Application for Federal Assistance SF-424

* 1. Type of Submission: ☐ Preapplication ☒ Application ☐ Changed/Corrected Application	* 2. Type of Application: ☒ New ☐ Continuation ☐ Revision	* If Revision, select appropriate letter(s): * Other (Specify):
* 3. Date Received: Completed by Grants.gov upon submission.	4. Applicant Identifier:	
5a. Federal Entity Identifier:	5b. Federal Award Identifier:	

State Use Only:

6. Date Received by State:	7. State Application Identifier:

8. APPLICANT INFORMATION:

* a. Legal Name: The Johns Hopkins University

* b. Employer/Taxpayer Identification Number (EIN/TIN): 52-0595110	* c. UEI: FTMTDMBR29C7

d. Address:

* Street1:	3400 North Charles Street
Street2:	
* City:	Baltimore
County/Parish:	
* State:	MD: Maryland
Province:	
* Country:	USA: UNITED STATES
* Zip / Postal Code:	212182625

e. Organizational Unit:

Department Name: Ctr. for Safe & Healthy School	Division Name: School of Education

f. Name and contact information of person to be contacted on matters involving this application:

Prefix:	Dr.	* First Name:	Odis
Middle Name:			
* Last Name:	Johnson		
Suffix:			

Title: Professor

Organizational Affiliation: Johns Hopkins University

* Telephone Number:	410-516-4925	Fax Number:	

* Email: ojohns06@jhu.edu

Application for Federal Assistance SF-424

*** 9. Type of Applicant 1: Select Applicant Type:**

O: Private Institution of Higher Education

Type of Applicant 2: Select Applicant Type:

Type of Applicant 3: Select Applicant Type:

* Other (specify):

*** 10. Name of Federal Agency:**

Office of Postsecondary Education

11. Assistance Listing Number:

84.116

Assistance Listing Title:

Fund for the Improvement of Postsecondary Education

*** 12. Funding Opportunity Number:**

ED-GRANT-26-038

* Title:

FY 26 Postsecondary Student Success Grant 84.116M

13. Competition Identification Number:

ED-GRANT-26-038

Title:

FY 26 Postsecondary Student Success Grant 84.116M

14. Areas Affected by Project (Cities, Counties, States, etc.):

Add Attachment | Delete Attachment | View Attachment

*** 15. Descriptive Title of Applicant's Project:**

Helping Students Stay in STEM: A Personalized AI Companion for Gateway Courses

Attach supporting documents as specified in agency instructions.

Add Attachments | Delete Attachments | View Attachments

Application for Federal Assistance SF-424

16. Congressional Districts Of:

* a. Applicant: MD-007

* b. Program/Project: US-all

Attach an additional list of Program/Project Congressional Districts if needed.

[Add Attachment] [Delete Attachment] [View Attachment]

17. Proposed Project:

* a. Start Date: 10/01/2026

* b. End Date: 09/30/2030

18. Estimated Funding ($):

* a. Federal	3,999,325.00
* b. Applicant	300,000.00
* c. State	0.00
* d. Local	0.00
* e. Other	100,000.00
* f. Program Income	0.00
* g. TOTAL	4,399,325.00

*** 19. Is Application Subject to Review By State Under Executive Order 12372 Process?**

[X] a. This application was made available to the State under the Executive Order 12372 Process for review on 06/29/2026.

[ ] b. Program is subject to E.O. 12372 but has not been selected by the State for review.

[ ] c. Program is not covered by E.O. 12372.

*** 20. Is the Applicant Delinquent On Any Federal Debt? (If "Yes," provide explanation in attachment.)**

[ ] Yes [X] No

If "Yes", provide explanation and attach

[Add Attachment] [Delete Attachment] [View Attachment]

21. *By signing this application, I certify (1) to the statements contained in the list of certifications and (2) that the statements herein are true, complete and accurate to the best of my knowledge. I also provide the required assurances** and agree to comply with any resulting terms if I accept an award. I am aware that any false, fictitious, or fraudulent statements or claims may subject me to criminal, civil, or administrative penalties. (U.S. Code, Title 18, Section 1001)**

[X] ** I AGREE

** The list of certifications and assurances, or an internet site where you may obtain this list, is contained in the announcement or agency specific instructions.

Authorized Representative:

Prefix:

* First Name: Denise

Middle Name:

* Last Name: Sparks

Suffix:

* Title: Senior Grants Associate

* Telephone Number: 667-208-8806

Fax Number:

* Email: dsparks4@jhu.edu

* Signature of Authorized Representative: Completed by Grants.gov upon submission.

* Date Signed: Completed by Grants.gov upon submission.

BUDGET INFORMATION - Non-Construction Programs

OMB Number: 4040-0006
Expiration Date: 06/30/2028

SECTION A - BUDGET SUMMARY

Grant Program Function or Activity (a)	Assistance Listing Number (b)	Estimated Unobligated Funds		New or Revised Budget		
		Federal (c)	Non-Federal (d)	Federal (e)	Non-Federal (f)	Total (g)
1. FIPSE-PSSG	84.116M	$	$	$ 3,999,325.00	$ 400,000.00	$ 4,399,325.00
2.						
3.						
4.						
5. Totals		$	$	$ 3,999,325.00	$ 400,000.00	$ 4,399,325.00

SECTION B - BUDGET CATEGORIES

6. Object Class Categories	GRANT PROGRAM, FUNCTION OR ACTIVITY				Total
	(1) FIPSE-PSSG	(2) N/A	(3) N/A	(4) N/A	(5)
a. Personnel	$ 468,424.00	$	$	$	$ 468,424.00
b. Fringe Benefits	84,894.00				84,894.00
c. Travel	16,500.00				16,500.00
d. Equipment	0.00				0.00
e. Supplies	5,300.00				5,300.00
f. Contractual	335,632.00				335,632.00
g. Construction	0.00				0.00
h. Other	71,443.00				71,443.00
i. Total Direct Charges (sum of 6a-6h)	982,193.00				$ 982,193.00
j. Indirect Charges	57,065.00				$ 57,065.00
k. TOTALS (sum of 6i and 6j)	$ 1,039,258.00	$	$	$	$ 1,039,258.00
7. Program Income	$ 0.00	$	$	$	$ 0.00

Authorized for Local Reproduction

Standard Form 424A (Rev. 7- 97)
Prescribed by OMB (Circular A -102) Page 1A

SECTION C - NON-FEDERAL RESOURCES

(a) Grant Program	(b) Applicant	(c) State	(d) Other Sources	(e)TOTALS
8. FIPSE-PSSG	$ 300,000.00	$ 0.00	$ 100,000.00	$ 400,000.00
9.				
10.				
11.				
12. TOTAL (sum of lines 8-11)	$ 300,000.00	$ 0.00	$ 100,000.00	$ 400,000.00

SECTION D - FORECASTED CASH NEEDS

	Total for 1st Year	1st Quarter	2nd Quarter	3rd Quarter	4th Quarter
13. Federal	$ 1,039,258.00	$ 259,814.00	$ 259,814.00	$ 259,814.00	$ 259,816.00
14. Non-Federal	$ 103,926.00	25,000.00	25,000.00	25,000.00	28,926.00
15. TOTAL (sum of lines 13 and 14)	$ 1,143,184.00	$ 284,814.00	$ 284,814.00	$ 284,814.00	$ 288,742.00

SECTION E - BUDGET ESTIMATES OF FEDERAL FUNDS NEEDED FOR BALANCE OF THE PROJECT

(a) Grant Program	FUTURE FUNDING PERIODS (YEARS)			
	(b)First	(c) Second	(d) Third	(e) Fourth
16. FIPSE-PSSG	$ 1,103,999.00	$ 951,038.00	$ 905,030.00	$
17.				
18.				
19.				
20. TOTAL (sum of lines 16 - 19)	$ 1,103,999.00	$ 951,038.00	$ 905,030.00	$

SECTION F - OTHER BUDGET INFORMATION

21. Direct Charges: 3793912

22. Indirect Charges: 205413

23. Remarks: Total Direct Costs - $3,793,912; MTDC - $2,567,662, IDC - $205,413; Cost Sharing - $400,000; Total Project Costs - $4,399,325

Budget Narrative File(s)

*** Mandatory Budget Narrative Filename:** FIPSE JHU Budget Narr_June 29.pdf

Add Mandatory Budget Narrative | Delete Mandatory Budget Narrative | View Mandatory Budget Narrative

To add more Budget Narrative attachments, please use the attachment buttons below.

Add Optional Budget Narrative | Delete Optional Budget Narrative | View Optional Budget Narrative

Budget Narrative

Title: *Helping Students Stay in STEM:A Personalized AI Companion for Gateway Courses*

A. Personnel - $1,679,740

The project will be led by a multidisciplinary team of researchers from JHU's School of Education, Whiting School of Engineering and Krieger School of Arts and Sciences.

Key Personnel –

Dr. Odis Johnson will be the Chief Scientific Officer/Principal Investigator and will be responsible for the overall management of the project, ensuring that the design and implementation phases align with the project's goals and anticipated outcomes. He will lead a multidisciplinary team to develop adaptive learning and tutoring systems that enhance educational experiences for students in introductory STEM courses. His responsibilities include overseeing project design, coordinating with institutional and educational partners, and managing the execution of evaluations to assess the effectiveness of the implemented tools. Dr. Johnson will also ensure adherence to compliance with ethical standards relating to human subjects and data management, including the protection of student information according to FERPA regulations. Furthermore, he will serve as the primary communicator with stakeholders, providing updates and insights about the project's progress and successes, while also identifying opportunities for scaling effective practices. He will commit 12.5% effort in each year of the project.

Dr. James Diamond will serve as co-Principal Investigator. Dr. Diamond will manage the design-based research components of the project, including co-developing the theoretical framework, co-designing and iteratively refining the STEM-AI Companion features in Year 1, and working with the other investigators to ensure alignment between the intervention, targeted constructs, and outcome measures across partner institutions. Over Years 2–4, he will support the following activities: integration of the Companion into gateway STEM courses; interpretation of mixed-methods data on student–AI interaction and learning outcomes; and dissemination of design principles and implementation guidance for instructors and institutions. He will commit 10% effort in each year of the project.

Dr. Hao (Frank) Yang will serve as co-Principal Investigator. Dr. Yang will lead the technical design, development, and evaluation of the AI companion. He will supervise one doctoral student throughout the project and a postdoctoral fellow in Years 1–2, coordinating their work on model architecture, personalization policies, integration with the LASSO infrastructure, and analysis of student–AI interaction data. Under Dr. Yang's direction, the team will implement and refine the orchestration layer on top of foundation models, develop and test algorithms for trustworthy, individualized feedback, and contribute to quantitative analyses of learning gains and usage patterns across the three experimental arms. He will also co-lead dissemination efforts, including joint conference presentations to present project findings and engage with the broader AI-in-education and STEM education communities. He will commit 10% effort in each year of the project.

Dr. Tak Igusa will serve as co-Principal Investigator. Dr. Igusa will serve as co-advisor to Dr. Yang's Ph.D. student and postdoctoral researcher developing the AI Companion, providing guidance on aligning the system's technical capabilities with the project's scaffolding, feedback, and self-regulated learning framework. He brings expertise in engineering methods for implementing complex, data-driven interventions, supporting the integration and deployment of the AI Companion in authentic instructional settings. He will commit 8.3% effort in each year of the project.

Dr. Lingxin Hao will serve as co-Principal Investigator. Dr. Lingxin Hao will oversee and supervise the Sociology PhD student to perform tasks related to the experimental design, methods, and statistical/machine-learning analysis of the data collected. Dr. Hao also collaborates with others in the data collection effort and the development and refinement of the AI companion tool. She will commit 10% effort in each year of the project.

Dr. William R. Gray-Roncal will serve as co-Principal Investigator. Dr. Gray-Roncal is a Research Engineer with expertise in AI and human-machine teaming programs. He will support the design and development of the AI companion tool and internal review and analysis of the student engagement transcripts. He will commit 10% effort in each year of the program.

Other Personnel –

A part-time **Project Manager** will be hired to assist the team with coordination and logistics with the many partners involved in the project. This individual will commit 50% of their time to the project.

Two **Postdoctoral Fellows** will be hired, one in Education and one in Engineering, to assist project faculty with the management and analysis of data. Both will commit 100% of their time to the project in Years 1 & 2.

Three **Doctoral Students** will be hired, in Education, Engineering and Arts and Sciences, to assist with data collection. Each will work 20 hours per week for each year of the project.

Salaries for personnel are increased 3% each year according to JHU policy.

- **Year 1: $468,424**
- **Year 2: $482,477**
- **Year 3: $359,034**
- **Year 4: $369,805**

B. Fringe Benefits – $296,089

We have used a fringe benefit rate for Faculty and Staff of 31.5%, and 21.1% for postdoctoral fellows.

- **Year 1: $84,894**
- **Year 2: $87,440**
- **Year 3: $60,963**

- **Year 4: $62,792**

C. **Travel – $66,000**
We have budgeted professional conference travel in each year to disseminate project findings. Expenses are estimated at $600 for airfare, $200/night for 3 nights of lodging, $80 for 2.5 days of per diem, and $100 for ground transportation for a total of $1500 per person, per trip. We estimate eleven members of the project team would travel to conferences each year.
    - **Year 1: $16,500**
    - **Year 2: $16,500**
    - **Year 3: $16,500**
    - **Year 4: $16,500**

D. **Equipment – $0** *none requested*

E. **Supplies – $21,200**
We anticipate needing materials and supplies for the training sessions and for data collection and analysis.
    - **Year 1: $5,300**
    - **Year 2: $5,300**
    - **Year 3: $5,300**
    - **Year 4: $5,300**

F. **Contractual – $1,272,000**
Web Services: We will engage a consultant to provide expert advice and strategic guidance on maintaining our project website.
    - **Year 1: $6,000**
    - **Year 2: $6,000**
    - **Year 3: $6,000**
    - **Year 4: $6,000**

Advisory Board: Seven national experts in STEM education and artificial intelligence will advise the project team. Each will receive $1000 in each year of the project.
    - **Year 1: $7,000**
    - **Year 2: $7,000**
    - **Year 3: $7,000**
    - **Year 4: $7,000**

Quality Measures LLC/External Evaluator: Gwen Lee-Thomas will serve as the project evaluator to provide and evaluation design, to generate data, formative evaluation and final evaluation report.
    - **Year 1: $55,000**
    - **Year 2: $55,000**
    - **Year 3: $55,000**
    - **Year 4: $55,000**

Iowa State University/Dr. Ben Van Dusen - Dr. Van Dusen will lead the design and implementation of the AI-supported instructional infrastructure, supervise research and evaluation activities, manage project personnel and partnerships, and ensure that assessment, dashboard, and AI Companion components function coherently to support student learning across participating STEM courses. Below is a detailed narrative of subgrantee expenses.

- **Year 1: $267,632**
- **Year 2: $258,833**
- **Year 3: $266,989**
- **Year 4: $206,546**

Iowa State Subaward Justification
Direct Project Costs: $1,000,000 plus $100,000 Cost Sharing

A. SENIOR PERSONNEL

A.1. Van Dusen, Ben, Principal Investigator

Funds are requested for three summer months per year for Dr. Ben Van Dusen, Associate Professor in the School of Education at Iowa State University and Director of the LASSO platform. Dr. Van Dusen will provide leadership for the Iowa State University portion of the project and will oversee the development, implementation, and evaluation of the project's assessment and AI-support infrastructure.

Dr. Van Dusen will lead the development of cognitive diagnostic assessments, oversee the integration of assessment and AI tutoring systems within the LASSO platform, supervise project personnel, coordinate activities with project partners, support the implementation of pilot and experimental studies, contribute to data analysis and dissemination, and ensure the successful completion of project milestones.

Salary requested:

Year	Salary
Year 1	$35,974
Year 2	$37,053
Year 3	$38,165
Year 4	$39,310
Total	$150,502

Salary increases are projected at approximately 3% annually in accordance with university policy.

B. OTHER PERSONNEL

B.1. Postdoctoral Research Associate

Funds are requested for one full-time Postdoctoral Research Associate for the duration of the project. The postdoctoral researcher will play a central role in the design, development, implementation, and evaluation of project activities.

Responsibilities include developing and validating cognitive diagnostic assessments; creating and calibrating assessment items; supporting development of the AI Companion; integrating assessment outputs with AI-supported learning tools; connecting assessment and tutoring infrastructure within the LASSO platform; conducting quantitative and qualitative research analyses; supporting pilot studies and randomized trials; preparing technical documentation; and assisting with dissemination activities.

Salary requested:

Year	Salary
Year 1	$63,631
Year 2	$65,561
Year 3	$67,527
Year 4	$69,553
Total	$266,292

Salary increases are projected at approximately 3% annually.

B.2. Graduate Research Assistant

Funds are requested to support one half-time Graduate Research Assistant during Years 1–3 of the project.

The Graduate Research Assistant will support assessment development, data collection, management of research databases, quantitative and qualitative analyses, participant recruitment, user testing, literature reviews, instructor support activities, and dissemination efforts. The GRA will also assist with the implementation of the AI Companion, coding and analysis of student interactions, preparation of reports and manuscripts, and support for project evaluation activities.

Salary requested:

Year	Salary
Year 1	$30.193
Year 2	$31,099

Year	Salary
Year 3	$32,032
Total	$93,324

B.3. Professional and Scientific Staff (Project Manager)

Funds are requested for two months of effort per year for a Professional and Scientific staff member serving as Project Manager.

The Project Manager will coordinate project operations across institutions, facilitate communication among project partners, monitor project timelines and milestones, organize meetings and reporting activities, assist with participant recruitment and onboarding, maintain project records, and support dissemination and sustainability activities.

Salary requested:

Year	Salary
Year 1	$15,341
Year 2	$15,801
Year 3	$16,275
Year 4	$16,764
Total	$64,181

C. FRINGE BENEFITS

Fringe benefits are specifically identified for each employee and charged as direct costs in accordance with Iowa State University policy. Fringe benefit rates are applied according to employee classification and university-approved rates.

Applicable fringe rates include:

Faculty: 31.8%

Postdoctoral Research Associate: University-approved postdoctoral rate

Professional & Scientific Staff: University-approved P&S rate

Graduate Research Assistant: 15.2%

Total fringe benefits requested: $185,767

D. EQUIPMENT

No equipment is requested.

E. TRAVEL

Funds are requested to support the dissemination of project findings at national conferences and project coordination activities.

E.1. Conference Dissemination

Travel funds are requested to support project personnel in presenting findings at the American Educational Research Association (AERA) Annual Meeting and related national conferences focused on STEM education, educational technology, learning sciences, and assessment.

Travel costs include airfare, lodging, meals, registration fees, ground transportation, and related travel expenses consistent with university travel policies and federal guidelines.

Year	**Travel**
Year 1	$10,000
Year 2	$7,500
Year 3	$7,500
Year 4	$5,000
Total	$30,000

Travel will support the dissemination of project findings, coordination with collaborators, and engagement with national communities of practice.

F. PARTICIPANT SUPPORT COSTS

None requested.

G. OTHER DIRECT COSTS

G.1. Materials and Supplies

Funds support essential project supplies (e.g., printing, data collection, and participant materials), totaling $10,805 ($3,000 in Years 1–3 and $1,805 in Year 4), reflecting reduced needs during closeout.

G.6. Tuition Remission

Tuition remission is requested for the Graduate Research Assistant in accordance with Iowa State University policy. Tuition costs are calculated using current university rates and projected increases during the project period.

Total tuition remission requested: $50,820

G.9. Other

G.9.1. Developers

Funds support developer services necessary for project implementation and technical development, totaling $67,000 ($27,000 in Year 1; $15,000 in Years 2–3; $10,000 in Year 4), reflecting higher initial development needs and reduced support during maintenance and closeout.

G.9.2. AI Tokens

Funds support AI token usage required for project-related data processing and technical operations, totaling $11,000 ($2,000 in Years 1–2; $3,000 in Year 3; $4,000 in Year 4), reflecting increasing usage as project activities expand.

H. TOTAL DIRECT COSTS

Total Direct Costs: $929,690

I. FACILITIES AND ADMINISTRATIVE COSTS

Facilities and Administrative (F&A) costs are calculated in accordance with Iowa State University's federally negotiated indirect cost rate agreement. In compliance with sponsor guidelines, F&A is limited to 8% of Modified Total Direct Costs (MTDC) and is applied to the appropriate MTDC base, resulting in total F&A costs of $70,310.

J. COST SHARE

Iowa State University (ISU) will provide a cost share of $100,000 in support of the project, representing 10% of the total requested funds. All cost-shared contributions will be tracked and documented in accordance with sponsor and institutional guidelines.

K. TOTAL PROJECT COSTS

Total project costs (including direct and indirect costs) are $1,100,000, of which $1,000,000 is requested from the sponsor, as detailed in the accompanying budget forms.

