

Leveraging Privileged Multimodal Data to Model Cognitive States in an Interactive Learning Platform

Deniz Sonmez Unal
University of Pittsburgh
Pittsburgh, PA 15260
des204@pitt.edu

Theresa Mowad
Lehigh University
Bethlehem, PA 18015
tgm320@lehigh.edu

Catherine M. Arrington
Lehigh University
Bethlehem, PA 18015
cma205@lehigh.edu

Erin Solovey
Worcester Polytechnic Institute
Worcester, MA 01609
esolovey@wpi.edu

Erin Walker
University of Pittsburgh
Pittsburgh, PA 15260
eawalker@pitt.edu

ABSTRACT

Student models in intelligent tutoring systems (ITS) often rely on large-scale behavioral data (e.g., correctness, response time), which provides only indirect evidence of cognition. While multimodal data such as think-alouds or screen recordings can offer richer insight into students' cognitive states, such data is rarely available at scale, limiting the practicality of multimodal models. To address this, we leverage Learning Using Privileged Information (LUPI) framework for cognitive state modeling in ITS. Our approach incorporates high-quality think-aloud data and written action descriptions from a subset of students at training time to inform model decision boundaries, while using only behavioral data at inference. This setup mimics real-world deployment where rich modalities are unavailable for most learners. We evaluate this method in modeling rule learning, an established framework for conceptual development during problem solving, and find improved predictive accuracy, particularly for nuanced cognitive states not well captured by behavioral logs alone.

Keywords

Learning Using Privileged Information, Rule Learning, Multimodal Learning Analytics

1. INTRODUCTION

Interactive learning platforms such as intelligent tutoring systems (ITS) aim to provide personalized learning support by modeling students' knowledge, behavior, and cognitive states [38]. These models are often built using large amounts of data collected during student interactions with the system. Among these, behavioral log data such as correctness, response time, or hint usage are the most widely used sources of information, largely because they are easy to collect and

available across many systems. However, while behavioral data are abundant and convenient, they are also discrete and indirect, often failing to capture the more subtle or continuous aspects of learners' cognitive processes [19].

In recent years, the field has increasingly turned to multimodal approaches to capture richer representations of student learning [3, 10, 13]. These approaches integrate behavioral logs with other data sources such as eye tracking, facial expressions, neural signals, and verbal protocols to gain deeper insight into students' thought processes and cognitive states. Among these, think-aloud protocols are particularly valuable for identifying real-time cognitive activity due to their interpretability [21]. Think-aloud protocols are widely used across HCI and cognitive science as a method for capturing users' internal reasoning in real time [14, 28]. They involve asking individuals to verbalize their thoughts while performing a task in order to gain access to moment-to-moment reactions and cognitive processing that are otherwise difficult to infer. In educational contexts, think-aloud protocols serve a similar purpose: they reveal how learners engage with problems, what they understand, and where they struggle. This form of verbal data has been shown to provide insights into learners' cognitive processes, and specifically in uncovering processes like self-regulated learning strategies [17, 43] and metacognitive monitoring and control [25, 31] that remain invisible in log data. Complementing these verbalizations, another heavily used modality is external observations which captures behaviors that learners may not articulate explicitly. This may include researcher coding of video data, screen recordings, or in situ observations of classroom behaviors. Prior work has shown such observations can reveal cognitive and affective states, and instances of disengagement [1, 27].

Even though multimodal methods are extremely helpful in providing a broader picture of students' cognition, a key limitation of many multimodal methods is that they rely on non-deployable signals or data that are easy to collect in controlled lab settings, but infeasible for real-world environments. As a result, models trained on rich multimodal data often suffer from limited generalizability and are difficult to apply in practice [26]. To address this challenge, we propose framing cognitive state modeling within the Learning

Using Privileged Information (LUPI) paradigm [39]. LUPI is a machine learning technique in which models are trained using additional information, called privileged information, that is available only at training time but not at test time. This privileged information can include richer or more explanatory features that help the model learn a better decision boundary during training. At inference time, the model operates solely on the standard input features. It ensures generalization while benefiting from the additional insight gained during training. In this work, we hypothesize that LUPI enables the use of rich but harder-to-collect information such as verbal think-alouds or action descriptions during model training, while maintaining deployability by relying only on behavioral data at inference time. In doing so, it offers a powerful approach to bridge the gap between what learners experience and what can be captured externally.

In this work, we model phases of rule learning [7] as our target cognitive states. The rule learning framework offers a formalization for cognitive states, such as moving from unfamiliarity to partial understanding to fluent application, and reflects core aspects of skill acquisition [33]. Our key contributions are as follows: (1) We introduce a novel application of LUPI to model rule learning states in ITS contexts, using multimodal data that combines behavioral logs with verbal protocol data. (2) We demonstrate that incorporating verbal protocol data during training significantly improves model performance, particularly for cognitive states that are challenging to infer from behavior alone.

2. RELATED WORK

2.1 Multimodal Student Modeling

Beyond behavioral log data, researchers have increasingly turned to multimodal approaches to capture a richer picture of learning processes. These approaches integrate multiple data streams such as eye-tracking, speech, hand-writing to improve student modeling and provide insight into cognitive and affective states that are difficult to detect from logged behavioral information alone [8].

Textual data has been widely explored within multimodal student modeling, since it provides direct evidence of learners’ reasoning and cognitive states in real time [2]. Within this broader category, think-aloud protocols represent a particularly valuable form of textual data. Recent advances in educational data mining and AI in education have increasingly incorporated think-aloud protocols as rich sources of cognitive and metacognitive data for machine learning applications. Zhang et al. demonstrated the potential of large language models to automatically detect self-regulated learning (SRL) behaviors from transcribed think-alouds, enabling predictive modeling of learning strategies without the need for manual annotation [43]. Similarly, Li et al. leveraged topic modeling and supervised classification on think-aloud transcripts to predict students’ reasoning activities during computer-assisted learning, showing that TAP-derived features can robustly model cognitive patterns [21]. Tabarsi et al. triangulated verbal and behavioral data, using coded think-aloud segments alongside video and interaction logs to train data-driven detectors for identifying programming difficulties among novice learners [35].

Another complementary modality is human observation, where

researchers code student behaviors to capture aspects of learning not visible in interaction logs. Early work in ITS relied heavily on expert observers to identify cognitive and affective states. For example, Baker et al. [1] used classroom observations of Cognitive Tutor users to code when students were gaming the system, and these observations provided the training data for generalizable automated detectors. Similarly, the Baker-Rodrigo Observation Method Protocol (BROMP) [27] has been used to code affective states such as boredom, frustration, and confusion. Other studies have relied on expert coding of more complex behaviors from screen replays or video recordings [34]. D’Mello et al. combined human annotation of affect during learning sessions with system logs to train multimodal models of affect detection [12]. Paquette et al. [29] examined expert coding of disengagement in online learning environments. They showed that coders rely on contextual cues rather than isolated actions when making judgments about student behavior. Recent studies have explored human-coded observational data as features in predictive models. Tian et al. demonstrated how textual artifacts such as rubric-based classroom observations and field notes can be encoded and used in supervised learning to predict instructional quality and engagement [36]. Wu utilized graph convolutional networks (GCNs) on structured classroom observation datasets to evaluate teaching effectiveness at scale [42]. In another study, Petković et al. combined manual coding of teacher nonverbal behaviors with computer vision and machine learning techniques to assess instructional immediacy and its impact on student engagement [30]. In our work, we use written notes from screen recordings as a form of human observation. This allows us to capture detailed traces of students’ actions during problem solving and relate them to underlying cognitive states.

