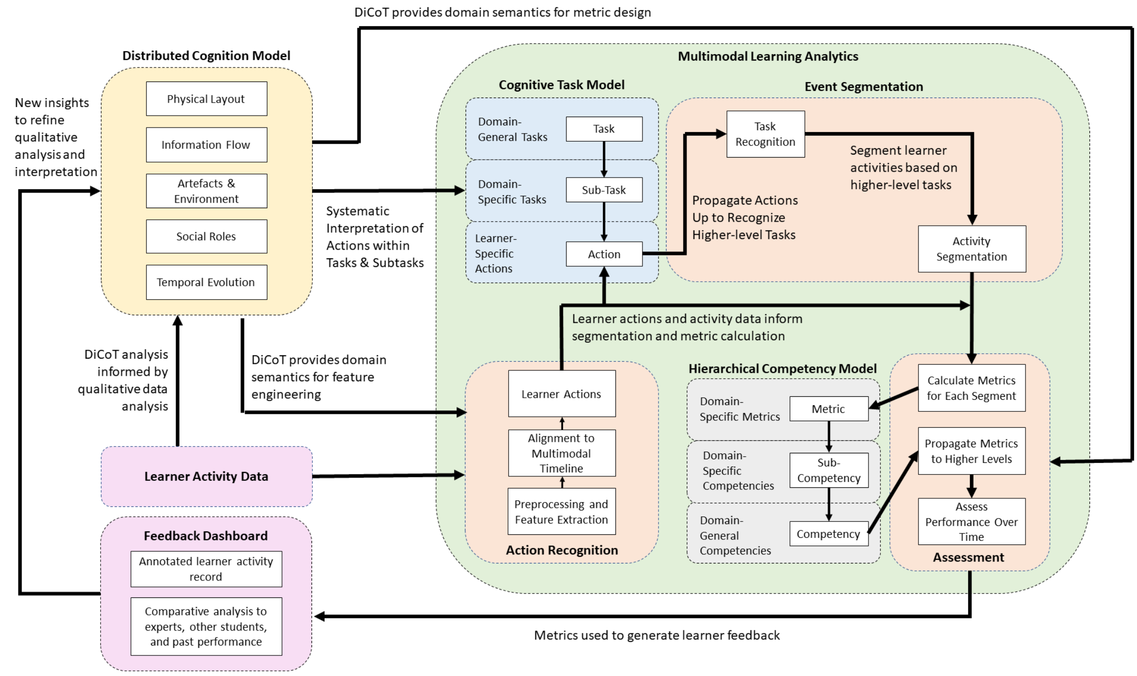

Figure 3. The complete theoretical framework for multimodal learner modeling and performance analysis

Now having laid the foundation for the theoretical framework for multimodal learner modeling and performance analysis at a high-level, we will examine each component and their associated computational processes in depth. The computational structure of the framework is shown in Figure 3 and breaks down into three primary high-level components: the learner data and feedback (purple); the distributed cognition model (yellow); and the multimodal learning analytics (green), which breaks down further into the cognitive task model (blue), the hierarchical competency model (gray), and the computational processes that connect them (orange).

4.1. Cognitive Task Analysis

At the center of the analysis framework is the cognitive task model, which represents the primary learner model of the system, and the first component of the larger MMLA block. A cognitive task analysis (CTA) of the training environment serves as the starting point for understanding the specific competencies and skills in which the learners have to become proficient. Cognitive task analysis is a structured approach to building a model with the help of human experts, which helps us explicate the cognitive processes and skills required to perform competently in a particular domain (Clark & Estes, 1996; Zachary et al., 2000). A cognitive task model is constructed by combining theoretical models and foundations of the analyzed domains with insights generated from interviews with domain experts and qualitative review of collected learner data. The task model assumes a hierarchical structure, with the upper levels representing more aggregated cognitive constructs, and each deepening level captures the breakdown of the aggregated cognitive process into component processes. The lowest levels of the hierarchy typically represent concrete and observable learner behaviors. High levels are decomposed into these lower levels such that the nodes present at level i are combined and sequenced to construct the nodes at level i+1. At the lowest level of the model, we call these nodes actions, which represent the smallest relevant and observable unit of learner behavior. Because of its hierarchical structure, the cognitive task model represents a method to generate inferences about high-er-level cognitive constructs, behaviors, strategies, and plans by understanding sequences of lower-level observable learner behavior. In other words, the cognitive task model takes low-level observable actions and links them to higher-level cognitive processes and behaviors. An example cognitive task model for the BC drill in our case, adapted from the domain-general H-ABC model (Vatral et al., 2022a), study is shown in Figure 4.