G. Construction – $0 *none requested*

H. Other – $458,883

Institutional Research: We will compensate Institutional Research departments at partner institutions to pull data for us.

- **Year 1: $15,000**

- **Year 2: $15,000**
- **Year 3: $15,000**
- **Year 4: $15,000**

Research Incentives: We anticipate needing funds for incentives to encourage students and instructors to participate in the project. We will provide $15 incentives for 500 participants in Year 1, and 1000 in Years 2, 3, and 4.

- **Year 1: $7,500**
- **Year 2: $15,000**
- **Year 3: $15,000**
- **Year 4: $15,000**

Single IRB: Single IRB – Funds are budgeted for an external IRB to serve as the single IRB of record for this multi-site project. Funds are estimated at $5000 in each year of the project.

- **Year 1: $5,000**
- **Year 2: $5,000**
- **Year 3: $5,000**
- **Year 4: $5,000**

Doctoral Student Tuition: Funds are budgeted to cover a portion (20%) of the Whiting School of Engineering doctoral student tuition and health insurance. This supports their participation as a Graduate Research Assistant on this project. The remaining portion of their tuition will be coved as cost sharing.

- **Year 1: $17,693**
- **Year 2: $18,400**
- **Year 3: $19,137**
- **Year 4: $19,903**

Training Stipends: We request funds to compensate course instructors to attend training to implement the STEM-AI Companion in their classes. We plan to provide $3750 to seven instructors in Year 1, and to 20 instructors in Years 2-5. These costs are deducted from MTDC

- **Year 1: $26,250**
- **Year 2: $75,000**
- **Year 3: $75,000**
- **Year 4: $75,000**

I. **Total Direct Costs: $3,793,912**
    - **Year 1: $982,193**
    - **Year 2: $1,046,950**
    - **Year 3: $905,923**
    - **Year 4: $858,846**

J. **Indirect Costs: $205,413-** *Indirect Costs are being charged at the allowable rate of 8% of Modified Total Direct Costs (MTDC).*

- **Year 1: $57,065**
- **Year 2: $57,049**
- **Year 3: $45,115**
- **Year 4: $46,184**

K. 12. Total Costs: $3,999,325

- **Year 1: $1,039,258**
- **Year 2: $1,103,999**
- **Year 3: $951,038**
- **Year 4: $905,030**

COST SHARING Committed by Applicant: $300,000

Johns Hopkins University (JHU) will provide a cost share of $300,000 in support of the project, representing 10% of the total requested funds. All cost-shared contributions will be tracked and documented in accordance with sponsor and institutional guidelines.

Note: Total cost share for the project is $400,000. $300,000 from lead applicant and $100,000 from subaward institution, Iowa State University.

Abstract

An abstract is to be submitted in accordance with the following:

1. Abstract Requirements

- Abstracts must not exceed one page and should use language that will be understood by a range of audiences.
- Abstracts must include the project title, goals, and expected outcomes and contributions related to research, policy, and practice.
- Abstracts must include the population(s) to be served.
- Abstracts must include primary activities to be performed by the recipient.
- Abstracts must include subrecipient activities that are known or specified at the time of application submission.

For research applications, abstracts also include the following:

- Theoretical and conceptual background of the study (i.e., prior research that the investigation builds upon and that provides a compelling rationale for this study).
- Research issues, hypotheses and questions being addressed.
- Study design including a brief description of the sample including sample size, methods, principals, and dependent, independent, and control variables, as well as the approach to data analysis.

[Note: For a non-electronic submission, include the name and address of your organization and the name, phone number and e-mail address of the contact person for this project.]

You may now Close the Form

You have attached 1 file to this page, no more files may be added. To add a different file, you must first delete the existing file.

* Attachment: FIPSE Abstract.pdf

Add Attachment | Delete Attachment | View Attachment

Abstract

This Postsecondary Student Success Grant (PSSG) proposal, ***Helping Students Stay in STEM: A Personalized AI Companion for Gateway Courses*** pursues absolute priority (AP1) early phase research with a project priority focus (AP3) in artificial intelligence (AI). A cross-disciplinary research team of engineers, computer scientists, sociologists, and education researchers from Johns Hopkins University and Iowa State University will generate new knowledge about the impact of AI-assisted learning on course non-completion for undergraduate students in physics, calculus, and chemistry. With the science, technology, engineering, and mathematics (STEM) workforce projected to grow by 8.1 percent through 2034, outpacing growth in non-STEM occupations, jobs may be left unfilled if we do not reverse our current status of granting bachelor's degrees in STEM to fewer than half of first-year undergraduate students who start a STEM program. With this challenge in mind, we test whether the STEM-AI Companion we have developed will serve as an AI co-regulator of students' self-regulated learning (SRL) using its hybrid intelligence and scaffolding adaptivity to achieve personalization of the student learning experience. Using a 3-arm randomization design aligned with WWC standards, this project will rigorously compare three models of support in 16 universities and community colleges for over 4500 Pell eligible, transfer-in, and first-generation students who: 1) utilize the LASSO assessments platform, 2) use LASSO with a non-personalized AI companion of their choice (e.g. ChatGPT, Claude, Copilot, Grok, or Meta AI), and 3) use LASSO and the personalized STEM-AI Companion. These experiences are related to course completion, performance, course repetition, and change of major. This project also advances the competitive preference priority, being undertaken by a land-grant institution within an EPSCoR (Established Program to Stimulate Competitive Research) state, Iowa State University in Ames, Iowa (OPEID 001869).

U.S. Department of Education Supplemental Information for the SF-424
Application for Federal Assistance

1. Project Director and Applicable Entity Identification Numbers:

Prefix:	* First Name:	Middle Name:	* Last Name:	Suffix:
Dr.	Odis		Johnson	

* Project Director Level of Effort (percentage of time devoted to grant): 13

Address:

* Street1: 2800 North Charles Street

Street2:

* City: Baltimore

County:

* State: MD: Maryland
* Zip Code: 212184026
* Country: USA: UNITED STATES

* Phone Number (give area code): 410-516-4925

Fax Number (give area code):

* Email Address: ojohns06@jhu.edu

Alternate Email Address:

OPE ID(s) (if applicable): 002077

NCES School ID(s) (if applicable):

NCES LEA/School District ID(s) (if applicable):

2. General Education Provision Act Section 427 Assurance:

☒ I assure that the proposed project complies with the requirements in section 427 of the General Education Provisions Act **(20 U.S.C. 1228a)**. Compliance can be found on the following page(s) of the application:

p. 3-4

3. New Potential Grantee:

☒ N/A. This item is not applicable because the program competition’s notice inviting applications (NIA) does not include a definition “New Potential Grantee.” This item is not applicable when the program competition’s NIA does not include the definition.

For NIA’s that include a definition of “New Potential Grantee,” complete the following:

Are you a new potential grantee as defined in the program competition’s NIA?

☐ Yes ☐ No

4. Human Subjects Research:

a. Are any research activities involving human subjects planned at any time during the proposed Project Period?

☒ Yes ☐ No

b. Are ALL the research activities proposed designated to be exempt from the regulations?

☐ Yes Provide Exemption(s) #(s): ☐ 1 ☐ 2 ☐ 3 ☐ 4 ☐ 5 ☐ 6 ☐ 7 ☐ 8

☒ No Provide Federal Wide Assurance #(s), if available:

FWA00005834

c. If applicable, please attach your "Exempt Research" or "Nonexempt Research" narrative to this form as indicated in the definitions page in the attached instructions.

Final Non exempt human subjects narrative_June 2 | Add Attachment | Delete Attachment | View Attachment

5. Infrastructure Programs and Build America, Buy America Act Applicability:

If the competition Notice Inviting Applications (NIA) in section III. 4. “Other” states that the program under which this application is submitted is subject to the Build America, Buy America Act (Pub. L. 117-58) (BABAA) domestic sourcing requirements, complete the following:

☒ This application does not include any infrastructure projects or activities and therefore **IS NOT** subject the BABAA domestic sourcing requirements.

☐ This application **IS** subject to the BABAA domestic sourcing requirements, because the proposed grant project described in this application includes the following infrastructure projects or activities:

☐ Construction

☐ Remodeling

☐ Broadband Infrastructure

If this application **IS** subject to the BABAA domestic sourcing requirements, please list the page numbers from within the application narrative where the proposed infrastructure project or activities are described:

Non-Exempt Human Subjects Research Narrative

Helping Students Stay in STEM: A Personalized AI Companion for Gateway Courses

1. **Human Subjects Involvement and Characteristics:**
   The proposed research involves human subjects who will participate in using a personalized artificial intelligence (AI) learning companion, designed for providing tutoring and analytics. The anticipated subject population includes undergraduate students enrolled in gateway STEM courses. It is expected that the participant number will be substantial, potentially ranging from hundreds to thousands, particularly in diverse educational settings.
   Age range: Likely to include undergraduate students, typically aged 18-24.
   Health Status: Participants are anticipated to be generally healthy; however, accommodations may be needed for students with disabilities.
   In addition to platform-based participation, a subset of students will be invited to participate in optional semi-structured interviews about their experiences using the AI learning companion. These interviews will be conducted only after IRB approval and will involve minimal risk.
   Inclusion criteria: Students enrolled in the specified courses. Exclusion criteria: students who opt out or do not meet course enrollment requirements. Special considerations for vulnerable populations (e.g., students with disabilities) will ensure that these groups are appropriately supported throughout the research.

2. **Sources of Materials:**
   Research materials will include data collected from the LASSO platform, specifically interaction logs and analytics generated from student usage. Data used will be primarily derived from existing records, though new data specific to the research objectives will also be collected. All data will be de-identified in accordance with ethical standards and privacy laws. We will also collect qualitative data through focus groups with participants to get a better understanding of their experience with the AI Companion.

   To supplement quantitative data, qualitative materials will also be collected through optional semi-structured interviews with a small subset of users. Interview data will be audio-recorded or transcribed only with participant consent, and all resulting qualitative information will be anonymized during transcription. No identifying information will be retained in any dataset used for analysis.

3. **Recruitment and Informed Consent:**
   Recruitment will involve outreach through course announcements, institutional emails, and professional organizations. Consent will be obtained prior to data collection, ensuring

that the nature of the study and data utilization are clearly explained to participants. Trained research staff will seek consent, and participants will receive a consent form detailing the study's purpose, risks, and benefits. Documentation of consent can be physical or electronic as determined by the IRB.

Recruitment for interviews will occur separately and only after IRB approval for the qualitative component. Students will receive an invitation describing the interview purpose, time commitment, and privacy protections. Consent for interviews will explicitly include permission to audio-record when applicable. Interview participation will not affect course standing or access to instructional tools, and non-participation will involve no penalty.

At this time, no waivers or modifications to the consent process are being requested; however, any future modifications will be submitted to and approved by the IRB before implementation.

4. **Potential Risks:**
   Potential risks include privacy concerns related to the data collected and possible psychological stress from engagement with tutoring assessments. The likelihood of serious risks is minimal, primarily due to de-identification protocols that enhance participant privacy. Alternative treatments may include less intensive forms of tutoring or support services, which could be beneficial for students needing additional help.

   Interview participation may involve mild reflection on personal learning experiences, which could cause slight discomfort; however, the likelihood of significant distress is low, and participants may decline to answer any question. No physical, legal, or economic risks are anticipated. Students who prefer not to discuss their learning processes may simply opt out of the interview portion.

5. **Protection Against Risk:**
   To protect against these risks, the project will implement robust data security measures, including encryption and limited access to identifiable data. All data protocols will follow FERPA guidelines to ensure confidentiality. Regular reviews and monitoring for any adverse effects on participants will also be in place, along with established procedures for professional intervention if required.

   Interview data will be stored securely in access-controlled locations. Audio recordings will be deleted immediately after transcription. Transcripts will be fully de-identified before analysis, and no direct identifiers will appear in published or shared materials. Participants will be reminded that they may skip any question or withdraw at any point.

IRB-approved procedures will guide all interactions with human subjects. Any unanticipated adverse events will be reported promptly to the IRB. Ongoing monitoring will ensure that data collection procedures remain minimal risk throughout the project.

6. **Importance of the Knowledge to be Gained:**
   The knowledge gained from this research will significantly contribute to understanding student engagement and learning processes in STEM education. Interview data will provide critical insight into students' perceptions of transparency, usability, workload, and trust in AI systems, complementing the quantitative engagement and learning outcomes data collected from the platform. The risks involved are reasonable compared to the anticipated benefits of enhanced teaching methods and student support systems. The project aims to enhance educational outcomes and develop tools for effective learning interventions.

7. **Collaborating Site(s):**
   The research will be led by Dr. Odis Johnson at Johns Hopkins University and potentially other partner institutions involved with the LASSO platform. Each site will participate by providing access to student subject pools, supporting recruitment efforts, and contributing resources for data collection and analysis. Collaborating sites will also assist with identifying potential interview participants and supporting logistics for qualitative data collection. The role of collaborating sites will be crucial for ensuring diverse participation and comprehensive data collection across different educational environments.

OMB Number: 4040-0013
Expiration Date: 06/30/2028

CERTIFICATION REGARDING LOBBYING

Certification for Contracts, Grants, Loans, and Cooperative Agreements

The undersigned certifies, to the best of his or her knowledge and belief, that:

(1) No Federal appropriated funds have been paid or will be paid, by or on behalf of the undersigned, to any person for influencing or attempting to influence an officer or employee of an agency, a Member of Congress, an officer or employee of Congress, or an employee of a Member of Congress in connection with the awarding of any Federal contract, the making of any Federal grant, the making of any Federal loan, the entering into of any cooperative agreement, and the extension, continuation, renewal, amendment, or modification of any Federal contract, grant, loan, or cooperative agreement.

(2) If any funds other than Federal appropriated funds have been paid or will be paid to any person for influencing or attempting to influence an officer or employee of any agency, a Member of Congress, an officer or employee of Congress, or an employee of a Member of Congress in connection with this Federal contract, grant, loan, or cooperative agreement, the undersigned shall complete and submit Standard Form-LLL, "Disclosure of Lobbying Activities," in accordance with its instructions.

(3) The undersigned shall require that the language of this certification be included in the award documents for all subawards at all tiers (including subcontracts, subgrants, and contracts under grants, loans, and cooperative agreements) and that all subrecipients shall certify and disclose accordingly. This certification is a material representation of fact upon which reliance was placed when this transaction was made or entered into. Submission of this certification is a prerequisite for making or entering into this transaction imposed by section 1352, title 31, U.S. Code. Any person who fails to file the required certification shall be subject to a civil penalty of not less than $10,000 and not more than $100,000 for each such failure.

Statement for Loan Guarantees and Loan Insurance

The undersigned states, to the best of his or her knowledge and belief, that:

If any funds have been paid or will be paid to any person for influencing or attempting to influence an officer or employee of any agency, a Member of Congress, an officer or employee of Congress, or an employee of a Member of Congress in connection with this commitment providing for the United States to insure or guarantee a loan, the undersigned shall complete and submit Standard Form-LLL, "Disclosure of Lobbying Activities," in accordance with its instructions. Submission of this statement is a prerequisite for making or entering into this transaction imposed by section 1352, title 31, U.S. Code. Any person who fails to file the required statement shall be subject to a civil penalty of not less than $10,000 and not more than $100,000 for each such failure.

* APPLICANT'S ORGANIZATION: The Johns Hopkins University

* PRINTED NAME AND TITLE OF AUTHORIZED REPRESENTATIVE

Prefix:
* First Name: Denise
Middle Name:
* Last Name: Sparks
Suffix:
* Title: Senior Grants Associate

* SIGNATURE: Completed on submission to Grants.gov
* DATE: Completed on submission to Grants.gov

Other Attachment File(s)

*** Mandatory Other Attachment Filename:** Priority Statement FIPSE 2026.pdf

Add Mandatory Other Attachment | Delete Mandatory Other Attachment | View Mandatory Other Attachment

To add more "Other Attachment" attachments, please use the attachment buttons below.

Add Optional Other Attachment | Delete Optional Other Attachment | View Optional Other Attachment

ABSOLUTE PRIORITY STATEMENT (AP1)/ (AP3) ARTIFICIAL INTELLIGENCE

This Postsecondary Student Success Grant (PSSG) proposal, Helping Students Stay in STEM: A Personalized AI Companion for Gateway Courses (hereafter, “AI Companion Project”) advances Evidence Absolute Priority 1 (AP1) in early phase research with a Project Pathway Absolute Priority 3 (AP3) in artificial intelligence (AI) by testing whether AI-assisted learning will reduce course non-completion for students in physics, calculus, and chemistry.

As growth in science, technology, engineering, and mathematics (STEM) jobs outpace expected growth in non-STEM jobs into the next decade, jobs may be left unfilled if we do not reverse our current course of granting bachelor’s degrees in STEM to fewer than half of first-year undergraduate students who start a STEM program. With this challenge in mind, we test whether the STEM-AI Companion we have developed will serve as an AI co-regulator of students’ self-regulated learning (SRL) using its hybrid intelligence and scaffolding adaptivity to achieve personalization of the student learning experience. Using a 3-arm randomization design aligned with WWC standards, this project will rigorously compare three models of support in 16 universities and community colleges for over 4500 Pell eligible, transfer-in, and first-generation students who: 1) utilize the LASSO assessments platform, 2) use LASSO with a non-personalized AI companion of their choice (e.g. ChatGPT, Claude, Copilot, Grok, or Meta AI), and 3) use LASSO and the personalized STEM-AI Companion. These experiences are related to course completion, performance, course repetition, and change of major.

This project aligns with AP1 by establishing a logic model and preliminary evidence of the potential impact of AI-assisted technologies on the learning outcomes of undergraduates in chemistry, calculus, and physics. Consistent with early phase research with the capacity to contribute to the WWC, this project features an early-phase (Absolute Priority 1) design, the stratified block randomization causally tests the short-term outcomes (productive AI use, self-direction, conceptual mastery, gateway-course completion, credit accumulation, and continuation to the next course) with a focus on Pell students and part-time or transfer-in students. The project tracks and reports the long-term outcomes, disaggregated by Pell status or other non-full-time statuses with baselines and targets. Improved economic mobility is identified as the ultimate, longer-horizon impact, rather than a measured project outcome. This proposal's conjecture mapping, synthesis of evidence-based WWC studies (personalized support, high-impact tutoring, and responsible, human-centered use of AI), and logic model maximize the feasibility of this study to link AI-assisted technologies to the improved outcomes of underserved populations.

The STEM-AI Companion project offers a transformative solution to these challenges by embedding a personalized, course-integrated AI tutor (Mousavinasab et al., 2021; Van Lehn, 2011; Ritter et al., 2007) directly within an established assessment platform—LASSO (LASSO, 2025; Nissen et al., 2018; Van Dusen, 2018; Van Dusen et al., 2021). Unlike traditional tutoring or generic AI tools, the STEM-AI Companion is designed to scaffold reasoning, guide solution planning, and prompt reflection, all while integrating domain-specific prompt optimization, retrieval from curated instructional resources, and adaptive reasoning strategies to personalize support based on students' demonstrated mastery (Zhao et al,, 2025). These capabilities are intended to strengthen the instructional mechanisms that prior research has shown to improve learning—including timely formative feedback, personalized scaffolding, and adaptive tutoring (Black & Wiliam, 1998; Hattie & Timperley, 2007; Ma et al., 2014; Nickow et al., 2020). Thus, the proposed intervention builds on established evidence regarding effective instructional practices while incorporating recent advances in trustworthy and adaptive AI, providing a clear rationale for expecting the personalized AI Companion to improve student learning, self-regulated learning (Zimmerman, 2002; Schunk & Zimmerman, 2013), and gateway-course success beyond what can be achieved using LASSO alone or LASSO paired with a general-purpose AI tool**.**

The mechanism proceeds in four linked steps. *First*, because the Companion offers personalized hints rather than standard answers, students learn to question, check, and reason with AI rather than copy from it, building AI literacy. *Second*, prompts to plan, self-check, and reflect strengthen self-directed learning. *Third*, personalized scaffolding that fades over time deepens conceptual understanding and problem-solving skills. *Fourth*, these gains raise the share of students who pass gateway courses and accumulate first-year credits (the earliest indicators of retention and completion), which institutional records then track through to retention, transfer, and completion, disaggregated by Pell status and inclusive of part-time and transfer-in students. Both the AI Companion and LASSO were developed by co-PIs of this project using federal funding streams, and LASSO is free to users online to support broad use, scalability, and sustainability.

The following attachment is not included in the view since it is not a read-only PDF file.

Upon submission, this file will be transmitted to the Grantor without any data loss.