2.2 Learning Using Privileged Information

The Learning Using Privileged Information (LUPI) paradigm was first introduced by Vapnik and Vashist [39]. The key idea is that during training, the model has access to additional *privileged* information that is not available at test time but can help the learner form a more accurate decision boundary. This setup mirrors real-world scenarios in which richer explanatory features (e.g., expert feedback, annotations, or multimodal data) are accessible during the development of a model but are impractical or impossible to collect consistently during deployment. Vapnik describes this idea as taking inspiration from human learning: a teacher may provide additional explanations, hints, or demonstrations when introducing a challenging concept, enabling the student to build a deeper understanding. However, once the student sits for an exam, these supports are no longer available, and the student must rely only on what they have internalized.

Researchers have used LUPI in a number of different HCI and educational settings. Vrigkas et al. [40] used LUPI for a human activity recognition task. They introduced a model that combined standard video features with additional data such as audio, pose data, or semantic annotations as privileged information. Their results showed that including this privileged information improved recognition accuracy across several activity datasets. Another line of work applies LUPI to affective computing, focusing on the

challenge of transferring models trained in the lab to real-world settings. Makantasis et al. [24] introduced the use of privileged information to bridge this gap. In their framework, high-quality multimodal signals available in controlled lab settings (e.g., audio, facial features, and physiological measures such as ECG and EDA) were treated as privileged information, while raw video frames were considered the only modality available “in the wild”. They showed that performance could approach that of full multimodal models, even in noisy, uncontrolled conditions. Wang et al. [41] also used LUPI in emotion recognition, where EEG signals and video features were treated as complementary modalities: each served as privileged information for improving the other task. This work illustrated how LUPI can exploit correlations across modalities to enhance models when one modality is unavailable at test time. In the education domain, Zhao et al. [44] explored LUPI for graduate admissions prediction. They proposed S3VM+, which combined semi-supervised learning with LUPI to handle biases in committee decisions across years and to exploit large amounts of unlabeled application data. By treating yearly committee differences as privileged information, their approach outperformed standard baselines in identifying top-performing students, demonstrating LUPI’s potential in educational data mining.

Our work builds on this line of research by applying the LUPI paradigm to model cognitive states in intelligent tutoring systems. Similar to the activity recognition and affective computing studies, our approach addresses the gap between rich but impractical signals and deployable features. Like Zhao et al., we operate in an educational context, but instead of admissions, we focus on modeling fine-grained cognitive states during problem solving.

2.3 Rule Learning

Inductive reasoning plays a central role in learning. It involves observing instances of a phenomenon, identifying patterns and relationships among them, and forming general rules that can be applied to new situations. The sequence of steps in this process is often described as the phases of rule learning. In cognitive science, these phases have traditionally been examined through variations of the rule attainment task [4, 7]. Such tasks are designed to capture three key phases of rule learning: rule search, rule discovery, and rule following. Typically, participants respond to a series of stimuli, indicating whether they believe the stimuli follow a rule or appear random. For example, in one version, learners see sequences of numbers that conform to a simple mathematical pattern (e.g., +2 leading to 32, 34, 36...) [20]. In another, they view numbered circles where the position of a blue circle shifts according to an underlying rule [5]. Across these tasks, specific behavioral markers make it possible to identify rule learning phases independently of the exact stimuli. This offers a standardized framework for analyzing learning processes across contexts.

Building on this tradition, our study adopts an extended set of five rule learning phases as the formalization of cognitive states (Table 1). These phases reflect how the constructs originally studied in abstract tasks translate into real-world educational problem solving:

Rule Search: Segments where learners express uncertainty or speculate without confidence about how to proceed.

Rule Discovery: Segments where learners suddenly recognize a useful principle or pattern that helps solve the problem.

Rule Following: Segments where learners apply a correct rule confidently and without hesitation.

Rule Violation: Segments where learners attempt to apply a previously correct rule but make a slip or error.

Follow Wrong: Segments where learners confidently apply a rule they believe is correct, but it reflects a misconception.

While most prior work has focused on abstract laboratory tasks, these phases have also been shown to manifest in real-world educational problem solving [37, 33]. In this study, we adopt the phases of rule learning as our formalization of cognitive states. This framing allows us to capture how learners progress from uncertainty to understanding and application. Higher-level abstractions, such as knowledge components or skills, are useful for structuring instruction but often too coarse to capture the nuanced, moment-to-moment shifts that occur during problem solving. In contrast, rule learning phases parallel these abstractions while making visible more subtle states that are otherwise difficult to define or detect. At the same time, the phases are sufficiently well-defined in cognitive science to provide a stable target for modeling.

3. LEARNING TASK DESIGN

Our study was conducted using the ASSISTments platform, an online learning environment that supports the creation of custom problems, interactive scaffolds, and adaptive feedback. [18]. We authored 9 probability problems in the platform. Each problem in our task included multiple probability rules, which corresponded to the underlying skills or concepts required for successful problem solving. The rules covered in our problem set were: 1) identifying the number of favorable outcomes, 2) identifying the number of possible outcomes, 3) calculating probability of an event, 4) finding the probability of only one event when two events can occur together, 5) finding the probability of either of the two events happening but not both, 6) finding the conditional probability of an event, and 7) finding the probability of dependent events. Students were given access to on-demand scaffolding that allows them to break down problems into subproblems as needed (see Figure 2). Each scaffolded subproblem was associated with a single probability rule only. In addition to requesting scaffolds, students could also ask for a hint or reveal the answer if they were completely stuck.

To capture students’ cognitive states throughout the learning process, we introduced a knowledge indication panel that prompts learners to self-report their current rule learning phase. Students respond to this panel whenever a new problem or scaffold appears on screen, before they can proceed with solving (see Figure 1). Their choices reflect phases such as “I know how to solve this problem” or “I don’t know how to solve this problem” to initially label students as being in a rule search or rule following state.