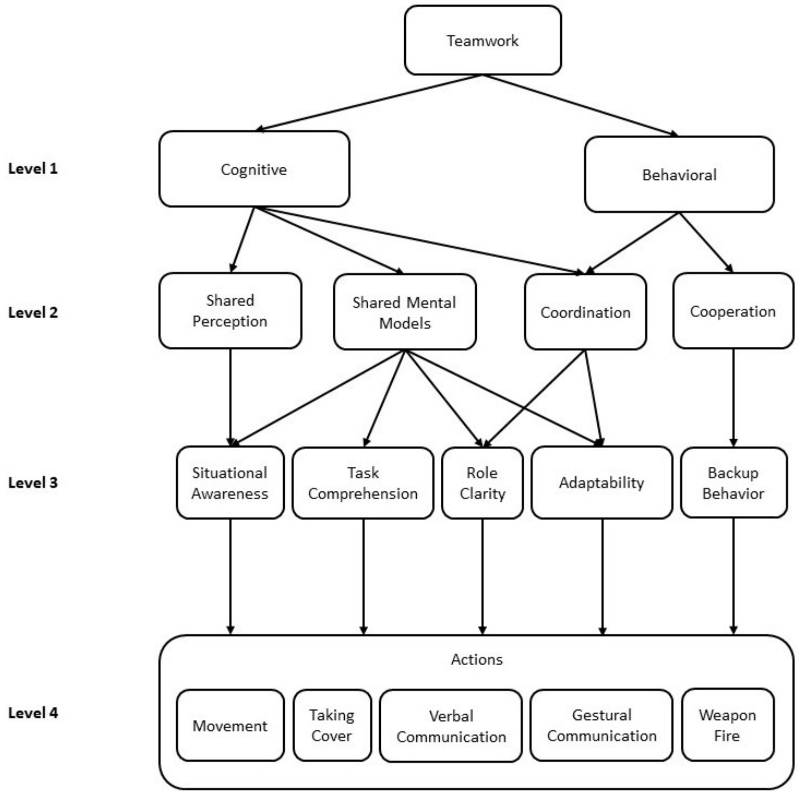

Figure 4. Example task model for the Break Contact drill in our case study

The cognitive task model takes low-level observable actions and links them to higher-level strategies and behaviors. Based on these higher-level strategies and behaviors, we can then perform Event Segmentation, which breaks down the complete training scenario into smaller sections that can be individually analyzed by the competency model, which will be discussed later. Event segmentation is an important step for the assessment process, as each higher-level behavior in the task model may have vastly different requirements, and thus be assessed vastly differently. For example, consider the ECR drill in our case study. The requirements of the team when entering the room, including rapid succession, active firing, moving along the walls, etc., are vastly different from the requirements when exiting the room, including taking the shortest path, disengaging weapons, and vocalizing exits. Evaluating the performance of the team in these highly different phases of the exercise requires the ability to segment the task into these separate phases. In other words, making sense of such com-plex behavior requires a systematic way of breaking it down into smaller units. This segmentation into smaller units would then be passed to the assessment model.

However, for complex and open-ended training environments, the cognitive task model alone does not provide sufficient information to completely disentangle high-er-level behaviors. For example, consider movement patterns in the BC drill from our case study. Without further context, we do not know whether simple movement of the team is part of the initial exploration phase or whether it is part of the break phase after contacting the enemy forces. In order to disambiguate to which phase this movement belongs, we have to introduce additional domain information.