FY26_PSSG_Project_Profile_Form_Johnson2.pdf

Department of the Treasury
Internal Revenue Service
Tax Exempt and Government Entities
PO Box 2508
Cincinnati, OH 45201

Date:
May 20, 2026
Employer ID number:
52-0595110
Form 990 required:
Yes
Person to contact:
Name: S. BEDFORD
ID number: 1000195682

JOHNS HOPKINS UNIVERSITY
3910 KESWICK ROAD STE N 4327B
BALTIMORE, MD 21211

Dear Sir or Madam:

We're responding to your request dated April 15, 2026, about your tax-exempt status.

We issued you a determination letter in January 1935, recognizing you as tax-exempt under Internal Revenue Code (IRC) Section 501(c)(3).

Donors can deduct contributions they make to you as provided in IRC Section 170. You're also qualified to receive tax-deductible bequests, legacies, devises, transfers, or gifts under IRC Sections 2055, 2106, and 2522.

In the heading, we indicated whether you must file an annual information return. If you're required to file a return, you must file one of the following by the 15th day of the 5th month after the end of your annual accounting period.

- Form 990, Return of Organization Exempt From Income Tax
- Form 990-EZ, Short Form Return of Organization Exempt From Income Tax
- Form 990-N, Electronic Notice (e-Postcard) for Tax-Exempt Organizations Not Required to File Form 990 or Form 990EZ
- Form 990-PF, Return of Private Foundation or Section 4947(a)(1) Trust Treated as Private Foundation

According to IRC Section 6033(j), if you don't file a required annual information return or notice for 3 consecutive years, we'll revoke your tax-exempt status on the due date of the 3rd required return or notice.

You can get IRS forms or publications you need from our website at www.irs.gov/forms-pubs or by calling 800-TAX-FORM (800-829-3676).

If you have questions, call 877-829-5500 between 8 a.m. and 5 p.m., local time, Monday through Friday (Alaska and Hawaii follow Pacific time).

Thank you for your cooperation.

Sincerely,

Stephen A. Martin

Stephen A. Martin
Director, Exempt Organizations
Rulings and Agreements

CERTIFICATION REGARDING LOBBYING

Certification for Contracts, Grants, Loans, and Cooperative Agreements

The undersigned certifies, to the best of his or her knowledge and belief, that:

(1) No Federal appropriated funds have been paid or will be paid, by or on behalf of the undersigned, to any person for influencing or attempting to influence an officer or employee of an agency, a Member of Congress, an officer or employee of Congress, or an employee of a Member of Congress in connection with the awarding of any Federal contract, the making of any Federal grant, the making of any Federal loan, the entering into of any cooperative agreement, and the extension, continuation, renewal, amendment, or modification of any Federal contract, grant, loan, or cooperative agreement.

(2) If any funds other than Federal appropriated funds have been paid or will be paid to any person for influencing or attempting to influence an officer or employee of any agency, a Member of Congress, an officer or employee of Congress, or an employee of a Member of Congress in connection with this Federal contract, grant, loan, or cooperative agreement, the undersigned shall complete and submit Standard Form-LLL, "Disclosure of Lobbying Activities," in accordance with its instructions.

(3) The undersigned shall require that the language of this certification be included in the award documents for all subawards at all tiers (including subcontracts, subgrants, and contracts under grants, loans, and cooperative agreements) and that all subrecipients shall certify and disclose accordingly. This certification is a material representation of fact upon which reliance was placed when this transaction was made or entered into. Submission of this certification is a prerequisite for making or entering into this transaction imposed by section 1352, title 31, U.S. Code. Any person who fails to file the required certification shall be subject to a civil penalty of not less than $10,000 and not more than $100,000 for each such failure.

Statement for Loan Guarantees and Loan Insurance

The undersigned states, to the best of his or her knowledge and belief, that:

If any funds have been paid or will be paid to any person for influencing or attempting to influence an officer or employee of any agency, a Member of Congress, an officer or employee of Congress, or an employee of a Member of Congress in connection with this commitment providing for the United States to insure or guarantee a loan, the undersigned shall complete and submit Standard Form-LLL, "Disclosure of Lobbying Activities," in accordance with its instructions. Submission of this statement is a prerequisite for making or entering into this transaction imposed by section 1352, title 31, U.S. Code. Any person who fails to file the required statement shall be subject to a civil penalty of not less than $10,000 and not more than $100,000 for each such failure.

* APPLICANT'S ORGANIZATION

Iowa State University of Science and Technology

* PRINTED NAME AND TITLE OF AUTHORIZED REPRESENTATIVE

Prefix: * First Name: Sara   Middle Name:

* Last Name: Fonseca Ricke   Suffix:

* Title: Senior Pre-Award Administrator

* SIGNATURE: Sara Fonseca Ricke — Digitally signed by Sara Fonseca Ricke Date: 2026.06.26 13:56:59 -05'00'   * DATE: 06/26/2026

ODIS JOHNSON JR., PhD

Center for Safe and Healthy Schools
Johns Hopkins University
2800 North Charles Street, Office 311
Baltimore, MD 21218
Email: ojohns06@jhu.edu

ETS Research Institute
Educational Testing Service (ETS)
660 Rosedale Rd
Princeton, NJ 08540
Email: ojohnson@ets.org

A. TRAINING

2004	*Spencer Postdoctoral Research Fellow*, University of Chicago, Sociology/Consortium on Chicago School Research. Advisors: Anthony Bryk and Edgar Epps
2003	Ph.D., Education and Social Policy, University of Michigan, Ann Arbor. Dissertation Committee: David K. Cohen, Sheldon H. Danziger, and Mary E. Corcoran
1995	B.A., Music and Education, University of Tulsa, OK

B. ACADEMIC APPOINTMENTS

08/24 – Present	*Edmund W. Gordon Chair of Policy Research and Evaluation*, ETS Research Institute, Educational Testing Service, Princeton, NJ.
01/21 – Present	*Bloomberg Distinguished Professor*, School of Education, Bloomberg School of Public Health Department of Health Policy and Management, and the Department of Sociology, Johns Hopkins University
01/23 – Present	*Core Faculty*, Data Science and Artificial Intelligence Institute, Johns Hopkins
01/*21* – Present	*Inaugural Executive Director,* Center for Safe and Healthy Schools, Johns Hopkins
08/19 – Present	*Founding Director,* Institute in Critical Quantitative and Computational Methodologies (ICQCM), Johns Hopkins University

2. RESEARCH AND SCHOLARLY ACTIVITIES

A. SUBFIELDS

Quantitative and Computational Methods (*Data Science, AI, Critical Quantification and Computation*)
Sociology of Education (*School-to-Prison Pipeline, STEM, AI Surveillance, Seasonal Learning*)
Social Policy (*Housing Policy, Education Policy, Policing*)
Urban and Community Sociology (*Neighborhoods, Metropolitan Organization*)

B. HONORS, AWARDS, and FELLOWSHIPS

Marsal Distinguished Education Researcher Alumni Award, The University of Michigan, 2025
National Science Foundation Advisory Committee, Directorate for STEM Education (EDU), 2024-25
National Academy of Education Advisory Committee, EMERG, 2023
AIR Scholar, American Institutes for Research, June, 2023 - present
National Science Foundation, Scholar in Residence, December, 2022
U.S. Department of Education/IES, HS&B:22 Technical Review Panel, 2021-25
U.S. Department of Housing and Urban Development, HCV Mobility Demonstration Evaluation Expert Panel, 2021-25
Panel of Visitors, Educational Testing Services, 2021-24
Outstanding Mentor Award, Graduate Student Senate, Washington University in St. Louis, 2019
Outstanding Author Contribution Award, Emerald Publishing Group, 2015
Outstanding Review of Research Award, American Educational Research Association, 2013

Spencer Postdoctoral Research Fellow, University of Chicago, 2002 - 2004
Ford Foundation Dissertation Fellow, National Research Council, Washington, DC, 2001

C. GRANTS (Partial List** - 23 awarded, 12 as Principal/Sole Investigator**,** range $15K - $5 million**)

07/25 – 06/26 Spencer Foundation – "ICQCM: Aligning Transformative Methodological Development with Science Policy Realities." PI: **Odis Johnson Jr.**, co-PI Ebony McGee. $15,000

11/24 – 10/26 William T. Grant Foundation (CBC-205399) – "Advanced Critical Data Science Methods to Build Capacity in Youth Research." Co-PI Ebony McGee, Co-PI: Odis Johnson Jr. $200,000

10/24 – 09/25 Spencer Foundation (202500083) – "Critical Data Science Methods to Build Capacity in Education Research." PI: **Odis Johnson Jr.**, co-PI Ebony McGee. $75,000

01/24 - 12/26 National Science Foundation (ECR-BCSER #2321179) – "Critical Data Science Methods to Broaden Capacity in STEM Education Research." PI: **Odis Johnson Jr**., Co-PI Ebony McGee. $999,980

10/23 – 09/25 National Science Foundation (#2322015) – "Promoting Research Using Adaptive Testing to Support Individualized Instruction at Scale." PI: Ben Van Dusen, Co-PIs Jayson Nissen, Paulette Vincent-Ruz, Darnell Leatherwood, Senior Personnel: Odis Johnson Jr., Hua Hua Chang, Jason Morphew. - $500,000

01/23 – 07/23 Spencer Foundation – "Critical Data Science Methods to Build Capacity in Education Research for Doctoral Students." PI: Ebony McGee, co-PI Odis Johnson Jr. $25,000

01/20 – 12/22 William T. Grant Foundation (#190932) – "Institute in Critical Quantitative and Computational Training (ICQCM)." (Co-PI: Ezekiel Dixon Roman [$161,610], Co-PI: Odis Johnson Jr., [$175,023], Co-PI: Ebony McGee, Vanderbilt [$63,367]) – $400,000

01/20 – 12/22 National Science Foundation (#EHR-BCSER - 1937687/1937490/1937391) – "Collaborative Research: Institute in Critical Quantitative, Computational, and Mixed Methods Training (ICQCM)." PI: **Odis Johnson Jr**., Co-PI: Ebony McGee, Co-PI: Ezekiel Dixon Roman – $ 1,001,309.

01/20 – 12/20 Spencer Foundation (#202000127) – "Institute in Critical Quantitative, Computational, and Mixed Methods Training (ICQCM)." (Co-PI: Ebony McGee, Co-PI: Odis Johnson Jr., Co-PI: Ezekiel Dixon Roman) – $149,996.

08/18 – 07/20 National Science Foundation (#DRL-1800199) – "Exploring ways to use National Datasets to Promote Broader Participation in STEM" (PI: **Odis Johnson Jr**., Co-PIs: None) – $299,999.

06/16 – 05/20 National Science Foundation (#EEC-1619843) – "Trajectories in Engineering: The Role of Social Control across Neighborhood and School Contexts" (PI: **Odis Johnson Jr.**, Co-PIs: None) – $617,202.

05/10 – 06/12 National Science Foundation (#DRL-0941014) – "Study of Ecological Determinants of the Achievement Gap" (PI**: Odis Johnson Jr.**, Co-PIs: None) - $34,999.

06/10 – 12/11 Spencer Foundation (#201000103) – "Study of Ecological and Seasonal Variation in Achievement Inequality" (PI: **Odis Johnson Jr**., Co-PIs: None) - $40,000.

D. PEER REVIEWED JOURNAL MANUSCRIPTS (Partial List)

(* = Publication with Postdoctoral Fellow, Graduate or Undergraduate Student)

Johnson, Jr., O. 2025. "Technology Innovates to Incarcerate: The Growing Reliance of Schools on AI Surveillance." In *Meena Dhanda (ed.), Oxford Intersections (Oxford, online edn, Oxford Academic, 20 Mar. 2025).* Oxford University Press.

Johnson, Jr., O. 2025. "What were Critical Quantitative Methods Back Then: Does it Inform what they are today?" *Current Opinion in Behavioral Sciences*. Volume 66, December 2025, 101609.

Lofton, R., Gong, Catherine, Priyanka Fernandes, Joshua M. Sharfstein, and **O. Johnson Jr**. 2025. "The

Physical Conditions of Schools, Absenteeism, and Test Scores in Maryland." *Policy Insights from Behavioral and Brain Sciences.* https://doi.org/10.1177/237273222251363170

*Huang, Wenrui, Jabbari, Jason, **O. Johnson Jr.,** and Chun, Y. 2025. "Can Certificate Programs Solve the Skills and Spatial Mismatch Problem? Job Portability and Residential Mobility in a Coding and Apprenticeship Program." *Urban Education,* doi: https://doi.org/10.1177/00420859251315480

Jabbari, Jason and **O. Johnson Jr**. 2024. "Multiplying Disadvantages in U.S. High Schools: An Intersectional Analysis of the Interactions among Punishment and Achievement Trajectories." *AERA Open*. https://doi.org/10.1177/23328584241230971

*Jabbari, Jason, W. Huang and **O. Johnson Jr**. 2023. "Broadening Participation in STEM through Alternative Preparation Programs." *Journal of Women and Minorities in Science and Engineering*, vol. 29, issue 6, pp. 1-47. 10.1615/JWomenMinorScienEng.2022041267

Jabbari, Jason and **O. Johnson, Jr**. 2023. "The Collateral Damage of In-School Suspensions: A Counterfactual Analysis of High-Suspension Schools, Math Achievement, and College Attendance." *Urban Education*, 58(5), 801-837 https://doi.org/10.1177/0042085920902256

Jabbari, Jason and **O. Johnson Jr.** 2022. "Perceptions of School Quality and Student Learning During the Pandemic: Exploring the Role of Students, Families, Schools, and Neighborhoods." *Socius*, 8, https://doi.org/10.1177/23780231221142955

*Jabbari, Jason and **O. Johnson Jr**. 2021. "The Process of 'Pushing Out': Accumulated Disadvantage across School Punishment and Math Achievement Trajectories." *Youth & Society,* https://doi.org/10.1177/0044118X211007175.

*Jabbari, Jason and **O. Johnson, Jr**. 2020. "Veering Off Track in US High Schools? Redirecting Student Trajectories by Disrupting Punishment and Math Course-taking Tracks." *Child and Youth Services Review*. https://doi.org/10.1016/j.childyouth.2019.104734

Johnson, Jr., O and M. Wagner. 2017. "Equalizers or Enablers? A Counterfactual Analysis of Residential Test-Score Gaps in Year-Round and 9-Month Schools." *Annals of the American Academy of Political and Social Science, 674, 1, 240-261*.

Johnson, Jr., O. 2015. "Responding to School Violence: Confronting the Columbine Effect." *Contemporary Sociology*, *44(4): 539-541*.

Johnson, Jr., O. 2012. "Relocation Programs, Opportunities to Learn and the Complications of Conversion." *Review of Educational Research*, *82 (2): 131-178*. **(Winner of the American Educational Research Association 2013 Outstanding Review of Research Award)**.

Johnson, Jr., O. 2012. "A Systematic Review of Neighborhood and Institutional Relationships Related to Education." *Education and Urban Society*, *44 (4): 477-511.*

E. SCIENTIFIC PRESENTATIONS, PANELS and KEYNOTES (Partial List)

Artificial Intelligence and Social Services Research and Practice. Speaker at Morehouse School of Medicine. NAARCF, August 2025.

AI and Ethics. Keynote. How can AI help improve educational assessments. International Association of Educational Assessment. Philadelphia, PA. Sept 24, 2024

From Principles to Practice: Operationalizing Responsible AI in Educational Measurement. Panelist at the Annual Meeting of the National Council on Measurement in Education. 4:30 PM - 6:00 PM Hilton Denver City Center Lower Level 2: Colorado

Critical Quantitative and Computational Methods: Rigor and an Emancipatory Education. August 18, 4:00 pm, 2023. Invited Presidential Session at the Annual Meeting of the American Sociological Association. Pennsylvania Convention Center, 200 Level, Ballroom A/B.

Quantitative Methods for Rigorous Research. April 8, 11am. Invited Presidential Session, Annual Meeting of the American Educational Research Association. San Diego, CA 2022.

Expert Testimony, "The Future of Methods and Measures in Education Research", National Academies of Science, Engineering, and Medicine**,** July 7, 2021.

James Diamond, Ph.D.
465 17th St., Apt. 4B • Brooklyn, NY 11215 • 646-515-3921 • jdiamo13@jhu.edu

EDUCATION

2012	Ph.D. in Educational Communication & Technology New York University, Steinhardt School of Culture, Education, and Human Development, New York, NY.
2003	Ed.M. in Educational Media & Technology Boston University, Wheelock School of Education, Boston, MA.
1994	B.A. in History and International Relations, *magna cum laude* Boston University, University Professors Program, Boston, MA.

PROFESSIONAL APPOINTMENTS

Aug 2018–Present	Assistant Professor, Johns Hopkins School of Education (SOE), Program in Digital Age Learning and Educational Technology, Baltimore, MD.
Jan 2018–Jul 2018	Research Scientist, Education Development Center | Center for Children and Technology (EDC|CCT), NY, NY.
2015–2017	Senior Research Associate, EDC|CCT.
2013–2014	Research Associate II, EDC|CCT.
2009–2012	Research Associate I, EDC|CCT.
2007–2008	Research Assistant II, EDC|CCT.

FUNDED RESEARCH AND DEVELOPMENT

2024	PI, *AI Synergy Summit,* JHU Nexus Convening Award, $99,000.
2023	Co-Principal Investigator (co-PI), *Fostering Digital Well-Being in Online Education at JHU,* JHU DELTA Award, $75,000.
2022	PI, *Exploring Virtual and Augmented Reality for SOE Classes with Meta Quest Pro*, JHU SOE Department Chairs Innovation Award, $2,000.
2022	PI, *Exploring Virtual and Augmented Reality for SOE Classes with Meta Quest Pro* JHU SOE Senate Pitch Grant Award, $1,500.
2022	Co-PI, *JHU Online Excellence: A Multi-Pronged Approach to Prepare Faculty for Excellence in Online Teaching and Learning,* JHU DELTA Award, $75,000.
2020	Co-PI, Computer science pre-service seed grant, Maryland Center for Computing Education, grant, $10,000.
2020	Principal Investigator (PI), *Using Gamification and the Adolescent Community Engagement (ACE) Framework to Support At-Risk Youth in Baltimore City*

	Schools During Online and Hybrid Classes, MD Governor's Emergency Education Relief (GEER) Fund, $374,926.
2020	Lead MOOC course developer, Social Entrepreneurship in Educational Technology, Johns Hopkins Provost's Office/FutureLearn grant, $20,000.
2019	Co-PI (formerly PI), *An exploratory study of classroom implementations of LEGO Education kits in Philadelphia public schools*, LEGO Education, evaluation, $150,000.
2019	PI (formerly co-PI), *Improving care and management of pediatric asthma through a gamified collaborative platform*, Johns Hopkins Discovery Award, grant, $100,000.
2018	PI, *Building CT Readiness: Refining and studying a framework for integrating computational thinking across subjects in high-poverty elementary school,* National Science Foundation, Award #1838523, $299,881.
2017	PI, *Identifying effective models for integrating computational thinking into NYC elementary schools*, Robin Hood Learning + Technology Fund, grant, $600,000.
2016	PI, *Investigating digital badges as alternative credentials to broaden STEM participation among underrepresented youth*, National Science Foundation, Award #1614727, $1,199,945.
2015	PI (formerly co-PI)*, Playing with the data: Developing digital supports for middle school science teachers using game-based formative assessment*, National Science Foundation, Award #1503255, $2,818,793.
2014	PI, *Planning a design-based implementation research agenda to investigate digital badges as transformative assessment in informal science learning*, National Science Foundation, Award #1451303, $115,000.
2012	PI, *A design-based research project to study the Who Built America? Teacher Mastery Badge System*, MacArthur Foundation Digital Media and Learning Research Competition on Badging and Badge Systems Development with funding from the Bill & Melinda Gates Foundation, grant, $195,000.
2012	Assessment Lead, *Who Built America? Online professional development from the American Social History Project,* MacArthur Foundation Digital Media and Learning Competition, grant, $175,000.

PUBLICATIONS AND PRESENTATIONS

Presentations

Harnett, C., Diamond, J., Friedlander, T., & Cooney, K. (2024/November,13). Belonging Matters in Digital Spaces. [Virtual Presentation]. Johns Hopkins University, Office of Diversity and Inclusion/ Spotlight on Belonging, Baltimore, Maryland.

Harnett, C., Diamond, J., Friedlander, T., & Cooney, K. (2024/May, 2). Fostering Digital Well-Being in Online Education: Wellness by Design. [Presentation]. Johns Hopkins University, Provost's DELTA Teaching Forum, Baltimore, MD.

Book Chapters

Diamond, J. (2021). Designing and teaching *Gaming and Simulations for Learning:* a syllabus analysis. In R. Ferdig, E. Baumgartner, & E. Gandolfi (Eds.), *Teaching the game: An interdisciplinary collection of game course syllabi.* ETC Press.

Joseph, R., & Diamond, J. (2017). *iDesign:* Designing and implementing a culturally relevant game-based curriculum. In A.D. Benson, R. Joseph, & J.L. Moore (Eds.), *Culture, learning, and technology: research and practice* (pp. 151–164). New York: Routledge.