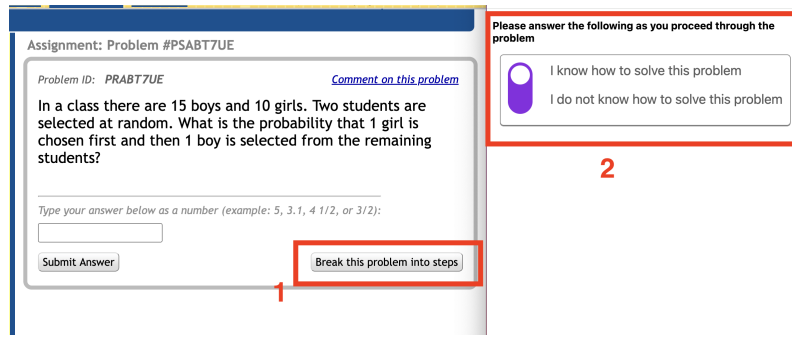


Figure 1: Screenshot of a sample problem. 1: Participants ask to divide the problem into smaller steps by clicking on this button. 2: The knowledge indication panel. Participants need to choose one of the two states on the toggle to answer the current problem on the screen.

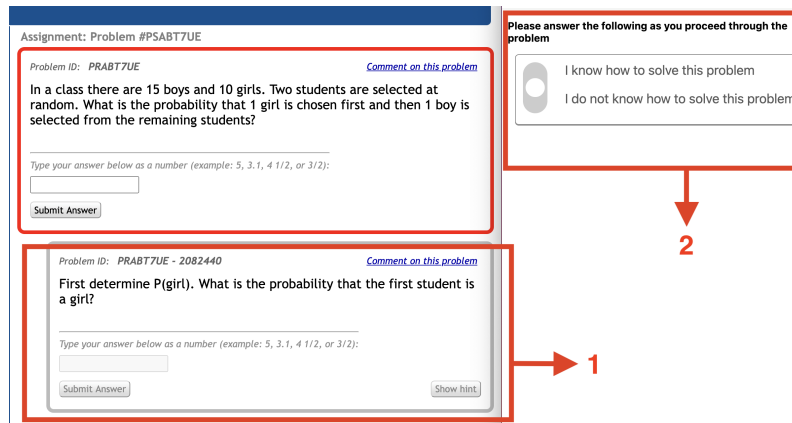


Figure 2: A sample problem showing how scaffolding is provided. 1: Once the participant clicks on the “Break this problem into steps” button shown in Figure 1, the first substep of the problem unfolds. 2: Participants continue to respond to the toggle to indicate their knowledge on the current problem for the substeps of the problem as well. For each new problem shown on the screen (either a main problem or a substep), the toggle is reset and the current problem is disabled until the participant switches the toggle.

In cases where a student’s problem-solving response contradicts their indicated knowledge state (e.g., incorrect answer despite claiming to know how to solve the problem), the system triggers additional prompts to refine our understanding of their learning status. These follow-up questions help disambiguate whether the student was having a misconception (follow wrong state), was just making a guess (rule search state), made a simple mistake like a typo or miscalculation (rule violation state), or encountering a different kind of difficulty (Figure 3).

4. DATASET

4.1 Experimental Procedures

42 university students completed a combination of problem solving tasks and a rule learning task. All participants gave their informed consent prior to their inclusion and received course credit. They agreed to their audio and visual information being recorded and the collected data was de-identified prior to analysis. All study procedures were reviewed and approved by the university’s Institutional Review Board. Participants began the experiment by solving a set of probability problems using ASSISTments. Initially,

they were introduced to a practice set of problems to familiarize themselves with the platform. After completing the practice problems, participants proceeded to the pretest, which included 7 probability questions designed to be completed within 15 minutes. Upon completion of the pre-test, participants engaged in the rule learning task. Participants were instructed on how they could access the on-demand scaffolding by clicking on “Break This Problem Into Steps” button as well as how to use the right hand panel to indicate whether they knew how to solve the problem. After the practice, participants proceeded to solve 9 probability problems that were designed to be completed within 20 minutes using ASSISTments. They were allowed to use pen, paper, and a calculator. They were instructed to think aloud while solving the problems. The final phase involved a post-test consisting of 7 problems similar to those in the pre-test.

We collected a multimodal dataset during a structured problem solving task in the ASSISTments online tutoring system. The dataset includes behavioral interaction logs, verbal think-aloud data, and screen recordings of participants’ sessions.

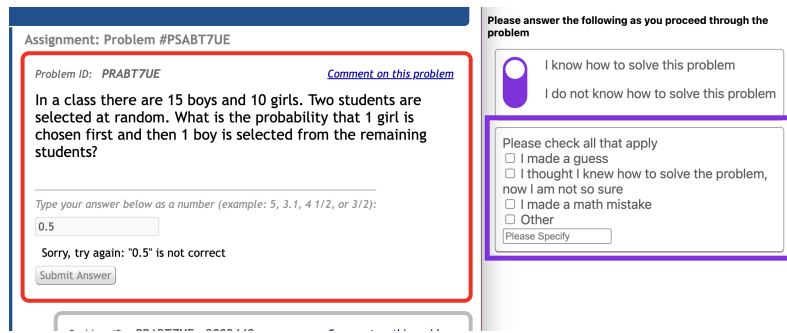


Figure 3: If the participant claims to know how to solve the current problem but provides an incorrect response or asks for scaffolding, another panel with additional options for the knowledge indication (highlighted in purple) unfolds to clarify the reason for the incorrect answer or the help request.

4.2 Behavioral Log Data

As participants interacted with the problems, we recorded detailed behavioral logs. These included the current action the participant was performing on the ASSISTments platform (submitting an answer, requesting a hint, requesting to break the problem into steps, or revealing the solution), correctness of the participants’ answers to the problems if the action was an answer attempt, the items participants click on the knowledge indication panel, and timestamps of each action. The resulting dataset contains 2942 rows, with each row representing an action taken by a specific participant during the session.

4.3 Verbal Think-Aloud Data

During the task, students were asked to think aloud, verbalizing their reasoning as they worked through each problem. These verbalizations were recorded and subsequently transcribed to produce sentence-level textual data aligned with the students’ actions. The transcriptions capture cognitive and metacognitive insights, such as moments of uncertainty, strategic decisions, and explanations of problem-solving steps.

4.4 Observation Notes from Screen Recordings

In addition to verbal data, we collected screen recordings of each session. These recordings were reviewed by trained coders to generate written descriptions of student actions, particularly for cases where system logs lacked sufficient context. These descriptions include observations such as hesitation before submitting an answer, switching between problem elements, or interpreting problem prompts. When combined with the think-aloud transcripts, these descriptions help contextualize the behavioral trace with higher granularity.

4.5 Labeling Cognitive States

To label cognitive states, we used a hybrid annotation approach based on both self-report data and manual coding. As mentioned in Section 3, Students were asked to complete a knowledge indication panel prior to solving each problem or scaffolded subproblem. This panel prompted them to indicate their perceived level of knowledge about the underlying rule, providing a continuous stream of self-reported rule learning states.

In cases where students provided verbal think-alouds aligned with specific actions, we used these transcriptions, along with the screen-recorded behaviors, to refine the state labels. Annotation was guided by a structured coding scheme for rule learning phases adapted from Unal et al. [33] (see Table 1), allowing coders to disambiguate ambiguous or inconsistent self-reports based on verbal evidence.