4.2. Distributed Cognition

In our theoretical framework, this information comes in the form of constraints created from the Distributed Cognition analysis using the DiCoT framework. Our theoretical framework extends the insights generated by CTA by combining the CTA model with additional domain information provided by a distributed cognition analysis. Distributed cognition analysis complements the CTA by examining how learners interact with the learning environment and distribute cognitive processes across tools, artifacts, and other individuals. This analysis helps uncover the distributed nature of learning and cognition, highlighting the ways in which learners leverage external re-sources and collaborate with others to enhance their performance (Hutchins, 2000). In our work, we perform a distributed cognition analysis of the training environment using the Distributed Cognition for Teamwork (DiCoT) framework, which provides a structured qualitative methodology for analyzing the distributed cognition of a group of trainees working together (Blandford and Furniss, 2006). DiCoT breaks down dis-tributed cognition into 5 themes: (1) information flow, (2) artefacts and environment, (3) physical layout, (4) social interactions, and (5) temporal evolution. Within each of these themes, DiCoT defines a set of principles, which represent a form of qualitative codes that can be identified, categorized, and analyzed by examining learners as they participate in their training. This qualitative analysis generates a set of domain knowledge which complements and extends the insights generated by CTA.

Continuing the example of BC, our DiCoT analysis might have revealed that in the training environment, the enemy threat is positioned on the screen. So, movement of the soldiers away from the screen would be interpreted as a retreat, which is part of the break phase. From this simple additional information, we can gather that, in our case study, if movement is toward the SAM-T screen, then it is part of the exploration phase, and if movement is away from the screen, it is part of the break phase. In practice, without this additional domain-specific information from the DiCoT model, the task model would be unable to tell what phase of the drill the soldiers are per-forming. These sorts of probabilistic constraints generated by qualitative DiCoT analysis offer a systematic method of applying a-priori domain knowledge to im-prove the overall learner model.

4.3. Multimodal Learning Analytics

However, in addition to these probabilistic constraints, the DiCoT analysis also pro-vides information to the computational processes of the MMLA pipeline, specifically providing domain semantics for feature engineering for action recognition, and metric design for the competency model. While CTA and DiCoT provide a strong theoretical background for our comprehensive framework, these two qualitative models alone do not provide the tools to analyze learner data at scale, nor do they provide a means of assessing learner performance. To fill this gap, our framework utilizes multimodal learning analytics (MMLA). MMLA is a field that explores the fusion of different data modalities to gain deeper insights into the learning process (Blikstein & Worsley, 2016; Ochoa et al., 2017). Traditionally, learning analytics focused primarily on analyzing data from a single modality, such as student assessments or clickstream data. However, with the advancement and proliferation of sensor technologies, researchers and educators have recognized the potential of incorporating multiple modalities to capture a more comprehensive view of learners' experiences. This is especially important for our framework, as traditional logs, clickstreams, and other single modalities do not provide enough information to capture the complexity of experiential learning, which often involves significant psychomotor and teamwork components.

Multimodal learning analytics leverages various data sources, including text, audio, video, physiological signals, and interaction data, among others. By integrating these diverse modalities, researchers can extract rich and contextualized information about students' cognitive, affective, and behavioral aspects during learning activities. This interdisciplinary approach draws from fields like machine learning, computer vision, natural language processing, and data mining to develop sophisticated algorithms and techniques that can analyze, interpret, and visualize multimodal data. The ultimate goal is to enhance educational practices, personalize learning experiences, and provide timely and targeted interventions to improve students' learning out-comes.

In our previous work, we have combined MMLA with cognitive task analysis to understand learner behavior in various learning environments, including traditional computer-based tasks and physical experiential learning tasks. In Biswas et al. (2020), we used a CTA approach to analyze trainee performance in a counterterror-ism simulation called UrbanSim. By recording trainee actions within the simulation environment, we observed their performance at various levels of abstraction and employed quantitative measures to capture their cognitive and metacognitive processes. In Vatral et al. (2021), we employed CTA to develop a hierarchical model and corresponding quantitative metrics for the Enter and Clear a Room (ECR) dismount-ed battle drill. Building upon this research, Vatral et al. (2022a) expanded the CTA methods to investigate teamwork behaviors alongside individual performance and develop the extensible H-ABC teamwork model, which connected overarching teamwork concepts to domain-specific performance metrics. This model was later applied in Vatral et al. (2022b) to analyze the ECR dismounted battle drill from a team perspective. In all of this previous work, we applied MMLA methods with various different modalities, including logs, audio, and video, to quantitatively evaluate performance on the given training tasks. Building upon this series of work, in this paper, we formalize the methods of applying MMLA to CTA and DiCoT into our comprehensive framework. Our framework breaks down the MMLA components into two processes: Action recognition and the hierarchical competency model.