Diamond, J., & Gonzalez, P. (2016). Digital badges for professional development: teachers' perceptions of the value of a new credentialing currency. In D. Ifenthaler, N. Bellin-Mularski, & D.K. Mah (Eds.), *Foundation of digital badges and micro-credentials: Demonstrating knowledge and competencies* (pp. 391–409). Switzerland: Springer International Publishing.

Schrier, K., Diamond, J., & Langendoen, D. (2010) *Using Mission US: For Crown or Colony?* to develop historical empathy and nurture ethical thinking. In K. Schrier & D. Gibson (Eds.), *Ethics and game design: teaching values through play* (pp. 255–273). Hershey, NY: Information Science Reference.

Invited Conference Proceedings

Diamond, J. (2024). Redefining Learning Goals in the AI Era. 25th Ireland International Conference on Education, Dublin, Ireland. [Invited Keynote Presentation]

Diamond, J. (2017). Learning how to help teachers use gameplay for formative assessment. [Conference presentation]. Annual Meeting of the National Center for Research on Evaluation, Standards and Student Testing Eastern Educational Research Association (CRESST), Los Angeles, CA.

Diamond, J. (2015). Playing with data: Understanding how teachers use videogame play for formative assessment. [Conference presentation]. Annual Meeting of the National Center for Research on Evaluation, Standards and Student Testing Eastern Educational Research Association (CRESST), Los Angeles, CA.

Conference Proceedings

Harnett, C., Diamond, J., Friedlander, T., & Cooney, K. (2024). Fostering Digital Well-Being in Online Education. 25th Ireland International Conference on Education Dublin, Ireland. [Conference Presentation]

BIOGRAPHICAL SKETCH

Provide the following information for the Senior/key personnel and other significant contributors.
Follow this format for each person. DO NOT EXCEED FIVE PAGES.

NAME: Yang, Hao Frank

eRA COMMONS USER NAME (credential, e.g., agency login):

POSITION TITLE: Assistant Professor

EDUCATION/TRAINING *(Begin with baccalaureate or other initial professional education, such as nursing, include postdoctoral training and residency training if applicable. Add/delete rows as necessary.)*

INSTITUTION AND LOCATION	DEGREE (if applicable)	END DATE MM/YYYY	FIELD OF STUDY
Beijing University of Posts and Telecommunications, Beijing, Beijing	BENG	06/2017	Electrical Engineering
University of London, London	BS	06/2017	Electrical Engineering
University of Washington, Seattle, Washington	MS	03/2020	Transportation Engineering
University of Washington, Seattle, Washington	PHD	07/2023	Civil Engineering (Transportation)

A. Personal Statement

PI Yang has led pioneering research on foundation models, especially large language models (LLMs) for decision-making and large-scale simulation, focusing on integrating domain knowledge to enhance predictive adaptability, interpretability, and reliability in public health and transportation applications. His recent studies demonstrate how LLMs can support trustworthy, data-driven decisions for public health and traffic safety. These include multimodal foundation model forecasting of hospitalization trends for infectious-disease intervention in Nature Computational Science (Editors' Highlights, 2025), and trustworthy traffic crash prediction with feature attribution of risk factors for safety decisions in Nature Communications (Editors' Highlights, 2025). Methodologically, his team has advanced LLM reasoning through pre-training knowledge detection and interpretation (ICLR 2025, Spotlight), spatial-knowledge integration and optimal reasoning for reliable decision-making (NeurIPS 2025, Spotlight), and symbolic learning for interpretable and closed-form decision modeling (NeurIPS 2025, Spotlight). Collectively, these works contribute to a new generation of health-focused decision-support systems powered by LLMs. Furthermore, Dr. Yang's prior foundational research on out-of-distribution detection, data bias mitigation, and Mixture-of-Experts (MoE) computing architectures, published in IEEE Transactions on Pattern Analysis and Machine Intelligence (TPAMI), CVPR, ICCV, and NeurIPS, which provides transferable technical underpinnings for efficient and robust large-scale LLM-based decision simulations to this project.

1. Linshen Liu, Boyan Su, Junyue Jiang, Guanlin Wu, Cong Guo, Ceyu Xu, Hao Frank Yang. Towards Accurate and Efficient 3D Object Detection for Autonomous Driving: A Mixture of Experts Computing System on Edge. Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV); 2025; arXiv; c2025. Available from: https://arxiv.org/abs/2507.04123 DOI: 10.48550/ARXIV.2507.04123
2. Ruimin Ke, Jerome M. Lutin, Yinhai Wang, Zhiyong Cui, Shuyi Yin, Yifan Zhuang, Hao (Frank) Yang. Transit Safety System Evaluation and Hotspot Identification Empowered by Edge Computing Transit Event Logging System. Transportation Research Record: Journal of the Transportation Research Board. 2024 January. DOI: 10.1177/03611981231171910
3. Hao (Frank) Yang, Jiarui Cai, Chenxi Liu, Ruimin Ke, Yinhai Wang. Cooperative multi-camera vehicle tracking and traffic surveillance with edge artificial intelligence and representation learning. Transportation Research Part C: Emerging Technologies. 2023 March. DOI: 10.1016/j.trc.2022.103982

4. Hao (Frank) Yang, Yifan Ling, Cole Kopca, Sam Ricord, Yinhai Wang. Cooperative traffic signal assistance system for non-motorized users and disabilities empowered by computer vision and edge artificial intelligence. Transportation Research Part C: Emerging Technologies. 2022 December. DOI: 10.1016/j.trc.2022.103896

B. Positions, Scientific Appointments and Honors

Positions and Scientific Appointments

2024 -	Assistant Professor, Johns Hopkins University, Department of Civil and System Engineering, Baltimore, MD
2024 -	Assistant Professor, Johns Hopkins University, Johns Hopkins Data Science and AI Institute, Baltimore, MD
2024 - 2024	Research Assistant Professor, Johns Hopkins University, Department of Civil and System Engineering, Center for Systems Science and Engineering (CSSE), Baltimore, MD
2023 - 2024	Research Scientist, NSF National AI Institute (Athena), Department of Electrical and Computer Engineering, Duke University, Durham, NC
2022 - 2022	Research Scientist Intern, Amazon Cloud Service (AWS) AI Research, Amazon, Seattle, WA
2021 - 2021	Research Scientist Intern, Microsoft Research (MSR), Seattle, WA

Honors

2025	Advisor of the Amazon AI PhD Fellowship Awardee, Amazon
2024	Best Paper Award, Transportation Research Board Annual Meeting, National Academies of Sciences, Engineering, and Medicine
2022	High Value Research Award, American Association of State Highway and Transportation Officials

C. Contribution to Science

1. Trustworthy AI for Decision Making
   a. Hongru Du, Yang Zhao, Jianan Zhao, Shaochong Xu, Xihong Lin, Yiran Chen, Lauren M. Gardner, Hao ‘Frank’ Yang. Advancing real-time infectious disease forecasting using large language models. Nature Computational Science. 2025 June. DOI: 10.1038/s43588-025-00798-6
   b. Yang Zhao, Pu Wang, Yibo Zhao, Hongru Du, Hao Frank Yang. SafeTraffic Copilot: adapting large language models for trustworthy traffic safety assessments and decision interventions. Nature Communications. 2025 October. DOI: 10.1038/s41467-025-64574-w
   c. Yang Zhao, Pu Wang, Hao Frank Yang. How to Auto-optimize Prompts for Domain Tasks? Adaptive Prompting and Reasoning through Evolutionary Domain Knowledge Adaptation. NeurIPS 2025; 2025; arXiv; c2025. Available from: https://arxiv.org/abs/2510.21148 DOI: 10.48550/ARXIV.2510.21148
   d. Yibo Zhao, Yang Zhao, Hongru Du, Hao Frank Yang. Personalized Decision Modeling: Utility Optimization or Textualized-Symbolic Reasoning. NeurIPS 2025 (Spotlight); 2025; arXiv; c2025. Available from: https://arxiv.org/abs/2511.02194 DOI: 10.48550/ARXIV.2511.02194
2. Mutimodal Infromation Integratation and Robustness Improvement
   a. Hongru Du, Yang Zhao, Jianan Zhao, Shaochong Xu, Xihong Lin, Yiran Chen, Lauren M. Gardner, Hao ‘Frank’ Yang. Advancing real-time infectious disease forecasting using large language models. Nature Computational Science. 2025 June. DOI: 10.1038/s43588-025-00798-6
   b. Guanlin Wu, Boyan Su, Yang Zhao, Pu Wang, Yichen Lin, Hao Frank Yang. Towards Physics-informed Spatial Intelligence with Human Priors: An Autonomous Driving Pilot Study. NeurIPS 2025 (Spotlight); 2025; arXiv; c2025. Available from: https://arxiv.org/abs/2510.21160 DOI: 10.48550/ARXIV.2510.21160

c. Wenqing Zheng, Hao (Frank) Yang, Jiarui Cai, Peihao Wang, Xuan Jiang, Simon Shaolei Du, Yinhai Wang, Zhangyang Wang. Integrating the traffic science with representation learning for city-wide network congestion prediction. Information Fusion. 2023 November. DOI: 10.1016/j.inffus.2023.101837
d. Hao (Frank) Yang, Jiarui Cai, Chenxi Liu, Ruimin Ke, Yinhai Wang. Cooperative multi-camera vehicle tracking and traffic surveillance with edge artificial intelligence and representation learning. Transportation Research Part C: Emerging Technologies. 2023 March. DOI: 10.1016/j.trc.2022.103982

3. Efficient Computing Systems for Large-scale Simulation
   a. Yibo Zhao, Yang Zhao, Hongru Du, Hao Frank Yang. Personalized Decision Modeling: Utility Optimization or Textualized-Symbolic Reasoning. NeurIPS 2025 (Spotlight); 2025; arXiv; c2025. Available from: https://arxiv.org/abs/2511.02194 DOI: 10.48550/ARXIV.2511.02194
   b. Linshen Liu, Pu Wang, Guanlin Wu, Junyue Jiang, Hao Yang. Towards Optimal Mixture of Experts System for 3D Object Detection: A Game of Accuracy, Efficiency and Adaptivity. IEEE Transactions on Pattern Analysis and Machine Intelligence. 2025. DOI: 10.1109/tpami.2025.3611795
   c. Linshen Liu, Boyan Su, Junyue Jiang, Guanlin Wu, Cong Guo, Ceyu Xu, Hao Frank Yang. Towards Accurate and Efficient 3D Object Detection for Autonomous Driving: A Mixture of Experts Computing System on Edge. Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV); 2025; arXiv; c2025. Available from: https://arxiv.org/abs/2507.04123 DOI: 10.48550/ARXIV.2507.04123
   d. Di Chen, Meixin Zhu, Hao Yang, Xuesong Wang, Yinhai Wang. Data-Driven Traffic Simulation: A Comprehensive Review. IEEE Transactions on Intelligent Vehicles. 2024 April. DOI: 10.1109/TIV.2024.3367919
4. Mobile Sensing and Edge Computing
   a. Linshen Liu, Boyan Su, Junyue Jiang, Guanlin Wu, Cong Guo, Ceyu Xu, Hao Frank Yang. Towards Accurate and Efficient 3D Object Detection for Autonomous Driving: A Mixture of Experts Computing System on Edge. Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV); 2025; arXiv; c2025. Available from: https://arxiv.org/abs/2507.04123 DOI: 10.48550/ARXIV.2507.04123
   b. Ruimin Ke, Jerome M. Lutin, Yinhai Wang, Zhiyong Cui, Shuyi Yin, Yifan Zhuang, Hao (Frank) Yang. Transit Safety System Evaluation and Hotspot Identification Empowered by Edge Computing Transit Event Logging System. Transportation Research Record: Journal of the Transportation Research Board. 2024 January. DOI: 10.1177/03611981231171910
   c. Hao (Frank) Yang, Jiarui Cai, Chenxi Liu, Ruimin Ke, Yinhai Wang. Cooperative multi-camera vehicle tracking and traffic surveillance with edge artificial intelligence and representation learning. Transportation Research Part C: Emerging Technologies. 2023 March. DOI: 10.1016/j.trc.2022.103982
   d. Hao (Frank) Yang, Yifan Ling, Cole Kopca, Sam Ricord, Yinhai Wang. Cooperative traffic signal assistance system for non-motorized users and disabilities empowered by computer vision and edge artificial intelligence. Transportation Research Part C: Emerging Technologies. 2022 December. DOI: 10.1016/j.trc.2022.103896

CURRICULUM VITAE
Lingxin Hao
Google Scholar
October 2025

PRESENT POSITION

2021-present	Benjamin H. Griswold III Professor in Public Policy, Department of Sociology, Johns Hopkins University
2003-present	Professor, Department of Sociology, Johns Hopkins University
2021-present	Scientific Core Director, Hopkins Population Center, Johns Hopkins University
2025-present	Member, Johns Hopkins Data Science and AI Institute

PREVIOUS POSITIONS

2015-2023	Director, Hopkins Population Center, Johns Hopkins University
1998-2003	Associate Professor, Department of Sociology, Johns Hopkins University
1996-1998	Assistant Professor, Department of Sociology, Johns Hopkins University
1990-1997	Assistant to Associate Professor (with tenure), Department of Sociology, University of Iowa

EDUCATION

Ph.D.	1990	Sociology, University of Chicago
M.A.	1985	Sociology, Sun Yat-sen University, China
B.A.	1982	English, South China Normal University, China

AWARDS AND FELLOWSHIPS

2025	Distinguished Social Sciences Scholar, University of Hong Kong
2010-2016	Wu Yuzhang Professor, Renmin University of China
2007	Resident Fellow, Spencer Foundation
2002-2003	Resident Fellow, Russell Sage Foundation
2003-present	Elected member, Sociological Research Association (SRA)

AREAS OF SPECIALIZATION

Sociology of Education, Social Demography; Migration; Family and Public Policy; Quantitative and Computational Methodology

PROFESSIONAL ACTIVITIES

2025	Reviewer, Science and Technology Center Program (NSF-STC)
2023-2025	Advisory Panelist, the Methodology, Measurement, and Statistics (NSF SES-MMS)
2022	Advisory Panelist, the Future of Work (NSF FW-HTF) program
2010-2017	Member, the Technical Review Committee for the National Longitudinal Surveys Bureau of Labor Statistics
2006-2011	Special Emphasis Panelist, NIH R03, R21, and R15 grant review committee (ZRG1 HOP B 90 S), the U.S. National Health Institute (NIH)
2009	Panelist, NIH P01 grant review committee (ZHD1-DSR-W (LN), the U.S. National Health Institute (NIH)
1998-2023	Editorial Board, *Sociology of Education, Sociological Methodology, Journal of Marriage and Family, Demography, American Journal of Sociology*

PROFESSIONAL ASSOCIATIONS

Population Association of America (PAA)

American Sociological Association (ASA)

EXTRAMURAL RESEARCH GRANTS IN THE LAST 10 YEARS

Hao, Lingxin (Principal Investigator). 2021-2026. "Hopkins Population Center." National Institute of Health (NIH) P2CHDHD042854. ($2,178,170).

Hao, Lingxin (Co-Principal Investigator). 2020-2024. "Methods and Applications for Massive One-mode and Bipartite Social Networks." National Science Foundation (NSF). SES1951005 (PI: Angelo Mele). ($549,592)

Hao, Lingxin (Principal Investigator). 2014-2020. "Research Infrastructure for the Hopkins Population Center." National Institute of Health (NIH) R24HD042854-13. ($300,000)

Hao, Lingxin (Principal Investigator). 2015-2017. "Agent-Based Modeling of Internal Migration." National Institute of Health (NIH) R21. ($438,709)

Hao, Lingxin (Principal Investigator). 2013-2016. "Student Migration and Education Segregation." National Science Foundation (NSF). ($348,666)

PUBLICATIONS RELEVANT FOR THE PROPOSAL

BOOKS

Hao, Lingxin and Daniel Naiman. 2010. *Assessing Inequality*. Thousand Oaks CA: Sage Publications. (Versions in English and Chinese)

Hao, Lingxin and Daniel Q. Naiman. 2007. *Quantile Regression*. Thousand Oaks CA: Sage Publications. (Versions in English and Chinese)

ARTICLES IN REFERRED JOURNALS

Zheng, Jiayu, Lingxin Hao, et al., Anqi Liu. Forthcoming. "Do students rely on AI? Analysis of student-ChatGPT conversations from a field study." *Proceedings of the Eighth AAAI/ACM Conference on AI, Ethics, and Society (AIES-25)*.

Lu, Kelun, Lingxin Hao, I-Jeng Wang, and Anqi Liu. 2024. "Eighth Graders and a Math Intelligent Tutoring System: A Deep Neural Network Analysis." A. M. Olney et al. (Eds.): AIED 2024 Workshops, CCIS 2151, pp. 38–46, 2024. https://doi.org/10.1007/978-3-031-64312-5_5.

Mele, Angelo, Hao, Lingxin, Cape, Joshua, Athreya, Avanti, and Priebe, Carey E. 2022. "Spectral Estimation of Large Stochastic Blockmodels with Discrete Nodal Covariates." *Journal of Business and Economics Statistics*. https://doi.org/10.1080/07350015.2022.2139709.

Mu, Cong, Mele, Angelo, Hao, Lingxin, Cape, Joshua, Athreya, Avanti, and Priebe, Carey E. Priebe. 2022. "On Spectral Algorithms for Community Detection in Stochastic Blockmodel Graphs with Vertex Covariates." IEEE *Transactions on Network Science and Engineering*, doi: https://doi.org/10.1109/TNSE.2022.3177708.

Hao, Lingxin and Dong Zhang. 2020. "China's College Expansion and the Timing of College-to-Work Transition: A Natural Experiment." *The ANNALS of the American Academy of Political and Social Science*, 688, 93-114. PMC7410005.

Fu, Zhaohao and Lingxin Hao. 2018. "Agent-Based Modeling of China's Rural-Urban Migration and Social Network Structure." *Physica A*, 490, 1067-1075. http://www.sciencedirect.com/science/article/pii/S0378437117308853?via%3Dihub.

Hao, Lingxin and Xiao Yu. 2017. "Sources of Unequal Cognitive Development of Middle-School Students in China's Rural-Urban Migration Era." *Chinese Journal of Sociology*, 3(1), 32-55. (DOI) https://doi.org/10.1177/2057150X16684115.

Hao, Lingxin and Wei-Jun Jean Yeung. 2015. "Parental Spending on School-Age Children: The Role of SES, Race, and Parental Expectation." *Demography* 52(3):835-860. (DOI) 10.1007/s13524-015-0386-1.

Hao, Lingxin, Alfred Hu, and Jamie Lo. 2014. "Two Aspects of the Rural-Urban Divide and Educational Stratification in China: A Trajectory Analysis." *Comparative Education Review* 58(3):509-536.

Hao, Lingxin and Han Soo Woo. 2012. "Distinct Trajectories in the Transition to Adulthood: Are Children of Immigrants Advantaged?" *Child Development* 83(5):1623-1639.

Hao, Lingxin and Suet-ling Pong. 2008. "The Role of School in Upward Mobility of Disadvantaged Immigrants' Children." *The ANNALS of the American Academy of Political and Social Science* 620(1):62-89.

Pong, Suet-ling and Lingxin Hao. 2007. "Neighborhood and School Factors in the School Performance of Immigrants' Children." *International Migration Review* 41(1):206-241.

Hao, Lingxin and Ross L. Matsueda. 2006. "Family Dynamics through Childhood: A Sibling Model of Behavior Problems." *Social Science Research* 35:500-524.

Pong, Suet-ling, Lingxin Hao, and Erica Gardner. 2005. "The Roles of Parenting Styles and Social Capital in the School Performance of Immigrant Asian and Hispanic Adolescents." *Social Science Quarterly* 86(4):928-950.

Portes, Alejandro and Lingxin Hao. 2004. "The Schooling of Children of Immigrants: Contextual Effects on the Educational Attainment of the Second Generation." *Proceedings of National Academy of Sciences* 101:11920-27. Spanish Translation: *Migraciones* 17 (2005): 7-44.

Hao, Lingxin, Nan M. Astone and Andrew J. Cherlin. 2004. "Adolescents' School Enrollment and Employment: Effect of State Welfare Policies." *Journal of Policy Analysis and Management* 23:697-721.

Portes, Alejandro and Lingxin Hao. 2002. "Linguistic Adaptation, Acculturation, and Gender in the Immigrant Second Generation". *Ethnic and Racial Studies* 25:889-912.

Hao, Lingxin and Guihua Xie. 2002. "The Complexity and Endogeneity of Family Structure in Explaining Children's Misbehavior." *Social Science Research* 31:1-28.

Hao, Lingxin and Melissa Bonstead-Bruns. 1998. "Parent-Child Difference in Educational Expectations and Academic Achievement of Immigrant and Native Students." *Sociology of Education* 71:175-198.