On top of the five rule learning states described above, we also observed a few instances of gaming the system in the screen recordings. Following Baker et al. [1], we flagged behaviors such as rapidly clicking through hints, repeatedly guessing until correct, or otherwise exploiting system feedback. These cases were coded but excluded from analysis, as gaming reflects disengagement rather than the cognitive processes underlying rule learning.

To ensure the reliability of our coding scheme, two coders independently annotated 20% of the dataset. Interrater agreement was assessed using Cohen’s kappa, yielding a value of 0.84. Frequency of each label is presented in Table 2.

5. METHOD

We implemented a LUPI framework using knowledge distillation which has become one of the standard ways for training models with information that is only available at learning time [22]. In our setup, different “teacher” models are trained with both behavioral features from interaction logs and privileged features from either student think-alouds or observation notes from screen recordings or a combination of both privileged data sources. The teacher models’ outputs are then used to guide a “student” model that only has access to behavioral features (see Section 5.2 for more details). This design allows the student models to benefit from the richer representations learned by their teachers, while remaining deployable in real-world settings where privileged features are not available.

5.1 Feature Representation

In this modeling scheme, we have 2 sets of features. The first set is regular features that are extracted from student interaction logs within the ASSISTments platform. These features reflect observable student behavior and are available for real-time inference. Our second feature set is called the privileged feature set. Privileged features consisted of

Table 1: Coding Scheme Labels and Descriptions.

Label	Description	Example
Rule Search	Segments where learners express uncertainty or lack of knowledge about how to proceed. They may be speculating, experimenting, or trying different approaches without confidence.	“I’m not sure if this is right... maybe I should divide these numbers first and see what happens.”
Rule Discovery	Segments where the learner suddenly recognizes a useful pattern or principle that helps solve the problem.	“Wait, I see now—we need to subtract the overlapping part from the total.”
Rule Following	Segments where the learner applies a correct rule with confidence, explaining their reasoning clearly and arriving at an answer without hesitation.	“Since the events are independent, I just multiply 3/10 by 2/5 to get 6/50.”
Rule Violation	Segments where the learner attempts to apply a previously correct rule but makes a mistake, often due to a slip or calculation error.	“Oops, I added those fractions and got 7/12... oh, I should have found a common denominator first.”
Follow Wrong	A case where the learner confidently applies a rule they believe to be correct, but it is actually a misconception. They consistently follow the wrong logic.	“To find the total probability, we always add the numbers—so 0.3 plus 0.6 gives 0.9.” (when multiplication is required).

Table 2: Frequency of cognitive states in the dataset.

Follow Wrong	Rule Discovery	Rule Following	Rule Search	Rule Violation
259	446	1248	898	91

verbal thinkaloud data, in which students articulated their reasoning while solving problems, and post-hoc observation notes transcribed from screen recordings. See Table 3 for the full list of features and their descriptions.

Verbal data streams were encoded using the all-MiniLM-L6-v2, a Sentence Transformer model that maps sentences to a 384 dimensional dense vector space to produce dense vector embeddings [32]. For instances where no think-aloud or observation note was available, we inserted a special “non textual” tag to mark the absence of verbal data.

5.2 Models & Training

The training pipeline involves two types of models: a teacher model trained with both behavioral and privileged features, and a student model trained solely on behavioral features using soft labels generated by the teacher. The idea is that the teacher, having access to richer information that privileged features provide, produces a probability distribution over the labels rather than a single hard label. The student model is then trained on this distribution instead of hard one-hot targets. Because the teacher’s output reflects how it weighs similarities between classes, for example, assigning probability mass to both Rule Search and Follow Wrong states when the evidence is ambiguous, the student model gains a better sense of which states are easily confused. This provides a smoother and more informative training signal, allowing the student model to learn about subtle distinctions that are especially difficult to detect from behavioral features alone. As a result, the student can generalize more effectively to uncertain or hard-to-classify cases.

Let $x \in \mathbb{R}^d$ represent the behavioral feature vector and

$x^* \in \mathbb{R}^{d^*}$ the privileged feature vector derived from participants’ verbal think-alouds and/or observation notes from screen recordings. The teacher model f_T , receives the concatenated input $x_{full} = [x, x^*] \in \mathbb{R}^{d+d^*}$ and produces logits $z_T = f_T(x_{full}) \in \mathbb{R}^k$ where k is the number of cognitive states. The teacher is trained to minimize the standard cross-entropy loss:

$$LCE = - \sum_i y_i \log(\text{softmax}(z_T)_i) \quad (1)$$

where $y \in \{0, 1\}^K$ is the one-hot encoded ground truth label.

For distillation, we applied temperature scaling to soften the teacher’s output distribution. The softened probabilities are defined as:

$$p_T^{(i)} = \frac{\exp(z_T^{(i)}/T)}{\sum_{j=1}^K \exp(z_T^{(j)}/T)} \quad (2)$$

where $T > 1$ is the temperature parameter. Higher temperatures produce smoother distributions, providing the student with more informative gradients.

The student model f_S is trained on behavioral features x and outputs logits $z_S = f_S(x)$. Its probability distribution with temperature scaling is given by:

$$p_S^{(i)} = \frac{\exp(z_S^{(i)}/T)}{\sum_{j=1}^K \exp(z_S^{(j)}/T)} \quad (3)$$

The student is optimized using a hybrid loss that combines cross-entropy with ground-truth labels and the KL diver-

Table 3: Regular and privileged features and their explanations.

Space	Feature	Description
x	Action	The type of the action on the current problem. Possible action types are answer attempts, hint or scaffolding requests, self reflection actions on the knowledge indication prompt, and next-button clicks.
	FirstAction	Whether the type of the first action on the current problem was an answer attempt or hint request.
	Rule	The rules associated with the current problem will be represented with a 7-dimensional vector (for the 7 distinct rules in our problem set). Each problem is tagged with 1 to 3 rules.
	TimeOnProblem	Total time spent on the current problem measured in milliseconds.
	FirstRT	First response time to the current problem measured in millisecond.
	IsMainProblem	Whether the current problem is a main problem or a sub-problem.
	AttemptCount	Number of attempts on the current problem.
	NumRule	A 7-dimensional vector representing the cumulative count of exposures to each rule up to the current problem.
	RollingMeanTime	The rolling mean of the time spent on previous NumRule problem steps for each rule that is associated with the current problem.
	SuccessRate	Proportion of correctly solved steps in the previous NumRule steps for each rule that is associated with the current problem.
x_1^*	Verbal thinkalouds	Sentences in the transcriptions of student think-alouds were represented with dense sentence embeddings.
x_2^*	Observation notes	Verbal descriptions of students' actions as seen on the screen recordings were represented using the same methods as we did for x_1^* .

gence with the teacher’s soft labels:

$$\mathcal{L} = \alpha \mathcal{L}_{\text{CE}}(y, \text{softmax}(\mathbf{z}S)) + (1 - \alpha) T^2 \mathcal{LKD}(p_T, p_S) \quad (4)$$

where

$$\mathcal{LKD}(p_T, p_S) = \sum_i i = 1^K p_T^{(i)} \log \left(\frac{p_T^{(i)}}{p_S^{(i)}} \right). \quad (5)$$

Here, $\alpha \in [0, 1]$ controls the balance between matching the hard ground-truth labels and distilling knowledge from the teacher’s softened predictions. The T^2 factor accounts for the scaling effect of the temperature on gradients.