4.3.1. Action Recognition

Action recognition, in the case of our theoretical framework, is the process of converting raw sensor data collected from sensors in the environment into interpretable learner actions. Continuing with the BC drill example, we previously discussed movement toward or away from the screen as an important indicator, but we did not specify how we were able to interpret that movement was occurring. Within the theoretical framework, the action recognition block would be responsible for this step. The exact specifics of the action recognition process depend very heavily on a combination of what sensors are equipped in the task environment and what actions need to be determined. In our previous work (Vatral et al., 2021; 2022a; 2022b), we used primarily computer vision algorithms for the multimodal analysis. However, these methods are designed to be general, so that are number of analysis algorithms could be applied based on the data available and the specifics of the desired outcomes.

Regardless of the sensors and algorithms employed in this step, it is the information from the DiCoT analysis which allows us to define a set of features and potential actions that can occur in the environment. For example, in the BC drill, we already showed the importance of movement patterns, but DiCoT could identify other important components of soldier behavior that require specific action recognition; for example, taking cover behind objects in the space, maintaining a low-to-ground posture when not moving, laying down cover fire, etc. The DiCoT analysis defines the important information, features, and potential actions that need to be calculated used the action recognition step.

4.3.2. Hierarchical Competency Modeling

The CTA and DiCoT analysis also provides domain-specific information to define the Hierarchical Competency Model and its associated assessment steps. The hierarchical competency model is a structured model for providing assessment of learner competency at multiple levels of abstraction. The competency model is structured very similarly to the cognitive task model, containing many of the same high-level components, but instead of actions at the low levels, the competency model has assessment metrics at the lowest levels. These assessment metrics are crafted based on the DiCoT analysis of the environment, as well as consultation with domain experts and relevant literature, and the metrics are designed to be computable based on the directly observable learner data and associated action recognition. For example, when evaluating the BC drill, one metric might evaluate whether the trainees are staying close to cover when they are not moving.

The parallel structure between the competency model and the cognitive task model allows us to take the event segmentation generated by the task model and generate the relevant assessment metrics for each segment depending on what task is being performed. These assessments for each segment are then propagated up the model to generate assessments of higher-level performance. To perform the propagation, we model the competency model as a dynamic Bayesian network, with directly computable low-level metrics representing the evidence variables, and higher-level competencies representing the unobservable variables conditionally linked to the evidence variables. By observing learner data, computing the relevant metrics, and propagating to high-level competencies over time during multiple training instances, we can then model the progression of learner competency over time. For more information on this Bayesian propagation, see Vatral et al. (2022b).

4.4. Feedback for After Action Review

Finally, the last component of the theoretical framework refers to the output given to the end user. The overall goal of this research, as well as research in AIED more generally, is ultimately to improve the learning outcomes of the students and trainees in these environments, so feedback to learners and instructors is a critical component of such learning analytics systems. In this work, we present such feedback using data dashboards which are designed to present conclusions from the analysis in a visual and understandable way to the stakeholders. The design of our feedback dashboard is grounded in theories of explainable AI applied to learning and training environments for debriefing and after-action review. The idea is that users can be presented with data and insights generated from the system in an annotated format that allows the users to understand both how the underlying algorithms work, as well as how the data might be useful for improving performance. Different learning environments often have differing requirements for an effective learning analytics dashboard, but three fundamental principles remain the same in the design of each system: (1) feed-back should be objective and data-driven; (2) users should understand how and why the feedback was generated; (3) all feedback should be designed to supplement and assist, but not replace, traditional instructors.