Portes, Alejandro and Lingxin Hao. 1998. "E Pluribus Unum: Bilingualism and Language Loss in the Second Generation." *Sociology of Education* 71:269-294.

Coleman, James S. and Lingxin Hao. 1989. "Linear Systems Analysis: Macro-Level Analysis with Micro-Level Data." *Sociological Methodology* 19:395-422.

BIOGRAPHICAL SKETCH

Provide the following information for the Senior/key personnel and other significant contributors.
Follow this format for each person. **DO NOT EXCEED FIVE PAGES.**

NAME: Igusa, Takeru

eRA COMMONS USER NAME (credential, e.g., agency login): TIGUSA1

POSITION TITLE: Professor, Department of Civil and Systems Engineering

EDUCATION/TRAINING *(Begin with baccalaureate or other initial professional education, such as nursing, include postdoctoral training and residency training if applicable. Add/delete rows as necessary.)*

INSTITUTION AND LOCATION	DEGREE *(if applicable)*	Completion Date MM/YYYY	FIELD OF STUDY
Harvard University	B.A.	06/1977	Applied Mathematics
University of California, Berkeley	Ph.D.	06/1983	Civil Engineering
University of California, Berkeley	Postdoctoral Fellow	06/1985	Civil Engineering

A. Personal Statement

During the past ten years, I have used my background in applied mathematics and engineering to apply systems science and artificial intelligence (AI) methods to a wide range of public health problems. These include participatory modeling, causal inference, and advanced machine learning (ML) techniques such as Bayesian kernel machine regression (BKMR) and epigenome-wide association studies (EWAS). This combination of systems approaches and AI-driven analytics is central to the current R01 proposal.

One of the first large, funded activities I helped initiate was an NIH center of excellence on Systems-Related Pediatric Obesity Research and Training, sponsored by NICHD and OBSSR through the U54 mechanism. I served as Director of the Education and Training Core, mentoring doctoral and postdoctoral trainees and leading participatory modeling initiatives. I have since applied participatory systems methods to public health challenges including food insecurity, immunization, suicide prevention, and maternal and child health. (Food insecurity: Estradé et al., 2023; Immunization: Schuh et al., 2017)

In parallel, I have contributed to AI/ML research in population health, including an EWAS study of maternal smoking and birthweight, and a BKMR study on in-utero exposure to metals and micronutrients. These efforts demonstrate my experience with flexible, high-dimensional ML methods for environmental and developmental epidemiology. (EWAS: Xu et al., 2021; BKMR: Huang et al., 2022)

I have also collaborated on the development of public-facing decision-support tools for health professionals and communities. These dashboards and mobile applications are described in more detail in Part C, Contribution to Science #4.

The most important hallmark of my work is collaboration – discovering ways to adapt and integrate novel methodological tools to address public health research questions. This has helped produce a wide array of funded projects, listed below, ranging from smoking cessation (P50), improvements to the food environment (R01), HIV/stigma (R01), and maternal and newborn health (Gates Foundation).

Ongoing and recently completed projects that I would like to highlight include:
NIMH 2P50MH115842-05
Daumit (PI), Role: Co-Investigator
06/01/2024 – 04/30/2029

Center to Accelerate Translation of Interventions to Decrease Premature Mortality in Persons with Serious Mental Illness

NIMHD 1R01 MD018022-01
09/26/2022-06/30/2027
Gittelsohn (PI), Role: Co-Investigator
Systems Science Approaches to Improve Access to Healthier Foods: The FRESH Trial

NIMH R01 MH132150-01
07/08/2022-05/31/2027
Baral (PI), Role: Co-Investigator
Integrating the Visualization and Use of Stigma Data to Maximize the Impact of the Ending the HIV Epidemic Initiative

NHLBI 1R34 HL161566-01
12/07/2022 - 11/30/2025
Gittelsohn, Barnett (MPIs), Role: Co-Investigator
Developing a Support Application for Food Pantries (SAFPAS) to Improve Client Access to Healthy Foods & Enhance Emergency Preparedness

INV-032846 (Bill & Melinda Gates Foundation)
06/01/2021-05/31/2024
Alonge (PI), Role: Co-Investigator
Modeling the impact of Service Delivery Redesign in different contexts

NICHD 1U54 HD070725-01 (parent project)
Igusa (PI), Role: Director of the Education and Training Core
09/22/2011-06/30/2016
Johns Hopkins Systems-Oriented Pediatric Obesity Research and Training (SPORT) Center, Education and Training Core

Citations:

1. Estradé M, Alarcon Basurto SG, McCarter A, Gittelsohn J, Igusa T, Zhu S, Poirier L, Gross S, Pardilla M, Rojo M, Lombard K. A Systems Approach to Identify Factors Influencing Participation in Two Tribally-Administered WIC Programs. Nutrients. 2023 Feb 28;15(5):1210.
2. Schuh HB, Merritt MW, Igusa T, Lee BY, Peters DH. Examining the structure and behavior of Afghanistan's routine childhood immunization system using system dynamics modeling. International Journal of Health Governance. 2017 Sep 4;22(3):212-27.
3. Xu R, Hong X, Zhang B, Huang W, Hou W, Wang G, Wang X, Igusa T, Liang L, Ji H. DNA methylation mediates the effect of maternal smoking on offspring birthweight: a birth cohort study of multi-ethnic US mother–newborn pairs. Clinical Epigenetics. 2021 Dec;13:1-3.
4. Huang W, Igusa T, Wang G, Buckley JP, Hong X, Bind E, Steffens A, Mukherjee J, Haltmeier D, Ji Y, Xu R. In-utero co-exposure to toxic metals and micronutrients on childhood risk of overweight or obesity: new insight on micronutrients counteracting toxic metals. Int. Journal of Obesity. 2022 Aug;46(8):1435-45.

B. Positions, Scientific Appointments, and Honors

Recent Positions

2015 -	Professor (Joint Appointment), Johns Hopkins University, Bloomberg School of Public Health, Department of Mental Health

2013 - Professor (Joint Appointment), Johns Hopkins University, Bloomberg School of Public Health, Department of International Health
1999 - Professor, Johns Hopkins University, Department of Civil and Systems Engineering
1996 - 1999 Professor, Northwestern, Department of Civil Engineering

C. Contributions to Science

1. Dashboards and Mobile Applications for Public Health Decision-Making

    1. Gittelsohn J, Poirier L, Zhu S and **Igusa T**. (2018). Staple Food Ordinance Model (Version 1.3). Johns Hopkins University, Baltimore, MD. [Mobile application software.] Available from: https://engineering.jhu.edu/tak/SFO/UI/index.html
    2. Sun S, Zhu S, **Igusa T**, Elmi N, Thaler K, Sufrin C, Burke A. (2020). LARC 411: A Guide to Long-Acting Reversible Contraception (Version 2.2). Johns Hopkins University, Baltimore, MD [Mobile application software.] Available from: https://engineering.jhu.edu/tak/LARC/LARC%20411_NEW.html
    3. Lightner A, Williamson S, Sundermeir SM, Lewis EC, Matsuzaki M, Poirier L, Thomas A, Reznar MM, Neff R, Barnett DJ, Oladimeji AT, **Igusa T**, Velez-Burgess, V, Attar N, Gagnon B, Moses L, Gittelsohn J. (2025). Formative research for an innovative digital application to strengthen healthy food access and resilience in food pantries: the Support Application for Food PAntrieS (SAFPAS) study. *Journal of Hunger & Environmental Nutrition*, Apr 2:1–8.
    4. Oladimeji, AT. (2025). *SAFPAS: Support Application for Food Pantries* (Version 1.0) [Mobile app]. Apple App Store. https://apps.apple.com/us/app/safpas/id6499470129

2. Systems dynamics research in medicine and public health

    1. Lyon AR, Maras MA, Pate CM, **Igusa T**, Vander Stoep A. Modeling the Impact of School-Based Universal Depression Screening on Additional Service Capacity Needs: A System Dynamics Approach. Adm Policy Ment Health. 2016 Mar;43(2):168-88. PMC4881856.
    2. Alonge O, Lin S, **Igusa T**, Peters DH. Improving health systems performance in low- and middle-income countries: a system dynamics model of the pay-for-performance initiative in Afghanistan. Health Policy Plan. 2017 Dec 1;32(10):1417-1426. PMC5886199.
    3. Schuh HB, Merritt MW, Igusa T, Lee BY, Peters DH. Examining the structure and behavior of Afghanistan's routine childhood immunization system using system dynamics modeling. International Journal of Health Governance. 2017 Sep 4;22(3):212-27.
    4. Squire MM, **Igusa T**, Siddiqui S, Sessel GK, Squire EN Jr. Cost-Effectiveness of Multifaceted Built Environment Interventions for Reducing Transmission of Pathogenic Bacteria in Healthcare Facilities. HERD. 2019 Apr;12(2):147-161. doi: 10.1177/1937586719833360. Epub 2019 Apr 16. PMID: 30991849.

3. Applications of systems science methods to research problems in medicine and public health

    1. Dugas AF, Jalalpour M, Gel Y, Levin S, Torcaso F, **Igusa T**, Rothman RE. Influenza forecasting with Google Flu Trends. PLoS One. 2013;8(2):e56176. PMC3572967
    2. Gittelsohn J, Mui Y, Adam A, Lin S, Kharmats A, **Igusa T**, Lee BY. Incorporating Systems Science Principles into the Development of Obesity Prevention Interventions: Principles, Benefits, and Challenges. Curr Obes Rep. 2015;4(2):174-81. PMC4452216.
    3. Xue H, Slivka L, **Igusa T**, Huang TT, Wang Y. Applications of systems modelling in obesity research. Obes Rev. 2018 Sep;19(9):1293-1308. doi: 10.1111/obr.12695. Epub 2018 Jun 25. PMID: 29943509.
    4. **Igusa T**, Hummers LK, Visvanathan K, Richardson C, Wigley FM, Casciola-Rosen L, Rosen A, Shah AA. Autoantibodies and scleroderma phenotype define subgroups at high-risk and low-risk for cancer. Ann Rheum Dis. 2018 Aug;77(8):1179-1186. doi: 10.1136/annrheumdis-2018-212999. Epub 2018 Apr 20. PMC6272061.

Ben Van Dusen is a national leader in STEM instructional improvement, large-scale assessment systems, and the integration of advanced analytics and AI into learning environments. His work focuses on designing and studying digital assessment systems, developing next-generation diagnostic models, and creating AI-supported tools that enhance students' learning experiences. Van Dusen serves as the founder and long-time director of the LASSO platform, a national research and instructional infrastructure funded by more than $5M in NSF awards, supporting instructors and researchers across over 250 institutions.

His contributions span three intersecting areas:

1. **Advanced Assessment and Measurement** — including IRT, diagnostic modeling, computer-adaptive tests, and large-scale online administration.
2. **AI-Enabled Learning Systems** — developing tutoring systems and guidance tools that draw on cognitive diagnostics, reasoning frameworks, and retrieval-augmented generation.
3. **STEM Education Infrastructure and Grant Leadership** — as PI or co-PI on multiple NSF projects, advancing national capacity for research-based assessments, instructor support tools, and data-driven improvement cycles.

Education

Ph.D., Science Education, University of Colorado Boulder, 2014
M.Ed., Educational Leadership, University of Oregon, 2005
B.S., Physics, University of California Berkeley, 2004

Professional Appointments

Associate Professor, Science Education, Iowa State University (2023–present)
Assistant Professor, Science Education, Iowa State University (2017–2023)
Postdoctoral Researcher, Colorado School of Mines (2014–2017)
Albert Einstein Distinguished Educator Fellow, National Science Foundation, 2010–2011
Teacher, High School Physics 2005-2010

Major Research Leadership and Grant Experience

2025 (under review): **Collaborative Research: Personalized Learning Pathways in Calculus (co-PI).** National Science Foundation. Amount: **$700,000**.
2025 (under review): **Collaborative Research: ASCEND: AI-Supported Cognitive Evaluation for Next-generation Diagnostics (PI).** National Science Foundation. Amount: **$2,000,000**.
2024–2025: **Building Partnerships to Recruit Recent STEM Graduates into a Masters of Arts in Teaching Program (co-PI, Award #2345165).** National Science Foundation. Amount: **$100,000**.
2023–2025: **Adaptive Testing to Support Individualized Instruction at Scale (PI, Award #2322015).** National Science Foundation. Amount: **$499,867**.
2022–2025: **Collaborative Research: Individualizing Instruction and Improving Research using Adaptive Testing (PI, Award #2141847).** National Science Foundation. Amount: **$441,494**.
2019–2024: **IUSE-HSI Collaborative Research: HDR: Developing Faculty Resources of Evidence-based Practices that Improve Learning in STEM (PI, Award #1928596).** National Science Foundation. Amount: **$2,050,450**.

2015–2024: **Improving Undergraduate STEM Education: Institutional and Community Transformation: Design and Development – Collaborative Research: Scaling Undergraduate STEM Transformation and Institutional Networks for Engaged Dissemination (SUSTAINED) (Co-PI, Award #1525338).** National Science Foundation. Amount: **$2,899,322**.

Research and Development Focus Areas
Assessment and Measurement Systems
Van Dusen specializes in designing, analyzing, and validating digital assessments used in large-enrollment STEM courses. His work includes:
• Development of **IRT- and CDM-based models** to estimate students' proficiency and underlying skills.
• Design of **computer-adaptive assessments** and algorithms for item selection and score estimation.
• Creation and validation of research-based assessments used across physics and mathematics.
• Leadership in **large-scale online administration**, data cleaning, reliability evaluation, and interpretation frameworks for instructors and researchers.

AI-Supported Learning Tools
Van Dusen leads efforts to integrate advanced AI methods with assessment infrastructure, including:
• Development of **AI tutoring companions** that guide students through conceptual reasoning.
• Design of **retrieval-augmented generation (RAG)** systems that ensure transparent grounding and reliable output.
• Collaboration on reasoning frameworks and prompt orchestration for multi-turn guidance tools.
• Research on how AI systems can support students' understanding of core ideas in introductory STEM courses.

Digital Infrastructure for STEM Education
As founder and director of LASSO for more than a decade, Van Dusen has:
• Built and maintained a **national assessment and data collection platform** serving instructors and researchers across physics, mathematics, engineering, and other fields.
• Managed platform development, data governance, and collaborative research partnerships.
• Integrated new capabilities such as automated scoring, instructor dashboards, flexible assessment workflows, and large-scale dataset pipelines.

Research Contributions and Scholarly Impact
Ben Van Dusen's research centers on assessment, quantitative methods, and scalable learning technologies in STEM education. Across 78 publications with more than **1,500 citations** (h-index **19**), his work advances modern measurement, diagnostic assessment, and large-scale data systems. His contributions span several interrelated areas:

Assessment, Measurement, and Psychometrics
Van Dusen has a sustained record of developing and evaluating STEM assessments, with work emphasizing item response theory, hierarchical modeling, and modern quantitative approaches. Highly cited publications include *Best Practices for Addressing Missing Data through Multiple*

Imputation (Woods et al., 2024), *Modernizing Use of Regression Models in PER* (Van Dusen & Nissen, 2019), and *Missing Data and Bias in PER* (Nissen, Donatello, & Van Dusen, 2019). His measurement work also includes studies of participation patterns, performance differences in online vs. in-class testing, and rating scale evaluation. This line of scholarship comprises more than **25 publications**, including work in *Physical Review Physics Education Research*, *Sociology of Education*, and *Journal of Chemical Education*.

Large-Scale Assessment Infrastructure and the LASSO Platform
As founder and director of the LASSO platform, Van Dusen leads the design of a national research and assessment system used across 150+ institutions. His scholarship documents the platform's development and use, including early analyses of learning outcomes (*LASSO study initial findings*, Van Dusen et al., 2015; *Student Outcomes Across Collaborative-Learning Environments*, Herrera et al., 2018) and studies comparing online and paper assessments (Nissen et al., 2018; Nissen et al., 2021). These works establish the foundation for scalable, instructor-friendly formative assessment systems and represent **15+ publications** tied directly to LASSO-enabled research.

Learning Environments, Instructional Models, and Student Outcomes
A major theme of Van Dusen's work examines learning in collaborative and interactive STEM classrooms. Highly cited examples include *Associations Between Learning Assistants and Passing Introductory Physics* (Van Dusen & Nissen, 2020), *Tools for identifying courses that support development of expertlike physics attitudes* (Van Dusen et. al, 2021), and analyses of learning assistant programs across large datasets (e.g., White, Van Dusen, & Roualdes, 2016; Van Dusen & Nissen, 2018). This body of research includes more than **20 publications** and has influenced national conversations around course transformation and student outcomes.

Diagnostic Modeling, Computer-Adaptive Testing, and AI-Supported Assessment
Recent scholarship extends Van Dusen's assessment work into diagnostic modeling and AI-supported tools. Notable publications include *Applying Cognitive Diagnostic Models to Mechanics Concept Inventories* (Le et al., 2025) and *Two-Phase Content-Balancing CD-CAT Calibration* (Huang et al., 2025). His work increasingly integrates assessment with generative AI, including projects on adaptive tutoring, retrieval-augmented generation, and cognitive diagnostic AI. This growing line of work includes **10+ publications**, many co-authored with collaborators across physics, engineering, and computer science.

Research Methodology and Interdisciplinary Approaches
Van Dusen has authored several papers advancing methodological rigor and analytical techniques. These contributions include *Comparison of Normalized Gain and Cohen's d* (Nissen et al., 2018), *Investigating Society's Educational Debts in Physics Attitudes* (Nissen, Her Many Horses, & Van Dusen, 2021), and broader methodological reflections such as *Quantitative Rigor Through Critical Consciousness* (Van Dusen & Nissen, 2025). Collectively, these works support researchers in using appropriate analytic tools and interpreting STEM-education data responsibly.

William R. Gray-Roncal, Ph.D.

Principal Professional Staff, Johns Hopkins University Applied Physics Laboratory
Visiting Assistant Professor, Johns Hopkins University School of Education
11100 Johns Hopkins Road, Laurel, MD 20723 • wgr@jhu.edu

PROFESSIONAL PREPARATION

Ph.D., Computer Science, Johns Hopkins University, 2016. Dissertation: *Enabling Scalable Neurocartography: Images to Graphs for Discovery* (Advisor: G. D. Hager).

Post-Master's Certificate, Electrical Engineering, Johns Hopkins University, 2010.

M.S., Electrical Engineering, University of Southern California, 2005.

B.E., Electrical Engineering (Mathematics minor), Vanderbilt University, 2003 — *Magna Cum Laude*.

PROFESSIONAL & ACADEMIC APPOINTMENTS

2007–Present	**Research Engineer & Project Manager**, Johns Hopkins University Applied Physics Laboratory, Asymmetric Operations Sector, Decision Systems Group, Laurel, MD. Currently **Principal Professional Staff** (highest technical-staff grade); joined in 2007 as Associate Staff and advanced through the staff ranks. Lead scalable data systems, AI, and human–machine teaming programs across defense, biomedical, and learning systems.
2026–Present	**Visiting Assistant Professor**, School of Education, Johns Hopkins University. Leading development of a next-generation learning engineering concentration within the M.Ed. in Learning Design and Technology program. *(Previously Assistant Research Professor, courtesy, School of Education, 2020–2025.)*
2017–Present	**Assistant Research Professor** (courtesy), Department of Computer Science, Johns Hopkins University.
2018–Present	**Lecturer**, Engineering for Professionals (Applied Biomedical Engineering; Lifelong Learning), Johns Hopkins University.
2017–2023	**Founder & Lead**, CIRCUIT Program, JHU Applied Physics Laboratory — cohort-based undergraduate research and mentoring program serving 220+ underserved and underrepresented STEM students embedded in sponsored research.
2009–2021	**Co-Founder & Co-Director**, College Prep Program at APL — volunteer postsecondary-readiness program serving 220+ students with >90% on track for four-year degrees.
2010–2017	**Research Scientist**, NeuroData / Open Connectome Project, Johns Hopkins University.
2001–2007	**Electrical & Systems Engineer**, Northrop Grumman Corporation (Space Technology; Information Technology).

TEACHING & AI-ENABLED INSTRUCTION

2022–Present	**Agentic AI Systems**, JHU Engineering for Professionals. Master class delivered at scale to ~2,000 learners; traces the evolution from classical computer vision to modern multimodal agent frameworks, with applied, hands-on AI practice — a working demonstration of AI-enabled instruction at scale.
2026–Present	**Learning Engineering concentration** (in development), JHU School of Education. Designing next-generation learning-engineering curriculum within the M.Ed. in Learning Design and Technology.

2018–Present **Frontiers in Neuroinformatics**, JHU Engineering for Professionals. Graduate survey of advanced computational neuroscience.