Both teacher and student models used the same feedforward neural network architecture with two hidden layers. Hidden layers had 128 units in teacher model and 64 units in the student model. Each hidden layer used ReLU activation functions and batch normalization. The output layer used a softmax activation to produce probability distributions over the cognitive states. Models were trained using the Adam optimizer with a learning rate of 0.001, batch size 32, and trained for 50 epochs. To select these settings, we performed a grid search over key hyperparameters, including learning rates $\{1e-2, 1e-3, 1e-4\}$, hidden layer sizes $\{64, 128, 256\}$, batch sizes $\{16, 32, 64\}$, temperature values $\{1, 2, 4, 8\}$, and hybrid loss weights $\alpha \in \{0.25, 0.5, 0.75\}$. The final configuration was chosen based on average F1 scores across folds on the validation set.

Because observation notes were produced by researchers, they were consistently high-quality and did not require fur-

ther stratification. In contrast, think-alouds varied in quality so we categorized participants’ think-aloud data into three groups to account for this variability. Participants who consistently articulated their reasoning and problem-solving strategies were classified as high-quality think-aloud participants ($n = 18$). Those who spoke but primarily read the problem text aloud or stated answers without reflecting on their reasoning were labeled as low-quality think-aloud participants ($n = 17$). Finally, participants who remained silent throughout were assigned to the silent group ($n = 7$). For model training, we stratified participants across these three groups when creating training and validation splits to ensure balanced representation. We used 5-fold cross validation at the participant level. We randomly oversampled participants who provided high quality think-alouds in the training folds to make sure that teacher models were exposed to richer privileged features and to strengthen the supervision signal for the student (LUPI) models. No data from the same participant appeared in both training and validation sets within a fold.

For evaluation, we computed per-class F1 scores on the validation set and. The baseline model was a behavioral-only network trained using cross-entropy loss with hard labels. Performance of the LUPI-trained student model was compared against this baseline and the privileged teacher to assess gains attributable to LUPI.

6. RESULTS

To evaluate the impact of incorporating privileged information during training, we compare performance of seven models using average F1 scores over folds. All models were trained and evaluated on the same participant-level cross-validation splits to ensure comparability. The teacher models differed in the privileged features (behavioral features were available to all models) they were given during training:

Teacher-Full had access to both think-alouds and observation notes, *Teacher-TA* had only think-alouds, and *Teacher-Obs* had only observation data. Each teacher produced soft labels (class probability distributions), which were then used to train the corresponding student (LUPI) models: *LUPI-Full*, *LUPI-TA*, and *LUPI-Obs*, that relied exclusively on behavioral features. In this setup, the student models did not differ in the features they used, but in the type of labels they were trained using: they learned from soft label distributions distilled from their respective teachers. Finally, we trained a baseline behavioral-only model using hard labels, without access to any privileged features or soft targets from a teacher.

Because evaluation was conducted using 5-fold participant-level cross validation on a relatively small dataset, we focus on the consistency and magnitude of observed performance trends rather than formal statistical significance testing. We first examined the overall performance of each model (Figure 4). The teacher models outperformed both the baseline and the LUPI models with *Teacher-Obs* achieving the highest average F1 score. This indicates that observation notes from screen recordings provided the strongest supervision signal. In contrast, *Teacher-TA* lagged behind the other teacher models which suggests think-alouds were a noisier source of privileged information.

The performance gains of the LUPI models over the baseline were modestly positive even though they did not match the teacher models' performance. This is natural as the LUPI models are smaller and rely only on behavioral features at inference. Importantly, the LUPI models achieved these gains while maintaining standard errors comparable to the baseline model, suggesting that the distillation process improved performance without increasing variability across folds. These results suggest that the distillation process provides a robust representation that generalizes well to unseen data, even in the absence of privileged information.

Having established the overall performance results, we next turn to a class-wise analysis of F1 scores to better understand where different types of privileged information contributed most to model improvements. As shown in Figure 5, performance varied across cognitive states. All models achieved very high F1 scores for Rule Following (above 0.94), suggesting that once participants are consistently applying the correct rule, behavioral information alone is largely sufficient for accurate classification. In Rule Discovery, performance differences were comparatively modest, with the baseline achieving an F1 score of approximately 0.81, LUPI models ranging between 0.79-0.81, and teacher models reaching 0.83-0.84. Unlike the more difficult states, the gains from privileged information in Rule Discovery were relatively small, suggesting that behavioral features already captured much of the information needed for classification. A similar pattern was observed in Rule Search, where the baseline already performed strongly (approximately 0.86 F1). Most LUPI models remained close to baseline performance, while *Teacher-Obs* achieved the strongest results (approximately 0.89 F1). This suggests that privileged information still provided additional benefit in distinguishing exploratory reasoning patterns, although the improvements were less pronounced than those observed in more subtle

cognitive states such as Follow Wrong and Rule Violation.

In contrast, privileged information had the strongest impact on the more subtle states. For example, in Follow Wrong, the baseline achieved an F1 of 0.78, while LUPI models improved slightly (0.79-0.80) and teacher models, particularly *Teacher-Obs* substantially higher performance (0.85). This reflects the fact that observation notes capture students' underlying strategies and misconceptions more directly than verbalizations or behavioral logs. Similarly, Rule Violation proved to be the most challenging state overall. The baseline was lowest at 0.59 but LUPI models improved to (0.68) and teacher models pushed F1 substantially higher with *Teacher-Full* at 0.82 and *Teacher-Obs* at 0.85. Interestingly, while *Teacher-TA* did not achieve comparable scores, the corresponding *LUPI-TA* model outperformed it. This can be attributed to the hybrid loss used in training, which allowed the LUPI model to balance between its teacher model's soft predictions and the direct behavioral signal. In cases where think-alouds were noisy or not informative, the LUPI model effectively "listened" more to the behavioral data, resulting in stronger generalization than its teacher.

Our results also highlight important differences between the types of privileged information used by the teacher models. Across most cognitive states, *Teacher-Obs* achieved the strongest or near-strongest performance, while *Teacher-Full* also consistently outperformed the baseline and LUPI models. This suggests that observation-based privileged information provided a particularly strong supervision signal for detecting students' cognitive states. At the same time, the competitive performance of *Teacher-Full* indicates that combining modalities can still provide complementary information, even if the additional think-aloud features did not always translate into higher F1 scores. When the modalities were examined separately, *Teacher-Obs* generally outperformed *Teacher-TA*. One explanation to why *Teacher-Obs* mostly outperformed *Teacher-TA* can be that the observation data was a stronger signal on its own. Even though think-alouds provided more raw data, they were often noisier, since many verbalizations are generic and occur across multiple states. For example, students may express surprise after giving an incorrect answer both when they are in a Follow Wrong phase (reflecting a deeper misconception or flawed mental model) and when they are in a Rule Violation phase (reflecting a simple slip despite otherwise correct knowledge). In contrast, observation data such as watching how students solve the problem on the calculator (on their computer screen) or sequence their inputs provides evidence of the strategies and mental models they are using. Because of this, observation features may have offered more discriminative information for distinguishing subtle differences between cognitive states.