In our case study, we facilitate the feedback generation using the Gamemaster interface in the Generalized Intelligent Framework for Tutoring (GIFT), whether our case-study system was implemented (Goldberg et al., 2021). In GIFT, the gamemaster is an interface where users can review performance on current and past training exercises. In our work, we have extended the Gamemaster interface to display additional performance metrics to the user alongside video which evidences how these metrics were calculated. The idea of this presentation is that the user could under-stand how and why the feedback was generated by looking at the video evidence. In addition to these metrics and videos, we also extended the Gamemaster interface to display a comparison between the trainees' performance and an expert model of what should have occurred in the situation. This form of expert comparison is a commonly used strategy for giving feedback to learners in many forms of learning and training environments. For more information on the implementation of our feed-back mechanisms in GIFT, see Vatral et al. (2021) and Vatral et al. (2023a).

5. Implementing the Framework

Now that we have described the general structure of the framework and each of its theory-grounded components, we will give a brief summary of how to implement the framework for a new training exercise or domain based on the lessons learned implementing it for both the ECR and BC drills.

The first step toward implementing this framework is building out the cognitive task model. For most training domains, there exists a set of best practices and procedures that trainees are trying to learn. These practices may be explicitly enumerated in literature related to the training environment, but they also may need to be elicited from experts in the field via interviews and co-design procedures. The important les-son to take away from our implementations is to keep the CTA model general at the high levels and domain-specific at the low levels. This structure allows parts of the model to be re-used across related domains and for multiple drills to evidence the same high-level competencies. An example of this kind of structure is the H-ABC model initially developed in Vatral et al. (2022a) and later utilized in Vatral et al. (2022b; 2023b).

After developing the CTA model, the next step is to collect some sample training data on which we can perform the qualitative DiCoT analysis. As previously de-scribed, the DiCoT analysis will break down the training domain into its five themes and code for their associated 18 principles. After performing the initial analysis, we then compare the identified DiCoT principles to the CTA model to find areas of over-lap that form the basis for the domain-specific constraints. To do this, we take a second pass through the sample training data and attempt to manually code event seg-mentation paths through the CTA model. When we reach an ambiguous point where multiple paths through the CTA model could be plausible, such as those described in the previous section, we find associated principles that have been identified in the DiCoT analysis that can help to clarify these ambiguities and generate a constraint. This process is repeated until all of the sample training data can be definitively ana-lyzed using the CTA model and the generated constraints.

With the completed joint DiCoT and CTA model, we then examine the low-level actions in the CTA model and use DiCoT constraints to develop action recognition features. As previously described, this process will differ depending on what data is available in the training environment, but the general idea is to determine how to calculate each of the defined actions and constraints using the available data. For example, in both of our case study drills, movement is a critically important action because of the physicality of the environment. So, we examine the data that was collected and discover that we can approximate the movement patterns of the soldiers using motion tracking on the video data. Then, we examine the constraints important constraints and define an algorithm to compute a path through the CTA model based on the actions (i.e., motion tracking movement data) and the constraints (e.g., movement toward or away from the screen). After defining these algorithms for each action and associated constraint, the event segmentation procedure can be computed for any new data.