2015–2022 **Introduction to Connectomics** (Intersession), Johns Hopkins University. Research-based course with student poster session; consistently among the highest-rated courses in the Whiting School of Engineering.

2011–Present **Research mentorship**, JHU and APL. Mentored ~500 students at the high-school, undergraduate, and graduate levels.

Convening: Co-organizer, AI Synergy Summit (JHU, 2025), on integrating generative AI and human-learning approaches; presenter, "Enhanced Learning Environment for Accelerating Training and Engagement" (Warfighter Training Modernization Conference, 2021).

PRODUCTS MOST CLOSELY RELATED TO THE PROPOSED PROJECT

1. **TeachMe** (significant product; learning-engineering prototype, 2026). AI-powered system applying learning science to support the training and operation of additive manufacturing; developed for the JHU–APL learning engineering collaboration.
2. **CIRCUIT Program** (significant product; founder & lead, 2017–2023). National cohort-based model integrating mentorship, research, and competency development; 220+ students served with longitudinal tracking of progression and placement. jhuapl.edu/circuit
3. **Agentic AI Systems** (significant product; instructional). Graduate / professional master class reaching ~2,000 learners through JHU Engineering for Professionals — AI-enabled postsecondary instruction delivered and refined at scale.
4. L. Huynh, K. Gray-Roncal, M. Roncal, and W. Gray-Roncal. "An accessible, distributed, technology-based approach for student and mentor engagement." *IEEE Integrated STEM Education Conference (ISEC)*, 2019.
5. H. P. Cowley, M. Natter, K. Gray-Roncal, et al., and W. Gray-Roncal. "A framework for rigorous evaluation of human performance in human and machine learning comparison studies." *Scientific Reports* 12(1):5444, 2022.

ADDITIONAL SIGNIFICANT PRODUCTS

1. M. Cervantes, S. Floryanzia, J. Sharp, E. C. Johnson, and W. Gray-Roncal. "Empowering trailblazers toward scalable, systematized, research-based workforce development." *ASEE Annual Conference & Exposition*, 2023.
2. E. C. Johnson, M. Villafañe-Delgado, et al., and W. R. Gray-Roncal. "An immersive curriculum to develop computational science and research skills in a cohort-based internship program." *IEEE ISEC*, 2023.
3. S. Floryanzia, A. Jayabharathi, K.-A. Carr, K. Gray-Roncal, and W. Gray-Roncal. "Toward a framework for providing equitable opportunities to activate STEM talent." *IEEE ISEC*, 2024.
4. K. Gray-Roncal, L. Huynh, T. Kolarik, et al., and W. Gray-Roncal. "The College Prep Program at APL: an experiential model to help high-achieving, underserved students trailblaze and achieve success." *IEEE ISEC*, 2018.
5. J. T. Vogelstein, W. G. Roncal, E. Perlman, B. Wester, et al., and R. Burns. "A community-developed open-source computational ecosystem for big neuro data." *Nature Methods*, 2018.
6. R. Hider, D. Kleissas, et al., W. Gray-Roncal, and B. Wester. "The Brain Observatory Storage Service and Database (BossDB): a cloud-native approach for petascale neuroscience discovery." *Frontiers in Neuroinformatics* 16:828787, 2022.
7. N. Kasthuri, K. J. Hayworth, et al., W. Gray Roncal, R. Burns, and J. W. Lichtman. "Saturated reconstruction of a volume of neocortex." *Cell* 162(3):648–661, 2015.

SELECTED SPONSORED RESEARCH

Sustained principal-investigator and technical-leadership experience managing federally funded programs (NIH, DARPA, IARPA, DoD) centered on scalable AI/data systems, rigorous evaluation, and technical talent development:

Co-Principal Investigator — NIH BRAIN CONNECTS: Center for High-Throughput Integrative Mouse Connectomics (2023–2028). Scalable circuit reconstruction and analysis.

Key Personnel — NIH bossDB, R24 (2018–2028). Community data infrastructure and storage paradigms for high-resolution connectomics.

Principal Investigator — BENCHMARK, NIH R01 (2021–2023). Community standards for high-resolution connectomics.

Principal Investigator — TITAN, DARPA AIE (2020–2021). Low-SWaP, robust properties of micro-insect brains to advance biological intelligence.

Technical Lead — IARPA MICrONS (2017–2022). Machine-learning algorithms for cubic-millimeter mammalian neuroscience.

Principal Investigator — SABER, NIH (2017–2020). Reproducible cloud-based processing pipelines for neuroscience.

Team Lead — CDAO Digital Talent Management Futures (2024). Diverse technical talent pipeline for the Department of Defense.

AWARDS, HONORS & FELLOWSHIPS

2024 Nexus Convening Award, Johns Hopkins University.

2021 Light the FUSE Award, JHU Applied Physics Laboratory.

2018 REDx Imagine Award (AI and connectomics), JHU Applied Physics Laboratory.

2015 Author's First Paper Publication Award, JHU Applied Physics Laboratory.

2014 Hart Prize for Best Research Project, JHU Applied Physics Laboratory.

2009 Diversity Leadership Award, Johns Hopkins University.

Fellowships Full-tuition Ph.D. Fellowship (APL); Harold Stirling Vanderbilt Scholarship, full tuition (Vanderbilt).

Honor societies Tau Beta Pi; Eta Kappa Nu.

RELEVANT EXPERIENCE & QUALIFICATIONS

For nearly two decades I have worked at the intersection of artificial intelligence and human learning — building scalable data and AI systems on one side, and designing, delivering, and *measuring* student-success interventions on the other. As a federally funded principal investigator across NIH, DARPA, and IARPA, I have led programs demanding exactly the capabilities this project requires: rigorous instrumentation, scalable infrastructure, and disciplined evaluation of human and machine performance. As founder of CIRCUIT and co-founder of the College Prep Program at APL, I built cohort-based models that served hundreds of students with documented progression outcomes, giving me direct experience translating educational goals into measurable results for underserved learners.

My current academic work places me squarely in this project's domain. At the JHU School of Education I am developing a next-generation learning engineering concentration; I teach an AI systems master class reaching ~2,000 learners; and I co-organized the 2025 AI Synergy Summit on integrating generative AI with human learning — practical experience deploying and convening AI in postsecondary education. My published work on the rigorous evaluation of human performance in human–machine comparison studies provides the methodological grounding to determine whether an AI intervention genuinely improves student performance rather than merely appearing to. I am well suited to lead the AI design, learning-engineering, and measurement components of this effort and to keep its claims evidence-based.

QUALITY MEASURES
131 Hanbury Road, West, Suite C1
Chesapeake, VA 23322-4379

June 29, 2026

Odis Johnson Jr., PhD
Johns Hopkins University
2800 N. Charles St.
Baltimore, MD 21218

Dear Dr. Johnson:

If the proposal submitted by you, and your Co-PIs titled "*Helping Students Stay in STEM: A Personalized AI Companion for Gateway Courses*" is selected for funding by the U.S. Department of Labor's Fund for the Improvement of Postsecondary Education – Postsecondary Student Success Program (FIPSE-PSSG) program, it is my intent that our organization, Quality Measures LLC, will collaborate as specified in the proposal as the external evaluators.

Sincerely,

Dr. Gwen Lee-Thomas, CEO

Project Narrative File(s)

*** Mandatory Project Narrative File Filename:** FIPSE Final 6.29.pdf

Add Mandatory Project Narrative File | Delete Mandatory Project Narrative File | View Mandatory Project Narrative File

To add more Project Narrative File attachments, please use the attachment buttons below.

Add Optional Project Narrative File | Delete Optional Project Narrative File | View Optional Project Narrative File

Table of Contents

SIGNIFICANCE

This Postsecondary Student Success Grant (PSSG) proposal, ***Helping Students Stay in STEM: A Personalized AI Companion for Gateway Courses*** (hereafter, "AI Companion Project") advances Evidence Absolute Priority 1 (AP1) in early phase research with a Project Pathway Absolute Priority 3 (AP3) in artificial intelligence (AI) by testing whether AI-assisted learning will reduce course non-completion for students in physics, calculus, and chemistry. With the science, technology, engineering, and mathematics (STEM) workforce projected to grow by 8.1 percent through 2034 (NSB, 2020), outpacing growth in non-STEM occupations, university gateway STEM courses offer a critical opportunity to address what former NSB Chair Gil Dario called "the domestic STEM talent crisis" (McKenzie, 2024). Gateway STEM courses such as introductory physics, calculus, and chemistry are well-documented choke points in the undergraduate pipeline, where nationally fewer than half of first-year undergraduate students who start a STEM program obtain a bachelor's degree in STEM six years later (Eagan et al., 2014; Leoni et al., 2023). Many capable students leave STEM fields, not only because of content difficulty, but because they struggle to adapt to fast pacing, impersonal large lectures, and competitive grading (Singer et al., 2012). Optimizing the nation's STEM workforce potential and improving the completion, earnings, and economic mobility of students depends on instructional innovations that address these student experiences, which are not only typical of university gateway courses, but also community college, and alternative (Jabbari et al., 2023; Huang et al., 2025) workforce credential programs.

Traditional high-touch supports such as peer-led team learning and small-group tutoring can improve performance, self-efficacy, and sense of belonging, but they are resource-intensive and difficult to scale to all learners in all course sections (Barrasso & Spilios, 2021; Chan & Bauer, 2015; Smith et al., 2014; Trujillo & Tanner, 2014) and program types (e.g. credential programs). This leaves a structural gap: many students face demanding courses with little timely, personalized formative feedback on their reasoning and strategies for how to study, persist, and recover from setbacks. This gap creates a clear opportunity for a high-impact technological intervention with well-designed systems that can replicate the personalized, scaffolded support of a human tutor at scale to support the self-regulation of learning (SRL) (Zimmerman, 2002; Schunk & Zimmerman, 2013). The *STEM-AI Companion* project will design, implement, and rigorously evaluate a personalized artificial intelligence (AI) learning companion for undergraduate STEM students in high-enrollment gateway courses (i.e., physics, chemistry, and calculus). The Companion will offer students personalized and differentiated instructional support as they master vital foundational concepts and procedures for a variety of STEM pathways into the workforce and success.

Improving STEM Gateways: High Impact Touchpoints Across College Types

We focus on gateway courses for several reasons. First, **these are high impact touchpoints within a critical education to career pathway**. While STEM degree completion rates within 2- and 4-year institutions remain similar to non-STEM degree programs, student transfers out of STEM degree programs early in their matriculation remain the highest (Chen, 2013; National Science Board, 2018; National Science Foundation, 2021). Often, students' decisions to withdraw from gateway courses are driven by factors unrelated to their academic preparation, among them, their motivation, self-efficacy, and perceived lack of support (Oliveira et al., 2025). Taken together, these factors highlight that students' capacity to plan, monitor, and adapt their learning (that is, their self-regulation), as well as institutional conditions, should be priority targets for research and evidence-based change (Winne, 2021). Second, **the fast pace of AI innovation has led to the uneven adoption of learning technologies across postsecondary institutions**, often with little guidance on how to use AI in classrooms, awareness of instructors' preparedness to use it effectively, and ability to determine if AI functions as planned (U.S. Dept. of Education, 2023). Third and finally, **the societal implications of STEM degree non-completion associated with gateway course challenges are serious, including the erosion of the nation's international competitiveness, national security, domestic innovation in AI (NASEM, 2024)**, **and the foregone possibility of economic mobility from STEM degrees' higher wages**. These national imperatives may be addressed if research establishes the effectiveness and reliability of AI instructional technologies to support the learning of all populations, including all students that are currently underserved across 2- and 4-year postsecondary institutions.

General Education Provision Act (GEPA) Statement: The project will ensure equitable access to high-quality postsecondary instructional supports and opportunities for all eligible students, including those who face barriers to persistence and completion. Activities will be designed to be accessible, practical, and responsive to student needs, with appropriate accommodations, clear communication, and targeted outreach to students who are most in need of support. University policies and practices also support non-discriminatory engagement with students, workplace professionals, and all participating in this research. The project will focus on measurable student success outcomes, efficient use of resources, and implementation strategies that expand access to education-to-employment pathways and strengthen student completion, transfer, and workforce readiness. Nonetheless, potential barriers to full participation may arise, including limited educator capacity, differential access to technology, time availability, and differences in familiarity with technology. To address these common barriers, this team will 1) coordinate technology access with participating university sites as feasible, 2) provide professional learning and coaching for instructors, 3) compensate instructors and students for the time they give to this project with survey and interview incentives, and up to $4000 to instructors for training and intervention implementation, and 4) use multiple communication methods to engage students and instructors to ensure their informed participation.

Evidence Supporting Assessment-Driven Learning and Personalized Tutoring

The proposed project builds on three evidence-supported approaches to improving student success: formative assessment and feedback, cognitive diagnostic assessment, and personalized tutoring. Decades of research demonstrate that formative assessment improves student learning by providing timely information about student understanding and opportunities for instructional adjustment. Black and Wiliam (1998) identified formative assessment as one of the most effective mechanisms for improving achievement, while Hattie and Timperley (2007) demonstrated that high-quality feedback helps students identify misconceptions and make progress toward learning goals. Freeman et al. (2014) found that active-learning environments incorporating frequent formative assessment significantly improve student performance and reduce failure rates in undergraduate STEM courses.

The project also builds on growing evidence supporting cognitive diagnostic assessment and mastery-based feedback. Unlike traditional assessments that provide a single summary score, cognitive diagnostic approaches generate detailed information about students' mastery of specific learning objectives and underlying skills. Tang and Zhan (2021) found that students receiving cognitive-diagnostic feedback outperformed students receiving traditional correct/incorrect feedback, particularly on difficult content. Similarly, Maas et al. (2022) demonstrated that cognitive diagnostic assessments can reliably classify student mastery and support effective learning decisions through learning-objective dashboards. Emerging work further suggests that the effectiveness of learning-objective dashboards depends not only on the quality of the underlying assessment models, but also on how information is visually represented and interpreted by educators (Salazar Morales et al., under review). These findings suggest that mastery-based feedback can help students and instructors identify specific conceptual barriers and target support more effectively.

A substantial body of evidence also supports personalized tutoring and AI-supported learning. Nickow et al. (2020) reported significant achievement gains across 96 randomized tutoring studies, while Ma et al. (2014) found that intelligent tutoring systems outperform traditional instruction and can approach the effectiveness of one-on-one human tutoring. Recent work by Kestin et al. (2025) further suggests that generative AI tutors can produce meaningful improvements in learning and engagement when designed around sound instructional principles. Together, these findings provide a strong rationale for integrating formative assessment, cognitive diagnostic feedback, and AI-supported tutoring within a unified instructional infrastructure designed to improve learning, persistence, and success in gateway STEM courses.

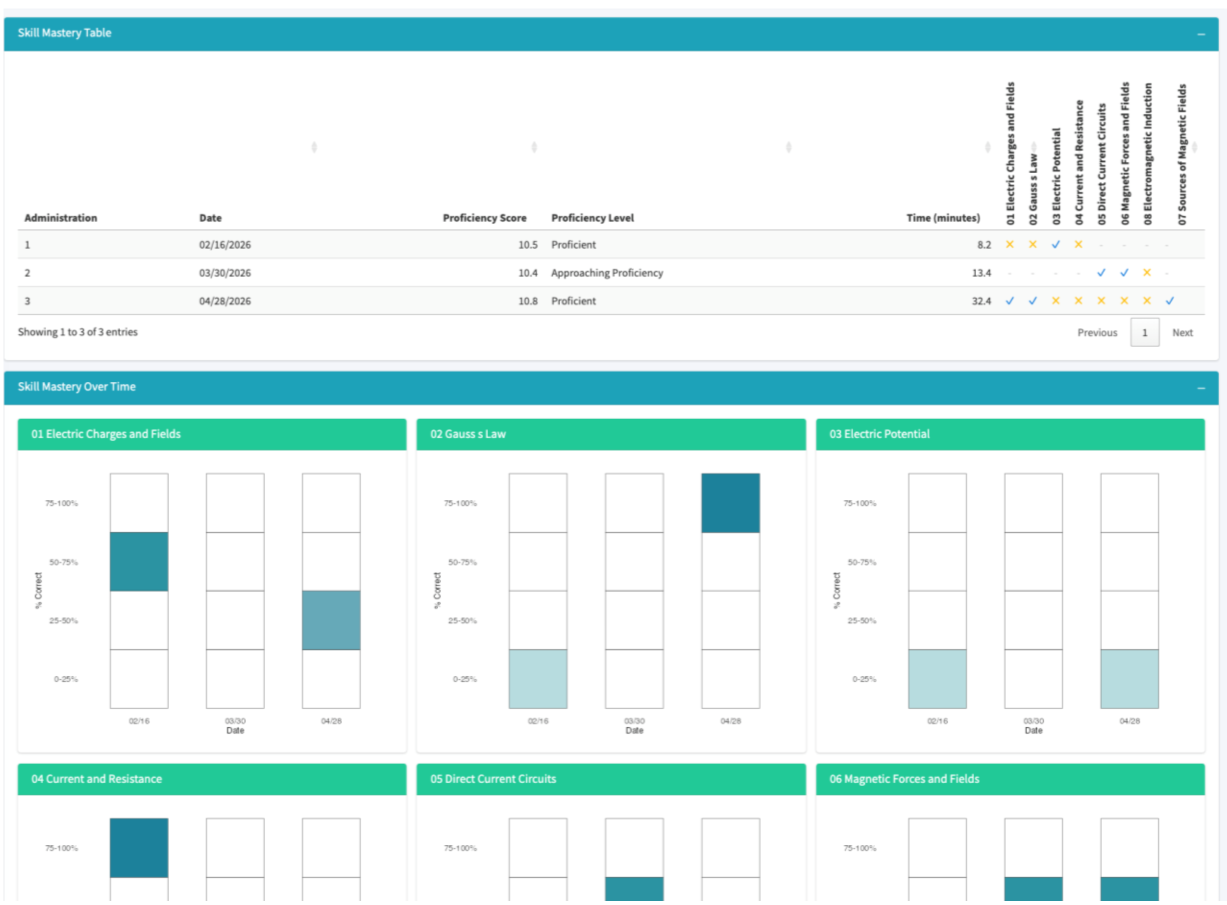
Improving Student Learning and Achievement in STEM and Beyond

The STEM-AI Companion project offers a transformative solution to these challenges by embedding a personalized, course-integrated AI tutor (Mousavinasab et al., 2021; Van Lehn, 2011; Ritter et al., 2007) directly within an established assessment platform—LASSO (LASSO, 2025; Nissen et al., 2018; Van Dusen, 2018; Van Dusen et al., 2021). **Unlike traditional tutoring or generic AI tools, the STEM-AI Companion is designed to scaffold reasoning, guide solution planning, and prompt reflection,** all while integrating domain-specific prompt optimization, retrieval from curated instructional resources, and adaptive reasoning strategies to personalize support based on students' demonstrated mastery (Zhao et al,, 2025). These capabilities are intended to strengthen the instructional mechanisms that prior research has shown to improve learning—including timely formative feedback, personalized scaffolding, and adaptive tutoring (Black & Wiliam, 1998; Hattie & Timperley, 2007; Ma et al., 2014; Nickow et al., 2020). Thus, the proposed intervention builds on established evidence regarding effective instructional practices while incorporating recent advances in trustworthy and adaptive AI, providing a clear rationale for expecting the personalized AI Companion to improve student learning, self-regulated learning (Zimmerman, 2002; Schunk & Zimmerman, 2013), and gateway-course success beyond what can be achieved using LASSO alone or LASSO paired with a general-purpose AI tool**.**

LASSO is an AI-based platform in which instructors create online STEM assessments, and data analysis is automated to generate interactive reports about student skills, mastery, and trends throughout the course (Le et al., 2025). Its adaptive support is grounded in validated diagnostic evidence (students' performance estimates and skill-mastery profiles generated through LASSO's cognitive diagnostic computer adaptive assessments) ensuring that personalization is driven by what students demonstrably know and need rather than generic model behavior. Its adaptive support is informed by both curated content and student engagement data, ensuring that assistance is relevant, timely, and effective. This approach leverages cutting-edge technology to provide real-time, individualized support for students as they engage with challenging material. Since LASSO is free for instructors' use, it provides an ideal cost effective and scalable platform for the distribution of AI assessment tools, and collection of data that will be generated during AI Companion implementation. Both the AI Companion and LASSO were developed by co-PIs of this project using federal funding streams.

Current deployments include the Mechanics Cognitive Diagnostic (MCD), a CD-CAT assessment spanning 20 learning objectives in introductory physics (Le et al., 2024) that has been developed and validated to provide fine-grained information about students' mastery of foundational physics concepts and procedures (Roberge et., 2025). A complementary calculus version is currently under development. Figure 1 illustrates an example of a student cognitive

Figure 1. Example of LASSO Student Cognitive Diagnostic Report



diagnostic report generated through the CD-CAT system, including proficiency estimates, learning-objective-level mastery information, and changes in mastery across test administrations.