These differences also carried over to the student models. While the gains over the baseline were more modest than those observed for the teacher models, the LUPI variants demonstrated that privileged information could still improve performance even when only behavioral features were available at inference. The effect was especially noticeable in the more challenging cognitive states, where the quality of the privileged supervision signal appeared to matter most. For example, in Follow Wrong, the *Teacher-TA* model per-

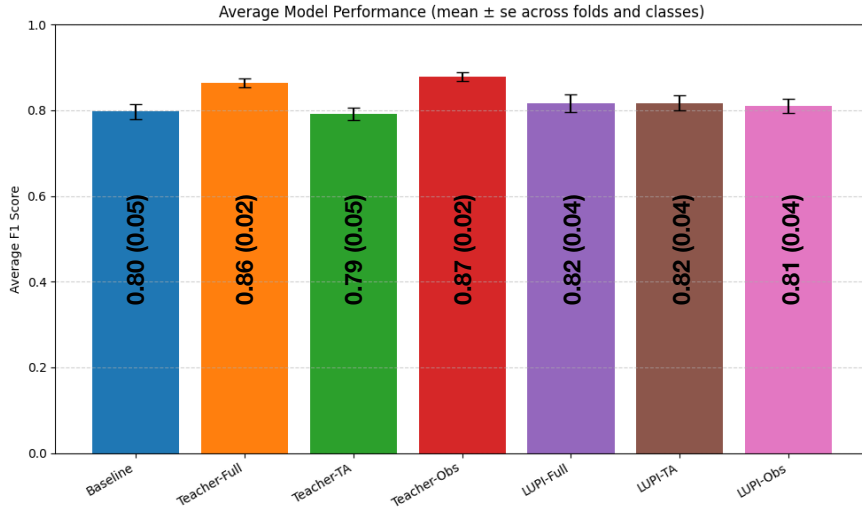


Figure 4: Comparison of Average F1 Scores and standard errors of the Models.

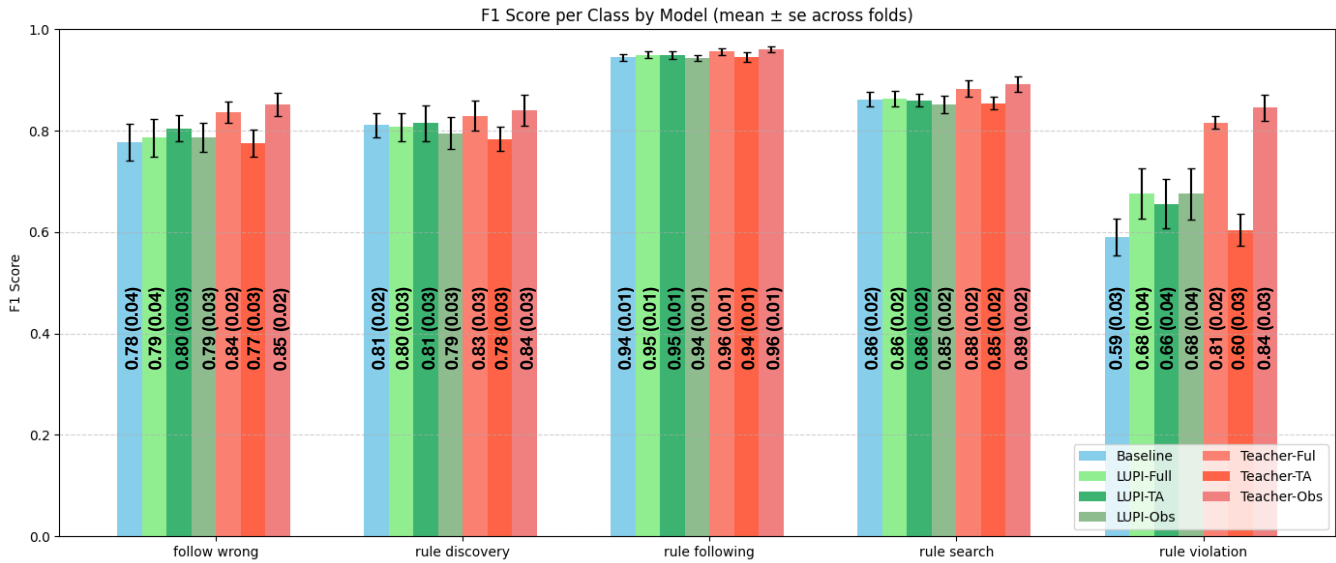


Figure 5: Mean class-wise F1 Scores and Standard Errors of the Models.

formed comparatively poorly, suggesting that think-alouds alone were often too noisy to reliably distinguish deeper misconceptions from neighboring states. However, the corresponding LUPI-TA model achieved stronger performance, likely because the hybrid distillation objective allowed it to balance between the teacher’s soft predictions and the ground-truth hard labels derived from behavioral data. In this sense, the student model was able to retain the most informative aspects of the privileged signal while reducing the impact of noisy verbalizations. A similar pattern ap-

peared in Rule Violation, the most difficult state overall. Here, observation-based privileged information proved especially valuable, as both Teacher-Obs and LUPI-Obs substantially outperformed their think-aloud counterparts. This suggests that observation data provided more discriminative cues about students’ underlying strategies and slips in reasoning, while think-alouds alone were often insufficiently specific. The strong performance of LUPI-Full further indicates that combining privileged modalities can still be beneficial when the student model learns to integrate comple-

mentary supervision signals through distillation.

7. DISCUSSION

In this work, our goal was to investigate whether incorporating privileged information through a LUPI framework can improve the modeling of cognitive states in a realistic learning environment. We used the phases of rule learning as our target cognitive states, because they offer an appropriate granularity for capturing nuanced, moment-to-moment shifts in cognition that are not easily represented in higher-level abstractions such as mastery / non-mastery in traditional knowledge tracing [6] or broader stages of skill acquisition in cognitive psychology such as cognitive and associative stages [15]. We compared the performance of teacher models with access to privileged information (think-alouds and observation data), student models trained on behavioral data only, optimized using soft labels distilled from those teachers, and a baseline model trained on the same behavioral data, optimized using hard labels. Our findings showed three main patterns. First, teacher models outperformed behavioral-only baselines, demonstrating that privileged information provides valuable insight into the learning process. Second, student models trained with teacher soft labels (LUPI) consistently outperformed the baseline despite relying only on behavioral data at test time. Third, these gains were most pronounced for subtle states such as Follow Wrong and Rule Violation, which are difficult to distinguish from behavioral features alone. The results also revealed differences between modalities. Observation data often provided a stronger standalone signal than think-alouds, but the best results were achieved by combining both. The models demonstrated strong generalization as all evaluations were conducted on unseen participants. The consistency of performance across folds highlights the robustness of our approach and the transfer of privileged information benefits to new learners is both reliable and effective.