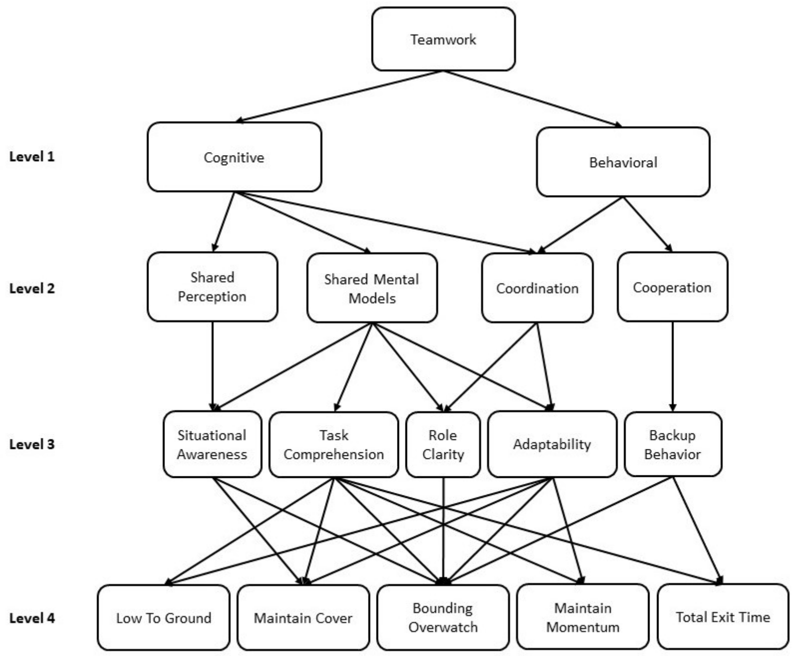

Figure 5. Hierarchical competency model for the Break Contact drill in our case study

Next, we define the hierarchical competency model (HCM). The HCM is structured in parallel to the CTA model, so much of the work for the high levels is already done. However, for the lowest levels of the model, we have to replace the actions in the CTA model with associated performance metrics that can be calculated for each event segment. This process is again largely based on the predefined best practices and procedures of the training task, which will be elicited from a combination of relevant literature, expert interviews and co-design, and qualitative analysis of the sample training data. In general, the metrics in the HCM will parallel the actions in the CTA model as well, but to a lesser degree than the higher levels which can be ported exactly. For example, compare the HCM for the Break Contact drill shown in Figure 4 to the CTA model for the Break Contact drill shown in Figure 5. There are clearly many parallels between the two, such as the Maintain Cover metric parallel to the Take Cover actions, for example, but the metrics are a more detailed breakdown of actual performance, rather than just the actions taken. Once defined, the low-level metrics can be computed for each event segment identified by the CTA model and the higher-level competencies can be inferred using the Bayesian inference procedures described in Vatral et al. (2022b; 2023b).

Finally, once all of the above components are defined, we can define the feedback component of the framework. This component will again depend heavily on the ex-act specifics of the training environment, but there are some strategies which are readily available from the framework's previous steps which can be easily leveraged to generate feedback. First, the framework provides a method of generating an annotated learner record using the action recognition and event segmentation. For example, in our case study we generate a video timeline which is annotated with the basic actions of the soldiers (e.g., they entered the room, they took cover, etc.), as well as the higher-level segments of the soldiers' actions (e.g., entering the room vs securing the room vs leaving the room, or exploration of the area vs retreat from contact). Second, the framework provides a method of generating a progress monitoring visualization. Since the basis for evaluation in the framework is a multi-level competency model, we can easily plot the performance at each level of the model over the course of training to give students and instructors insights into how they are progressing at multiple levels of abstraction. Finally, the framework also provides a method of generating performance comparisons. These comparisons could include trainees to their own past performance, trainees to one another, or trainees to expert models. Regard-less of the type of comparison, since competency evaluations are data-driven met-rics, it is easy to compare performance on similar training drills.

6. Conclusions

In this paper, we presented a comprehensive theoretical framework for analyzing trainee performance across multiple levels of abstraction using multimodal data from the training environment. Based on the theories of CBEL, CTA, DiCoT, and MMLA, the framework provides a complete analysis path to take raw learner data and develop computational models that generate both learner competency evaluations learner feedback. Throughout the paper, we presented the framework through the lens of an example implementation of soldiers training on dismounted battle drills. We concluded the work by outlining the complete process for implementing the framework with suggestions and lessons learned based on our case-study implementation. Future work should focus on implementing this framework in a variety of settings other than Army battle drill training, which would further validate the frame-work's efficacy and ensure that it is comprehensive. In addition, future work should focus on developing open-source tools to support the common components of the framework, such as the event segmentation and hierarchical competency modeling steps. These open-source tools would lower the barriers to entry for the framework and allow for consistent implementations that can plug and play between different training domains. With continued development, we hope this framework can represent a comprehensive analysis routine for a wide variety of training domains and play an integral part in standardizing assessment and feedback in competency-based experiential learning environments.

7. References

[1] Blandford, A., and Furniss, D. (2006). "Dicot: a methodology for applying distributed cognition to the design of teamworking systems," in Interactive Systems. Design, Specification, and Verification, eds S.W. Gilroy and M. D. Harrison (Berlin; Heidelberg: Springer Berlin Heidelberg), 26–38.