Although CD-CAT is not the core innovation of the present project, it provides the high-resolution formative data (ability estimates and skill-mastery vectors) that make personalized AI tutoring feasible, enabling the Companion to detect specific areas of struggle, scaffold reasoning, and tailor instructional support throughout the course. As shown in Figure 1, the report provides proficiency estimates, learning-objective-level mastery information, and changes in mastery across administrations. These diagnostic profiles identify strengths and weaknesses at the skill level and provide the foundation for personalized instructional recommendations and AI-supported tutoring.

This project will rigorously compare three models of support: students who utilize the LASSO platform, those who use LASSO with a non-personalized AI companion of their choice (e.g. ChatGPT, Claude, Copilot, Grok, or Meta AI), and students using LASSO and the personalized AI Companion. By evaluating these approaches across institutions, courses, and subjects, the project will determine whether curated, in-situ personalized AI can match or surpass the effectiveness of "platform only" or "non-personalized AI" models in improving mastery, reasoning, course completion, and progression. We propose a three-phased project, with Year 1 comprising design-based research of the AI Companion tool, a pilot evaluation in year 2 and a mixed method stratified block randomization across years 3 and 4. With this design we inform the **theory of change** that the STEM-AI Companion will serve as an AI co-regulator of students' self-regulated learning (SRL) using its hybrid intelligence and scaffolding adaptivity to achieve personalization of the student learning experience. This will support students in planning, monitoring, and reflecting on learning tasks without doing the work for students, and at times, provide opportunities for students to critique the quality of AI generated responses. The outcome will be increased conceptual understanding of the STEM gateway course content, stronger senses of subject area self-efficacy, and enhanced AI literacy in learning, and improvements in the distal outcomes of course completion, course repetition, grades, and changes in major.

The benefits of this approach are numerous. First, the AI Companion is designed to scale and pair with a free access platform (LASSO: www.lassoeducation.org), enabling institutions—including those with limited resources—to extend high-impact support to more students without proportional increases in staffing. Costs and limited reach hinder the usefulness of traditional academic supports, particularly for less resourced institutions (Barrasso & Spilios, 2021; Chan & Bauer, 2015; Smith et al., 2014; Trujillo & Tanner, 2014). By integrating individualized scaffolding directly into LASSO, the STEM-AI Companion reduces barriers to access and expands the availability of evidence-based instructional practices. This aligns with federal priorities to advance artificial intelligence in education and prepare U.S. students for the demands of the modern workforce.

Second, we anticipate a substantial impact: Over the course of the project, we will engage 48-54 instructors at 16 – 18 institutions, including both two-year and four-year colleges, and reach a minimum of 4580 students. Our sample recruitment approach benefits from a partnership with the National Science Foundation ECR Hub and the Learning Assistance Alliance, a STEM focused instructional design network. Outcomes will include improved course completion rates, increased enrollment in subsequent STEM courses, and measurable gains in mastery and reasoning, especially for underserved students. For faculty, this project will provide training for LASSO use, guidance for AI implementation, and through the Companion, actionable insights into student learning, enabling targeted instructional improvements and more effective teaching practices. Consistent with prior research on graph literacy, graph comprehension, and learning analytics dashboards (Friel et al., 2001; Oslund et al., 2021; Salazar Morales et al., under review), the project emphasizes presenting assessment information in forms that support interpretation and instructional decision making.

A third contribution of our approach is that is provides rigorous causal evidence by considering multiple instructional experiences, employing randomization to address troubling selection effects, and allowing each participant to experience all three treatments in a sequenced fashion. This research approach is supported with administrative data, instructor surveys and student interviews, and the cataloguing of user and AI generated content to permit the analysis of scaffolding and discourse for a greater understanding of AI effectiveness. This study will therefore answer longstanding questions and offer new knowledge about the interface of AI tools with postsecondary institutions, instructors, and all students, including those underserved.

PROJECT DESIGN

Theory of Action and Logic Model

The project's theory of action begins from a well-documented barrier: gateway STEM courses—introductory physics, calculus, and chemistry—are where many students leave the STEM pipeline, and failure or withdrawal in these courses falls most heavily on at-risk students (Pell-eligible, first-generation, and community-college students), who have the least access to personalized academic support. The primary constraint is rarely raw student ability. Instead, it is the absence of timely, individualized feedback and the difficulty of remaining organized and self-directed in large, fast-paced courses. The STEM-AI Companion addresses this barrier as an instructor-overseen student-coordinated AI tutoring intervention embedded in the LASSO platform. It provides each student with personalized hints, questions, and feedback rather than standard answers, guided by skill-by-skill diagnostic results, and fades scaffolding support as competence grows, within responsible-use safeguards (human review, privacy and bias review, accessibility, and AI-literacy/fluency instruction).

Central hypothesis. *If* at-risk students in gateway STEM courses receive this personalized, instructor-overseen AI tutoring, *then* they will use AI more productively, become more self-directed, and understand course material more deeply, so that more of them pass these courses and accumulate credits, and more are retained and continue in STEM pathways. The mechanism proceeds in four linked steps. *First*, because the Companion offers personalized hints rather than standard answers, students learn to question, check, and reason with AI rather than copy from it, building AI literacy. *Second*, prompts to plan, self-check, and reflect strengthen self-directed learning. *Third*, personalized scaffolding that fades over time deepens conceptual understanding and problem-solving skills. *Fourth*, these gains raise the share of students who pass gateway courses and accumulate first-year credits (the earliest indicators of retention and completion), which institutional records then track through to retention, transfer, and completion, disaggregated by Pell status and inclusive of part-time and transfer-in students. Table 1 presents the full logic model.

Table 1. STEM-AI Companion logic model

Resources / Inputs >	Activities (Project Components) >	Outputs >	Short-Term Outcomes >	Long-Term Outcomes |
• **LASSO platform -** free; used at 300+ institutions; validated skill-by-skill diagnostic assessments in physics and calculus (chemistry version in development) • **STEM-AI tutor-** gives hints, questions, and	• **Identify each student's skill mastery profile** using cognitive diagnostic computer adaptive testing • **Provide personalized AI tutoring** in gateway physics, calculus, and chemistry (personalized	• **At-risk students served** in AI-supported sections: ~4580 students, 48-54 instructors, 16-18 institutions; each student's Pell, first-generation, part-time, transfer-in, and 2-year status recorded • Diagnostic assessment banks in three subjects;	*Causally tested in the stratified block randomization:* • Students use the AI tutor productively and build AI literacy • Students plan, self-check, and reflect (self-directed learning) • Students understand gateway content better and solve	*Required performance measures, reported by Pell status (baselines and targets set):* • **Annual retention and persistence** • **Graduation and upward transfer** (2-year institutions) • **Completion and credentials conferred**

feedback, rather than answers; draws on vetted course materials; instructor-overseen; protects privacy and exam integrity • **Team and independent evaluator**—JHU (lead), LASSO/Iowa State, Quality Measures LLC • **Recruiting networks**—NSF ECR Hub/BCSER and the Learning Assistance Alliance (626 institutions), reaching 2- and 4-year schools serving many Pell students • **Research base and safeguards**—evidence on feedback, diagnostic assessment, tutoring, and self-directed learning; IRB approval; FERPA-aligned data governance	hints and feedback) • **Build in responsible use**—instructor review, privacy and bias review, accessibility, AI-literacy instruction • **Train instructors** and help students learn to use AI well • **Develop and test in stages**—Year 1 design-based research, Year 2 multi-site pilot, Years 3–4 stratified block randomization, plus independent evaluation reported to ERIC • **Disseminate and scale** to 16-18 institutions	instructor dashboards • Longitudinal, fine-grained student assessment data. • AI tutoring sessions and assessments completed; student-AI interaction patterns revealed by analyzing randomly selected usage logs • Instructors trained; responsible-use procedures in place • Evaluation reports and publications; independent evaluation submitted to ERIC	problems more skillfully • **More students pass gateway courses** (fewer failures and withdrawals) • **Students accumulate more first-year credits** • Students continue to the next STEM course	• **Time to credential** *Ultimate impact:* • Broadened participation in STEM workforce • Stronger footing for STEM degrees, careers, and economic mobility

Consistent with an early-phase (Absolute Priority 1) design, the stratified block randomization causally tests the short-term outcomes (productive AI use, self-direction, conceptual mastery, gateway-course completion, credit accumulation, and continuation to the next course) with a focus on Pell students and part-time or transfer-in students. The project tracks and reports the long-term outcomes, disaggregated by Pell status or other non-full-time statuses with baselines and targets. Improved economic mobility is identified as the ultimate, longer-horizon impact, rather than a measured project outcome.

Each key project component in the logic model is informed by research suggesting it is likely to improve student retention and completion, satisfying the Absolute Priority 1 “demonstrates a rationale” standard (Table 2). These component-to-outcome links are elaborated at finer grain by the conjecture map in the Design-Based Research subsection, which connects specific Companion features to hypothesized self-regulation processes and proximal outcomes.

Table 2. Evidence supporting each key project component (Absolute Priority 1 rationale)

Project component	Research it builds on	Outcome it is expected to improve
AI tutoring that guides rather than answers	Nickow et al. (2020), 96 trials; Ma et al. (2014); Kestin et al. (2025)	More students pass gateway courses > retention and completion
Frequent feedback on student work	Black & Wiliam (1998); Hattie & Timperley (2007); Freeman et al. (2014)	Fewer failures and withdrawals > credit accumulation and retention
Skill-by-skill diagnostic feedback	Tang & Zhan (2021); Maas et al. (2022)	Stronger mastery > course completion and progression
Support that builds self-direction and fades over time	Zimmerman (2002); Schunk & Zimmerman (2013)	Persistence and completion, especially for at-risk students
Building AI literacy and responsible use	Nguyen & Barbieri (2025); Anders & Dux Speltz (2025)	Productive AI use > progression and retention

Core assumptions. The theory of action rests on three assumptions. First, instructors integrate and oversee the Companion within their courses, and recruited sites enroll sufficient Pell-eligible, first-generation, part-time, transfer-in, and 2-year students. Second, the diagnostic results are accurate enough to guide personalized support, and enough students use the tools as assigned to estimate credible effects. Third, gains in gateway-course completion and credit accumulation carry forward into retention, transfer, and completion, as tracked through institutional records.

Theoretical Framework: Self-Regulation, Co-Regulation, and AI for STEM Learning

We frame our proposed project using Zimmerman's social-cognitive model of self-regulated learning (SRL) (Zimmerman, 2002). SRL refers to the self-initiated thoughts, emotions, and actions that students use to plan, monitor, and adapt their approach to achieving academic goals, organized into three cyclical phases of forethought, performance, and self-reflection (Schunk & Zimmerman, 2013). Rather than a fixed trait, SRL comprises proactive, cyclical processes that can be improved by targeted instruction.

In gateway STEM courses, students struggle not only with regulating their learning but also with mastering critical core conceptual and procedural knowledge that impact future success in their chosen STEM pathways. From the project's outset, we will design the Companion to support both disciplinary learning and SRL with prompts, explanations, and representations tailored to common conceptual and procedural bottlenecks in physics, calculus, and chemistry. Consistent with contemporary models and empirical work in SRL and metacognition, we treat SRL processes and discipline-specific knowledge building as tightly coupled: planning, monitoring, and reflection are designed to occur in relation to concrete domain-specific tasks (Greene & Azevedo, 2007; Panadero, 2017).

Many students enter with weak SRL, setting vague goals rather than process goals, relying on passive strategies, struggling to monitor whether they truly understand new problems, and often interpreting setbacks as evidence of fixed inability, rather than improvable skills (Dai & Cromley, 2014; Limeri et al., 2023; Sebesta & Speth, 2017). These gaps in students' self-regulation, rather than differences in raw ability, are key drivers of attrition (Park et al., 2019). Effective interventions must therefore support not just "more and better explanations," but better planning, monitoring, and reflection around conceptual and procedural problem solving.

Generative AI offers a novel way to provide this kind of support at scale, but only if it is designed as an *augmenting co-regulator* rather than an answer engine (Nguyen & Barbieri, 2025). By “augmenting co-regulator,” we mean that the AI acts as an additional mentoring layer that scaffolds students’ self-regulation (i.e., setting goals, monitoring progress, and reflecting) while leaving high-stakes evaluation and relational support to human instructors. Drawing on Nguyen and Barbieri’s (Nguyen & Barbieri, 2025) *Mentoring and SRL Pyramid Model*, the AI supplements, rather than replaces, human mentoring by filling feedback gaps and supporting co-regulated learning on the way to greater student autonomy.

Traditional intelligent tutoring systems typically could only offer fixed hints in narrow domains, whereas large language models can engage in flexible, dialogic tutoring specific to the individual learner (Letourneau et al., 2025; Ma et al., 2014). They also introduce new risks, however, including producing hallucinations and facilitating negative aspects of cognitive offloading, where learners outsource thinking and bypass the productive struggle required for learning (Rohilla, 2025). We therefore adopt a “constraint-based co-regulation (Nguyen, & Barbieri, 2025) approach in which the AI is explicitly designed to withhold direct answers, emphasize hints, questions, and strategy prompts, and structure interactions so that students remain cognitively engaged rather than passively consuming solutions.

To operationalize our approach, we draw on work in hybrid human-AI systems and shared regulation (Holstein et al., 2020). In this view, the STEM-AI Companion, instructors, and students together form a shared “seeing and doing” system: the Companion helps notice patterns in students’ language and problem-solving (for example, premature help-seeking), interprets these through an SRL lens (such as weak planning or maladaptive attributions), and then responds with regulatory approaches, not just content explanations, that are grounded in the disciplinary concepts and procedures students are working on. At *forethought,* it can prompt students to set specific goals and select active strategies before they begin; during *performance,* it can ask them to explain their reasoning, highlight recurring error patterns, and suggest alternative approaches when they are stuck; and in *self-reflection,* it can guide brief, structured debriefs that emphasize controllable causes and concrete strategy changes for next time. The primary aim is to move students, over time, from the system doing much of this regulatory work, to shared regulation between student and Companion, and ultimately toward more fully self-regulated learning and stronger mastery of gateway STEM content.

A complementary outcome of students becoming more self-regulated learners with AI is that it directly addresses calls for enhanced AI literacies in education and aligning with Priority 1(a)'s focus on responsible AI use. SRL and co-regulation skills are essential to engaging AI as a partner with an ability to make errors, rather than a shortcut (Anders & Dux Speltz, 2025).

Scaffolding and Learner Feedback as Core Mechanisms

A final piece of the framework concerns scaffolding and fading. From a Vygotskian perspective, the AI initially acts as an always-available "more knowledgeable other" that helps students work in their zone of proximal development (Cai et al., 2024). But constant, fully formed solutions risk impeding development, where learners feel fluent but fail to learn. To avoid this outcome, we will design the Companion to provide adaptive scaffolding that is contingent on student need and gradually faded as competence grows. For example, as student competency expands, the AI moves from fully worked examples to targeted hints and Socratic questions to high-level checks only. This fading, combined with interaction designs that require students to analyze, explain, and create, is central to how we expect the intervention to strengthen problem-solving skill and SRL, rather than undermine it.

The core of our approach to learner support with the Companion, therefore, is based on two well-established mechanisms for supporting learning: scaffolding and feedback. By *scaffolding*, we mean temporary structures, provided by more knowledgeable others, that guide how students approach and reflect on problems in ways that are gradually withdrawn as competence grows (Wood et al., 1976). By *feedback*, we mean timely information about the quality of students' thinking and performance that helps them close the gap between current and desired understanding (Hattie, & Tuimperley, 2007). Decades of research on tutoring, formative assessment, and problem-solving instruction suggest that well-timed scaffolds and feedback, especially when focused on process and strategy rather than answers alone, are among the most effective levers for improving conceptual understanding and problem solving in STEM (Black & Wiliam, 1998; Singer et al., 2012; VanLehn, 2011). In large, fast-paced gateway STEM courses, instructors may face structural constraints that make it difficult to provide frequent, individualized feedback on students' problem-solving processes, and students who most need guidance are often the least likely to receive it. As a result, many students have limited opportunities to get timely, targeted support and formative feedback on how to engage with complex problems and learn from mistakes as they work through them (Schaffer et al., 2017).

The use of generative AI to support instruction and tutoring raises an important design question (Steinert et al., 2024): Can an AI system deliver scaffolding and feedback that are not only frequent and personalized, but also instructionally sound, aligned with course goals, supportive of self-regulation, and conducive to durable learning rather than shortcutting it? **Our design treats scaffolding and feedback delivered through AI-mediated formative assessment as the primary levers through which the Companion is expected to influence both SRL processes and core STEM learning outcomes.** In doing so, the project will generate and test concrete design principles for constraint-based, SRL-aware AI tutoring, thereby advancing how educational AI systems themselves are built to support learning.

The conjecture map in Table 4 on page 22 makes these links explicit by connecting specific features of the Companion to hypothesized changes in students' planning, monitoring, and reflection, and in turn to gains in disciplinary conceptual understanding, problem solving, and persistence in STEM pathways. Year 1 design-based research (DBR), described below, will therefore focus not only on whether students use these AI-provided supports, but on which patterns and qualities of scaffolding and feedback promote deeper learning and self-regulation in our targeted courses, as well as on the factors that influence them.

In sum, the theoretical framework positions the STEM-AI Companion as a hybrid human–AI co-regulator grounded in SRL, co-regulation, and hybrid intelligence. It explains why gateway courses are such a risk point (namely, fragile SRL under high pressure), how a generative AI system can be designed to support planning, monitoring, and reflection without doing the work for students, and why adaptive, fading scaffolds are needed to prevent dependency. **This framework informs our core research questions:** 1) How does personalized, in-situ AI scaffolding from the STEM-AI Companion relate to undergraduate students' learning outcomes and retention (e.g. course, program) in physics, calculus, and chemistry in 2- and 4-year institutions?; 2) In both 2- and 4-year institutions, how do patterns of student interaction with a personalized AI Companion differ from interactions with general purpose AI?; 3) How much awareness of AI literacy related to responsible use and trustworthiness can be gained through engagement with personalized AI companions?; and, 4) To inform implementation and scaling, what individual, platform, or institutional factors inform the effectiveness of the AI Companion relative to other AI and non-AI options?

Design Principles for the STEM-AI Companion

Guided by the theoretical framework, in this section we articulate five core principles for the Companion. These principles translate self-regulation, co-regulation, and hybrid human-AI ideas into concrete expectations for how the Companion should behave in gateway STEM courses in undergraduate settings. The principles align with Absolute Priority 1 by emphasizing personalized support, high-impact tutoring, and responsible, human-centered use of AI.

Principle 1: AI as an Augmenting Co-Regulator, Not an Answer Engine

As described above, the Companion will function as an *augmenting co-regulator* of learning (Nguyen & Barbieri, 2025), not a replacement for human teaching or a shortcut to answers. In practice, this means the Companion's role is to help students formulate goals, select productive strategies for solving core conceptual and procedural problems in their STEM courses, notice when their understanding is fragile or unclear, and reflect on what to change next time. Its prime directive is to be constrained from freely giving answers and instead support students in doing the thinking themselves, always in ways that keep students cognitively engaged, rather than sidelined.

Principle 2: Constraint-Based Co-Regulation to Keep Students Doing the Work

Building on this constraint-based co-regulation approach, we will explicitly constrain the Companion to keep students doing meaningful cognitive work. Consistent with emerging AI-in-education scholarship on human-centered generative AI literacies, we will design the Companion to default to hints, questions, partial explanations, worked examples, and strategy prompts, rather than fully formulated solutions. Direct answers will be withheld or delayed until after a student has attempted a problem, articulated a plan, or explained their reasoning. These constraints are intended to preserve productive struggles with key conceptual and procedural steps, reduce over-reliance on AI, and support the development of SRL.

Principle 3: Alignment to SRL Phases Across the Learning Cycle

Following Zimmerman's (Zimmerman, 2002) cyclical model of SRL, the Companion will structure its interactions with students around the following three phases. **Forethought:** Before a study session or problem set, the Companion will prompt students to set specific, process-focused goals (e.g., mastering free-body diagrams) and to choose at least one active strategy (e.g., comparing solutions to a worked example and explaining each step aloud). **Performance:** While students are working, the Companion will encourage the use of cognitive, metacognitive, and resource management strategies (Xu et al., 2023). It will also provide metacognitive monitoring and domain-accurate explanations and representations that target known conceptual bottlenecks by asking students to explain their steps, justify their choices, and consider alternative approaches when they are stuck. **Self-reflection:** Following practice and formative assessment sessions, or other low-stakes assessments, the Companion will guide a brief debrief with the student, asking what worked, what did not (and why), and what strategy the student will change next time, encouraging attributions toward controllable causes (e.g., strategy, effort, or time management). By aligning the Companion's features with these phases, the tool targets SRL processes that support growth in content knowledge and problem solving. These processes are more frequently observed among students who persist in gateway STEM courses than among those who leave (Hawkins et al., 2021).

