GLOW: Implementation and Programmatic Details

Corresponding Author: Ethan Dickey (dickeye at purdue.edu)

Original GLOW team, developed in Summer 2025: Ethan Dickey (+Rubric co-author), Ashok Saravanan, Alexander Siladie, Houyame Lkhider (+Rubric co-author), Quiondriya Gee, Andres Bejarano

Introduction

This section outlines the implementation model, design assumptions, and operational considerations for deploying GLOW within a GTA onboarding program. The intent is to provide sufficient transparency for academic stakeholders while offering concrete guidance for program leads responsible for logistics, evaluation alignment, and resource allocation.

GLOW is designed as a structured rehearsal system for instructional communication in STEM contexts. It is not a replacement for mentoring or observation, but a scalable complement that enables repeated practice, rubric-aligned feedback, and program-level analytics.

Implementation Model

Delivery Format

- Web-based platform (standalone link).

- Typically deployed during onboarding or pre-semester training.

- Participants complete a defined number of required simulations aligned to a program benchmark.

- Each simulation ends with rubric-anchored feedback at the domain level.

- Retakes are permitted to encourage iterative improvement.

Rubric Structure

- Core rubric domains are stable and discipline-agnostic (e.g., clarity, de-escalation, professionalism).

- Programs do not generally modify rubric criteria.

- Primary customization occurs through the scenario bank (course context, policies, disciplinary content).

This preserves cross-cohort comparability while allowing local relevance.

Program-Level Configuration

Program leads typically determine:

- Required number of simulations to "pass" onboarding

- Benchmark threshold (e.g., minimum domain scores)

- Persona mix (e.g., proportion of aggressive vs. confused cases)

- Scenario alignment to specific courses (e.g., CS182 vs. CS253)

- Whether retries are limited or unlimited

Dashboards provide aggregated metrics, including:

- Completion rates

- Attempt counts

- Domain-level performance distributions

- Persona-specific difficulty trends

- … and much more

These data support targeted coaching and resource allocation without replacing human evaluation.

Scope and Boundaries

GLOW is designed as structured rehearsal under controlled variation. It does not claim to:

- Replace in-person classroom observation

- Measure long-term instructional effectiveness

- Capture nonverbal communication or embodied classroom dynamics

Simulations are text-based and intentionally introduce randomized contextual variation (e.g., deadlines, room intensity, persona stance) to broaden exposure.

Performance analytics should therefore be interpreted as indicators of readiness under simulated conditions, not summative certification of teaching quality.

AI Architecture (Transparency Note)

GLOW uses a two-stage model architecture:

1. One model generates persona-consistent dialogue within a defined scenario.

2. A second model evaluates responses using rubric anchors to produce interpretable, domain-level feedback.

This separation is intended to reduce self-grading artifacts and increase consistency in feedback application.

Data and Privacy

- "Student" dialogue is AI-generated and not derived from real office-hour transcripts.

- Excerpts used for research or dissemination are anonymized and de-identified.

- Programs retain control over local deployment decisions and data access policies.

- AI model backend is easily swappable to local deployments or other models.

Cost and Resource Considerations

For planning purposes:

- API costs for a mid-sized onboarding cohort (e.g., ~250 GTAs required to pass ~6 simulations each) are typically modest (on the order of low hundreds of USD), depending on model selection and transcript length.

- Additional costs may include hosting, integration, and staff time for scenario development and dashboard review.

Programs should treat GLOW as a scalable practice layer that reduces live role-play time while preserving coaching capacity for higher-need or extreme cases.

Recommended Positioning Within GTA Training

GLOW is most effective when framed as:

- Structured rehearsal before live teaching

- A low-stakes environment for practicing difficult interactions

- A diagnostic tool to identify where mentoring time should be concentrated

- A complement to human-led workshops and observation

Programs that integrate simulation, human discussion, and follow-up reflection typically realize the strongest instructional alignment.