[2] Blikstein, P., & Worsley, M. (2016). Multimodal learning analytics and education data mining: Using computational technologies to measure complex learning tasks. Journal of Learning Analytics, 3(2), 220-238.

[3] Biswas, G., Rajendran, R., Mohammed, N., Goldberg, B. S., Sottilare, R. A., Brawner, K., & Hoffman, M. (2019). Multilevel learner modeling in training environments for complex decision making. IEEE Transactions on Learning Technologies, 13(1), 172-185.

[4] Clark, R. E., & Estes, F. (1996). Cognitive task analysis for training. International Journal of Educational Research, 25(5), 403-417.

[5] Gervais, J. (2016). The operational definition of competency-based education. The Journal of Competency-Based Education, 1(2), 98– 106. doi: 10.1002/cbe2.1011

[6] Goldberg, B., Owens, K., Hellman, K., Robson, R., Blake-Plock, S., Hoffman, M., Gupton, K. (2021). Forging Competency and Proficiency through the Synthetic Training Environment with an Experiential Learning for Readiness Strategy. In Proceedings of the 2021 I/ITSEC. Orlando, FL.
[7] Hernandez, M., Goldberg, B., Robson, R., Owens, K., Blake-Plock, S., Welch, T., & Ray, F. (2022). Enhancing the Total Learning Architecture for Experiential Learning. In Interservice/Industry Training, Simulation, and Education Conference (I/ITSEC).
[8] Hutchins, E. (2000). "Distributed cognition," in International Encyclopedia of the Social and Behavioral Sciences (Amsterdam: Elsevier Science), 138.
[9] Kolb, D. A. (2014). Experiential learning: Experience as the source of learning and development. FT press.
[10] Lewis, L. H., & Williams, C. J. (1994). Experiential learning: Past and present. New directions for adult and continuing education, 1994(62), 5-16.
[11] Ochoa, X., Lang, A. C., & Siemens, G. (2017). Multimodal learning analytics. The handbook of learning analytics, 1, 129-141.
[12] Reynolds, S. (1981). A marriage proposal: Competency based education and experiential learning. Journal of Experiential Education, 4(2), 34-39.
[13] Vatral, C., Mohammed, N., Biswas, G., & Goldberg, B. S. (2021). GIFT External Assessment Engine for Analyzing Individual and Team Performance for Dismounted Battle Drills. In Proceedings of the Ninth Annual Generalized Intelligent Framework for Tutoring Users Symposium (GIFTSym9) (pp. 107-127). US Army DEVCOM – Soldier Center. (ISBN 13: 978-0-9977257-9-7).
[14] Vatral, C., Biswas, G., & Goldberg, B. S. (2022a). Multimodal Learning analytics using hierarchical models for analyzing team performance. In Proceedings of the 15th International Conference in Computer Supported Collaborative Learning (pp. 403-406). International Society of the Learning Sciences.
[15] Vatral, C., Mohammed, N., Biswas, G., & Goldberg, B. S. (2022b). Multimodal Learning analytics using hierarchical models for analyzing team performance. In Proceedings of the 2022 Interservice/Industry Training, Simulation and Education Conference (I/ITSEC). National Training and Simulation Association.
[16] Vatral, C. Mohammed, N., Biswas, G., Roberts, N., & Goldberg, B. S. (2023a). A Comparative Analysis Interface to Streamline After-Action Review in Experiential Learning Environments. In Proceedings of the 2023 Annual Generalized Intelligent Framework for Tutoring Users Symposium (in press). US Army DEVCOM – Soldier Center.
[17] Vatral, C., Mohammed, N., Biswas, G., & Goldberg, B. S. (2023b). A Framework for Performance Assessment Across Multiple Training Scenarios Using Hierarchical Bayesian Competency Models. Under Review for Proceedings of the 2023 Interservice/Industry Training, Simulation and Education Conference (I/ITSEC). National Training and Simulation Association (Under Review).
[18] Zachary, W. W., Ryder, J. M., & Hicinbothom, J. H. (2000). Building cognitive task analyses and models of a decision-making team in a complex real-time environment. Cognitive task analysis, 365-384.