Beyond predictive performance, the utility of modeling rule learning phases lies in the potential to inform feedback. By detecting nuanced states such as Follow Wrong or Rule Violation in real time, systems could provide targeted interventions that directly address misconceptions or slips without needing additional context about the specific problem, rather than treating all errors as equivalent. Similarly, identifying transitions into Rule Discovery opens opportunities for reflective prompts that help consolidate new understanding. This phase-level precision creates possibilities for adaptive support that is both more timely and more tailored than what is achievable with traditional correctness-based models.

Despite these promising results, our work has several limitations. First, the sample size was modest, which is common in this line of research [11]. While participant-level cross-validation mitigated the risk of overfitting, larger studies will be needed to validate the robustness of our findings. Second, while observation data proved valuable, it was generated through researcher coding of screen recordings, which may limit scalability. Future work should explore automated extraction of similar data. Third, think-aloud data were highly variable in quality, consistent with prior findings that verbal protocols can range from deeply reflective to surface-level or uninformative. Although we stratified participants by

think-aloud quality in our cross-validation splits, the variability highlights the need for methods to reduce noise.

Our research can lead to several future directions. Firstly, while this work focused on modeling rule learning phases, the proposed learning framework is not limited to this context. In principle, any type of privileged information that provides deeper access to students' cognitive or affective states could be incorporated at training time. For example, sensor-based modalities such as facial expressions, posture, or physiological signals have long been shown to capture affective dynamics relevant for learning [16], but their reliance on specialized equipment limits classroom deployment. Sensor-free affect detectors developed using behavioral proxies such as timing patterns, hint use, or keystroke dynamics already aim to approximate such states [9]. A promising direction for future work is to train LUPI models on multimodal or sensor-based privileged signals, while using sensor-free behavioral detectors as the deployable counterpart. This would allow researchers to take advantage of the richness of lab-collected affective data while maintaining scalability in real-world systems. A natural extension is to neuroscience, where privileged signals such provide direct access to neural mechanisms that underlie rule learning and other higher-order cognitive processes. Prior research has shown that distinct phases of rule learning are associated with signature patterns of brain activation, including shifts between prefrontal and parietal regions as learners move from rule search to rule following [20]. While such data are difficult to collect outside of controlled laboratory environments, they offer a uniquely fine-grained and objective window into cognition. A LUPI framework could allow us to leverage these neural signatures as privileged information during training, improving the ability of deployable models to infer subtle cognitive states in real time from behavioral logs alone.

Furthermore, an additional line of inquiry is to explore participant level training, where models are individualized based on richer data collected from each learner. Prior work in multimodal learning analytics and neuroscience has shown that such approaches can improve classification accuracy when sufficient data per participant is available [23]. While our study was not designed at that level of granularity, this type of approach could be especially valuable in settings with larger datasets per learner.

Finally, another promising line of research lies in improving the quality of verbal data used as privileged information. Thinkalouds are valuable for revealing reasoning processes, but as observed in our study and others [45], their quality varies widely across learners. Large language models (LLMs) have recently shown promise in evaluating, coding, and even scaffolding verbal explanations in educational contexts [43]. Leveraging LLM based evaluation could help filter noisy or surface level thinkalouds to enhance the reliability of verbal data as privileged information, and even provide real time feedback to learners as they articulate their reasoning. These directions point to a broader role for LUPI in learner modeling, serving as a bridge between modalities that are rich but impractical at scale and the behavioral data streams that are readily available in classrooms and tutoring systems.

8. ACKNOWLEDGMENTS

This work was supported by the National Science Foundation under awards #1835307 and #1912474.

9. REFERENCES

- [1] R. S. Baker, A. T. Corbett, K. R. Koedinger, and A. Z. Wagner. Off-task behavior in the cognitive tutor classroom: When students’ game the system”. In *Proceedings of the SIGCHI conference on Human factors in computing systems*, pages 383–390, 2004.
- [2] P. Blikstein. Multimodal learning analytics. In *Proceedings of the third international conference on learning analytics and knowledge*, pages 102–106, 2013.
- [3] P. Blikstein and M. Worsley. Multimodal learning analytics and education data mining: Using computational technologies to measure complex learning tasks. *Journal of learning analytics*, 3(2):220–238, 2016.
- [4] P. W. Burgess and T. Shallice. Bizarre responses, rule detection and frontal lobe lesions. *Cortex*, 32(2):241–259, 1996.
- [5] B. Cao, W. Li, F. Li, and H. Li. Dissociable roles of medial and lateral pfc in rule learning. *Brain and behavior*, 6(11):e00551, 2016.
- [6] A. T. Corbett and J. R. Anderson. Knowledge Tracing: Modeling the Acquisition of Procedural Knowledge. *User Modeling and User-Adapted Interaction*, 4(4):253–278, 1994.
- [7] C. Crescentini, S. Seyed-Allaei, N. De Pisapia, J. Jovicich, D. Amati, and T. Shallice. Mechanisms of rule acquisition and rule following in inductive reasoning. *Journal of Neuroscience*, 31(21):7763–7774, 2011.
- [8] M. Cukurova, M. Giannakos, and R. Martinez-Maldonado. The promise and challenges of multimodal learning analytics. *British Journal of Educational Technology*, 51(5):1441–1449, 2020.
- [9] R. S. d Baker, S. M. Gowda, M. Wixon, J. Kalka, A. Z. Wagner, A. Salvi, V. Alevan, G. W. Kusbit, J. Ocumpaugh, and L. Rossi. Towards sensor-free affect detection in cognitive tutor algebra. *International Educational Data Mining Society*, 2012.
- [10] D. Di Mitri, J. Schneider, M. Specht, and H. Drachslar. From signals to knowledge: A conceptual model for multimodal learning analytics. *Journal of Computer Assisted Learning*, 34(4):338–349, 2018.
- [11] X. Du, J. Yang, J.-L. Hung, and B. Shelton. Educational Data Mining: A Systematic Review of Research and Emerging Trends. *Information Discovery and Delivery*, 48(4):225–236, 2020.
- [12] S. K. D’mello, S. D. Craig, A. Witherspoon, B. McDaniel, and A. Graesser. Automatic detection of learner’s affect from conversational cues. *User modeling and user-adapted interaction*, 18(1):45–80, 2008.
- [13] A. Emerson, W. Min, J. Rowe, R. Azevedo, and J. Lester. Multimodal predictive student modeling with multi-task transfer learning. In *LAK23: 13th International Learning Analytics and Knowledge Conference*, pages 333–344, 2023.
- [14] K. A. Ericsson and H. A. Simon. Verbal reports as data. *Psychological review*, 87(3):215, 1980.
- [15] P. M. Fitts. Perceptual-Motor Skill Learning. In *Categories of Human Learning*, pages 243–285. Academic Press, 1964.
- [16] M. Giannakos, M. Cukurova, and S. Papavlasopoulou. Sensor-based analytics in education: Lessons learned from research in multimodal learning analytics. In *The multimodal learning analytics handbook*, pages 329–358. Springer, 2022.
- [17] J. A. Greene, V. M. Deekens, D. Z. Copeland, and S. Yu. Capturing and modeling self-regulated learning using think-aloud protocols. In *Handbook of self-regulation of learning and performance*, pages 323–337. Routledge, 2017.
- [18] N. T. Heffernan and C. L. Heffernan. The assistments ecosystem: Building a platform that brings scientists and teachers together for minimally invasive research on human learning and teaching. *International Journal of Artificial Intelligence in Education*, 24(4):470–497, 2014.
- [19] K. Holstein, Z. Yu, J. Sewall, O. Popescu, B. M. McLaren, and V. Alevan. Opening up an intelligent tutoring system development environment for extensible student modeling. In *International conference on artificial intelligence in education*, pages 169–183. Springer, 2018.
- [20] F. Li, B. Cao, H. Gao, L. Kuang, and H. Li. Different brain potentials evoked at distinct phases of rule learning. *Psychophysiology*, 49(9):1266–1276, 2012.
- [21] S. Li, X. Huang, T. Wang, J. Zheng, and S. P. Lajoie. Using text mining and machine learning to predict reasoning activities from think-aloud transcripts in computer assisted learning. *Journal of Computing in Higher Education*, 37(1):477–496, 2025.
- [22] J. Liu, D. Capurro, A. Nguyen, and K. Verspoor. Improving text-based early prediction by distillation from privileged time-series text. *arXiv preprint arXiv:2301.10887*, 2023.
- [23] R. Liu, E. Walker, L. Friedman, C. M. Arrington, and E. T. Solovey. fnirs-based classification of mind-wandering with personalized window selection for multimodal learning interfaces. *Journal on multimodal user interfaces*, 15(3):257–272, 2021.
- [24] K. Makantasis, K. Pinitas, A. Liapis, and G. N. Yannakakis. From the lab to the wild: Affect modeling via privileged information. *IEEE Transactions on Affective Computing*, 15(2):380–392, 2023.
- [25] M. Montague and B. Applegate. Middle school students’ mathematical problem solving: An analysis of think-aloud protocols. *Learning Disability Quarterly*, 16(1):19–32, 1993.
- [26] J. Ning, Z. Ma, J. Yao, Q. Wang, and B. Zhang. Personalized learning supported by learning analytics: a systematic review of functions, pathways, and educational outcomes. *Interactive Learning Environments*, pages 1–23, 2025.
- [27] J. Ocumpaugh et al. Baker rodrigo ocumpaugh monitoring protocol (bromp) 2.0 technical and training manual. *New York, NY and Manila, Philippines: Teachers College, Columbia University and Ateneo Laboratory for the Learning Sciences*, 60, 2015.
- [28] E. L. Olmsted-Hawala, E. D. Murphy, S. Hawala, and K. T. Ashenfelter. Think-aloud protocols: a

- comparison of three think-aloud protocols for use in testing data-dissemination web sites for usability. In *Proceedings of the SIGCHI conference on human factors in computing systems*, pages 2381–2390, 2010.
- [29] L. Paquette, A. M. de Carvalho, and R. S. Baker. Towards understanding expert coding of student disengagement in online learning. In *CogSci*, 2014.
- [30] U. Petković, J. Frenkel, O. Hellwich, and R. Lazarides. Modeling teacher nonverbal immediacy: Refinement and analysis. *Adaptive Behavior*, page 10597123251374920, 2025.
- [31] A. Ramachandran, C.-M. Huang, E. Gartland, and B. Scassellati. Thinking aloud with a tutoring robot to enhance learning. In *Proceedings of the 2018 ACM/IEEE international conference on human-robot interaction*, pages 59–68, 2018.
- [32] N. Reimers and I. Gurevych. Sentence-bert: Sentence embeddings using siamese bert-networks. In *Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing*. Association for Computational Linguistics, 11 2019.
- [33] D. Sonmez Unal, E. Solovey, C. M. Arrington, and E. Walker. Modeling the phases of rule learning during problem solving with an interactive learning environment. *User Modeling and User-Adapted Interaction*, 35(1):7, 2025.
- [34] Ö. Sümer, P. Goldberg, S. D’Mello, P. Gerjets, U. Trautwein, and E. Kasneci. Multimodal engagement analysis from facial videos in the classroom. *IEEE Transactions on Affective Computing*, 14(2):1012–1027, 2021.
- [35] B. Tabarsi, H. Reichert, R. Qualls, T. Price, and T. Barnes. Exploring novices’ struggle and progress during programming through data-driven detectors and think-aloud protocols. In *2023 IEEE Symposium on Visual Languages and Human-Centric Computing (VL/HCC)*, pages 179–183. IEEE, 2023.
- [36] Z. V. Tian, M. Sun, A. Liu, S. Sarkar, and J. Liu. Instructional improvement: leveraging computer-assisted textual analysis to generate insights from educational artifacts. In *Research Handbook on Classroom Observation*, pages 190–207. Edward Elgar Publishing, 2025.
- [37] D. S. Unal, C. M. Arrington, E. Solovey, and E. Walker. Using thinkalouds to understand rule learning and cognitive control mechanisms within an intelligent tutoring system. In *International Conference on Artificial Intelligence in Education*, pages 500–511. Springer, 2020.
- [38] K. VanLehn. The behavior of tutoring systems. *International journal of artificial intelligence in education*, 16(3):227–265, 2006.
- [39] V. Vapnik and A. Vashist. A new learning paradigm: Learning using privileged information. *Neural networks*, 22(5-6):544–557, 2009.
- [40] M. Vrigkas, E. Kazakos, C. Nikou, and I. A. Kakadiaris. Human activity recognition using robust adaptive privileged probabilistic learning. *Pattern Analysis and Applications*, 24(3):915–932, 2021.
- [41] S. Wang, Y. Zhu, L. Yue, and Q. Ji. Emotion recognition with the help of privileged information. *IEEE Transactions on Autonomous Mental Development*, 7(3):189–200, 2015.
- [42] S. Wu. Application of artificial intelligence graph convolutional network in classroom grade evaluation. *Scientific Reports*, 15(1):32044, 2025.
- [43] J. Zhang, C. Borchers, V. Aleven, and R. S. Baker. Using large language models to detect self-regulated learning in think-aloud protocols. In *Proceedings of the 17th international conference on educational data mining*, pages 157–168, 2024.
- [44] Y. Zhao, B. Lackaye, J. G. Dy, and C. E. Brodley. A quantitative machine learning approach to master students admission for professional institutions. *International Educational Data Mining Society*, 2020.
- [45] L. Zheng, A. He, C. Qi, H. Zhang, and X. Gu. Cognitive echo: Enhancing think-aloud protocols with llm-based simulated students. *British Journal of Educational Technology*, 2025.