Principle 4: Adaptive, Fading Scaffolds Within the Zone of Proximal Development (ZPD)

Building on the scaffolding and fading framework discussed above, we will design the Companion to adjust the amount and type of support based on student need and to fade that support as learners' competence grows. A steady stream of fully formed solutions can encourage an illusion of fluency, so the Companion will prioritize partial solutions, hints, and strategy prompts early on, and gradually shift toward brief checks and high-level prompts as students demonstrate more consistent understanding. In practice, this means that students who show recurring errors or fragile understanding will receive more structured, step-by-step guidance, while students who show consistent success will be moved toward independent problem solving with only light monitoring and feedback. This adaptive fading is intended to keep students working in their zone of proximal development while reducing dependence on the system.

Principle 5. Supporting Learning-with-AI Literacies and SRL in AI-Supported Environments

As noted above, effective learning in AI-rich environments requires students to engage AI as a partner rather than a bypass, drawing on SRL and co-regulation skills to do so (Anders & Dux Speltz, 2025). Accordingly, we will design the Companion to prompt and model "learning-with-AI moves," such as planning productive queries and prompts, choosing hints and explanations over direct answers, monitoring the plausibility of AI responses, comparing AI output to their own work and course resources, and reflecting on how the AI dialogue influenced their reasoning or decision-making. These interactions are intended to strengthen disciplinary problem solving and students' competence with regulating their learning when AI tools are available.

Core Functional Features of the STEM-AI Companion

Building on the theoretical framework and the core principles, this section specifies the baseline functional capabilities the STEM-AI Companion must support in gateway physics, calculus, and chemistry courses, and our iterative, stakeholder-engaged design processes. We focus on what the Companion needs to do for students and which interaction signals we will log.

By “interaction signals” we mean loggable features of learner and system behavior that can be used to personalize the Companion’s scaffolding and to analyze how it supports learning. These features enact the core principles discussed above. Table 3 summarizes the initial feature set that will be refined and implemented in Year 1 through design-based research cycles.

Table 3 Initial Companion Feature Set

SRL Phase	Companion Feature (Learner-Facing)	Key Interaction Signals Logged	Intended Function / Outcome
Forethought	Brief goal setting and strategy prompt at the start of a session or problem set	Stated goals; chosen strategies; frequency of skipping or engaging prompts	Support specific, process-focused planning before work begins
Performance	Constrained hints and “explain your thinking” prompts when help is requested	Hint vs. answer requests; explanation content; time between attempts	Keep students doing the cognitive work; surface reasoning and misconceptions
Performance	Error-pattern detection and targeted feedback on recurring mistakes	Repeated error types; changes in responses after feedback	Identify conceptual gaps and unproductive patterns to tailor support
Self-reflection	Post-task reflection mini prompts (e.g., “What was hard, and why?” “What will you change?”)	Reflection text; attribution patterns; link between stated plans and subsequent behaviors	Encourage adaptive attributions and concrete strategy changes over time
Crosscutting	Logging of “learning-with-AI moves” (e.g., queries, hints vs. answers, checks, and comparisons)	Query types; hint-to-answer ratio; instances of checking or challenging AI output; sequences of moves within sessions	Characterize and improve AI-supported SRL and co-regulation patterns

These entries are functional requirements, not rigid technical specifications. Each feature can be implemented with simple interface elements and prompting logic in the initial pilots, then iteratively refined.

Design-Based Research to Refine the Companion (Year 1)

Year 1 efforts will be centered in design-based research (DBR): a series of iterative, theory-driven design and study cycles in authentic gateway STEM courses. The goal is not to deliver a fully optimized STEM-AI Companion immediately, but to co-develop and refine it in context so that its core features, interaction patterns, and personalization logic are pedagogically sound and practically usable at scale. To that end, we will implement minimal versions of these features and instrument the associated interaction signals (e.g., help-seeking patterns and error profiles). Using DBR methods, we will analyze how these signals relate to proximal outcomes such as problem-solving performance, concept inventories, course assessments, and persistence in the course sequence. The goal is to identify which signals most meaningfully inform effective personalization and productive “learning-with-AI" SRL patterns, and to use those findings to refine the feature set and personalization logic for the larger-scale impact studies in Years 2–4. Table 4 summarizes these links. In line with DBR, we use a conjecture map—a structured representation that links design features to hypothesized mediating processes and desired outcomes—to guide iterative refinement of the Companion (Sandoval, 2014). DBR cycles in Year 1 will explicitly test and refine these conjectures by examining how actual interaction patterns with the Companion mediate changes in SRL processes and outcomes.

Table 4 Conjecture Map Linking Companion Features, Processes, and Outcomes

Companion Feature	Intended SRL / Co-Regulation Process	Anticipated Proximal Outcomes	Anticipated Distal Outcomes
Goal setting and strategy prompts at session start	More specific, process-focused planning in the forethought phase	Higher-quality study plans; earlier and more focused engagement	Improved performance on course exams and conceptual inventories in physics, calculus, and chemistry (conceptual understanding and procedural fluency); persistence to the next course in the sequence
Scaffolding: constrained hints and “explain your thinking” prompts	Increased strategic problem solving and metacognitive monitoring in performance	Better problem setup; fewer repeated conceptual errors	Higher scores on multi-step exam problems and standardized concept inventories in physics, calculus, and chemistry
Feedback: error-pattern detection and targeted explanations	Greater awareness of recurring misconceptions and strategy gaps	Targeted corrections; shifts to more productive strategies	Reduced failure and/or withdrawal rates in gateway courses
Scaffolding: post-task reflection mini prompts	More adaptive attributions and concrete strategy	Specific “next time” plans; reduced	Improved performance on later problems covering previously weak concepts;

	adjustments in reflection	avoidance after setbacks	stronger STEM identity; greater likelihood of staying in STEM
Logging and analysis of "learning-with-AI moves"	More productive engagement with AI (e.g., hints over answers, checking with AI)	Increased responsible AI use; richer explanations and comparisons	Better long-term regulation of study with AI across STEM courses; stronger conceptual gains across the gateway sequence

Iterative Design Cycles in Authentic Settings

Year 1 will also consist of three increasingly larger design cycles across introductory physics, calculus, and chemistry sections at Johns Hopkins University. We will carry out the DBR cycles in close partnership with course instructors, who will help determine how the Companion is introduced, when students use it, and how its outputs are interpreted and acted on.

Cycle 1: Feasibility and Usability. Over two months, we will deploy minimal viable versions of the core features in the platform in target domain courses. Data will include interaction logs, brief student surveys, and instructor feedback. Analyses will focus on basic feasibility (e.g., technical integration), usability (e.g., whether students engage with hints and prompts), and perceived value. Findings will inform immediate refinements to prompt wording, interface placement, and constraints on when and how help is offered. We will also provide brief instructor orientations and collect their feedback on when and how the Companion fits into existing teaching routines, and what support they need to feel comfortable encouraging its use.

Cycle 2: Refinement of Scaffolds and Signals. In the second cycle, we will refine the Companion's scaffolding and logging based on Cycle 1 results, with particular attention to: (a) the quality and timing of hints; (b) the design of explanation and reflection prompts (i.e., which questions elicit meaningful reasoning rather than perfunctory responses); and (c) the usefulness of different interaction signals for distinguishing productive from unproductive "learning-with-AI" patterns. We will use mixed-method analyses (including patterns in logs, qualitative coding of explanations and reflections, and user interview data) to identify which features and signals best support the intended SRL and co-regulation processes described in Table 2 above, and to clarify what information instructors find most actionable for adjusting instruction or outreach.

Cycle 3: Consolidation and Preparation for Impact Study. In the final two-month cycle, we will consolidate the most effective and feasible features into a stable Year 2–4 version of the Companion. We will confirm that the refined features can be deployed in multiple course sections and institutions with acceptable instructor workload and minimal disruption to existing course structures, and that instructors can readily interpret and use the Companion's summaries in their ongoing teaching. Across all cycles, a cross-disciplinary team of learning scientists, STEM teaching faculty in physics, chemistry, and calculus, computer scientists, and assessment experts will make design decisions collaboratively. Disciplinary faculty will provide input on which concepts to target, how to phrase scaffolds, and how to integrate the Companion into existing course structures, ensuring that refinements are both theoretically grounded and technically feasible.

DBR Products

By the end of Year 1, the DBR will yield the following: (a) a refined feature set for the STEM-AI Companion that is acceptable to instructors and students, includes simple instructor-facing summaries of key signals, and is fully integrated with the LASSO platform so that student mastery profiles and ability estimates seamlessly inform the Companion's guidance; (b) a prioritized set of interaction signals (e.g., help-seeking patterns or error profiles) that meaningfully inform personalization when combined with LASSO-derived diagnostic information and are empirically linked to problem-solving gains and early indicators of persistence; and (c) a revised conjecture map specifying which Companion features are most closely associated with desired SRL and co-regulation processes and which of those, in turn, are most predictive of learning and persistence outcomes when aligned with students' evolving diagnostic profiles.

The Year 1 DBR work ensures that large-scale evaluation of the STEM-AI Companion in Years 2–4 is conducted on a theoretically principled, empirically informed design, rather than on an untested set of features. It also directly supports the USED's emphasis on advancing AI in ways that improve instruction, support personalization, and can be scaled responsibly across diverse postsecondary contexts. The Companion will not replace the productive struggle of learning. Rather, it structures it by transforming the solitary struggle of a potentially at-risk student into a shared, co-regulated endeavor, providing the support necessary to turn a potential dropout into a persistent, self-regulated learner with stronger mastery of STEM content.

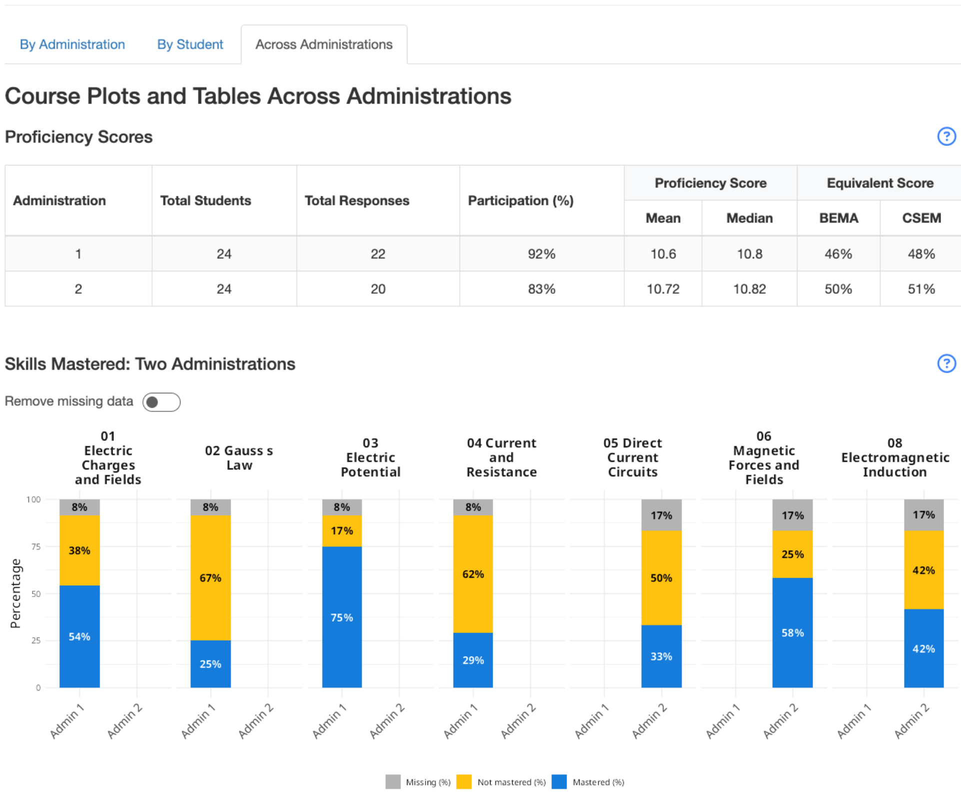
Together, these Year 1 products define a concrete specification for what the Companion must do. They spell out the functional requirements for the technical system: which types of assessments and item metadata are needed, how student states and "learning-with-AI" moves must be represented, which signals must be logged in real time, and how scaffolding and feedback must be orchestrated within authentic courses. In doing so, the Year 1 cycles will specify not only how the Companion responds to student signals, but also what instructors need to see and do with those signals for the system to function as a shared "seeing and doing" partnership among students, instructors, and AI (Holstein et al., 2020).

The next section describes how we will translate these design requirements into an integrated instructional infrastructure built on the LASSO platform. We outline how assessment delivery, IRT-based scoring, tagged knowledge bases, prompt orchestration, personalization logic, and mixture-of-experts routing will be combined to implement the STEM-AI Companion at scale. In this way, the infrastructure is not a separate technical layer, but the concrete instantiation of Year 1 design principles and conjectures, enabling us to test, refine, and ultimately deploy the Companion across diverse institutions in the Years 2–4 impact studies.

Iterative Design and Testing of the Infrastructure

The proposed project will design and implement an integrated instructional infrastructure that combines cognitive diagnostic (CD) computer adaptive tests (CAT), learning-objective-level feedback, and an AI-enhanced tutoring companion to support students in introductory physics, calculus, and chemistry. This work builds directly on the LASSO platform, which has supported over 700 instructors across more than 300 institutions and securely collected millions of student responses. LASSO's established capabilities for assessment delivery, psychometric scoring, data governance, and reporting allow the project to focus on development on new domain-specific assessment banks, learning-objective reporting, and the AI Companion. The design of these reporting tools builds on prior work examining how educators interpret and use graphical representations of assessment information to support instructional decision making (Salazar Morales et al., under review). Figure 2 presents an example instructor dashboard generated through LASSO's cognitive diagnostic reporting system, illustrating how instructors can monitor proficiency scores, participation rates, and learning-objective mastery across administrations to identify conceptual bottlenecks and guide instructional decision making.

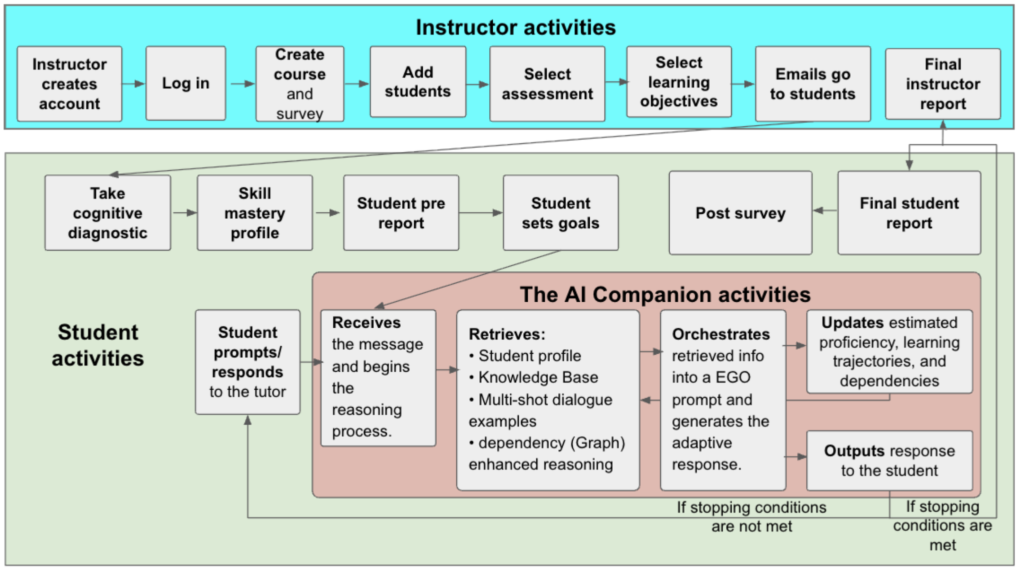
Figure 2. Example LASSO-Generated Instructor Cognitive Diagnostic Dashboard.



Administration	Total Students	Total Responses	Participation (%)	Proficiency Score		Equivalent Score	
				Mean	Median	BEMA	CSEM
1	24	22	92%	10.6	10.8	46%	48%
2	24	20	83%	10.72	10.82	50%	51%

Based on these existing capabilities, Figure 3 illustrates the proposed workflow for the integrated STEM-AI Companion system. In this workflow, instructors configure assessments and learning objectives; students complete a CD-based diagnostic and set goals; and the AI Companion uses assessment data and curated exemplars to provide multi-turn reasoning support before generating student and instructor reports. Because LASSO already hosts CDs in introductory physics and calculus (Le et al., 2025; Huang et al., 2025; Morphew, 2024), these subjects will adopt CD-based diagnostics immediately. Students in these courses will receive fine-grained skill-mastery profiles from the start of implementation, enabling early personalization and precise learning-objective–level feedback. Chemistry will begin with fixed-form assessments while a CD item bank is developed.

Figure 3: Activity Overview



Assessment development follows a staged psychometric plan aligned with the CD-CAT architecture. In Year 1, chemistry items will be authored and administered in fixed-form mode to support initial calibration. As calibration progresses, chemistry will transition to CD-CAT using the same procedures validated in physics and calculus. This sequencing enables immediate impact in physics and calculus while establishing a clear development path for chemistry.

The AI Companion integrates diagnostic outputs directly into the tutoring workflow. In physics and calculus, where CDs are available at launch, the Companion will use skill-mastery profiles to align scaffolds, examples, and reasoning supports to students' demonstrated needs. As CD-CAT becomes available, the Companion will draw on even more efficient, high-precision mastery estimates. Chemistry will adopt the same model once its CD bank and routing parameters are complete.

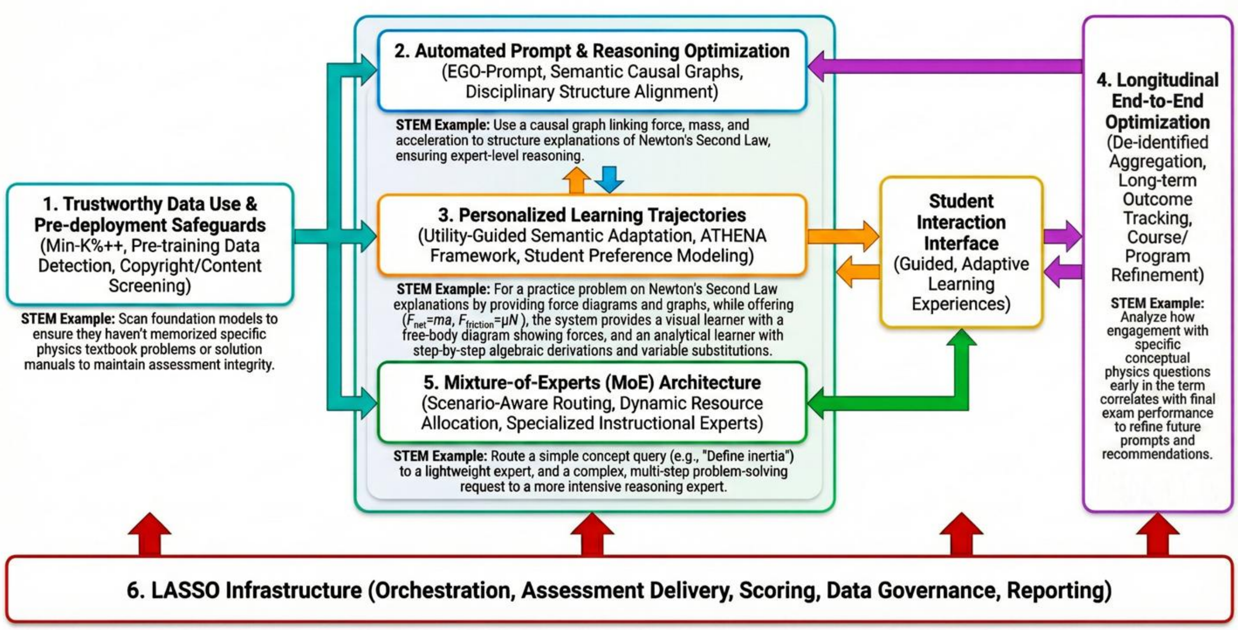
AI Companion Architecture for Trustworthy Personalized Learning

The proposed AI Companion is designed as a course-integrated orchestration layer that couples validated assessment data from LASSO with domain-aware, policy-driven personalization to support students in gateway STEM courses. Unlike generic chatbots, which provide one-shot answers drawn from broad training data, the Companion grounds its explanations in retrieval-augmented generation (RAG) using curated physics, calculus, and chemistry materials aligned with course learning objectives. This approach enables the system to provide explanations, scaffolds, and reasoning steps that are specific, trustworthy, and tailored to the disciplinary and instructional context, rather than generic responses derived from large-scale internet corpora. By "trustworthy," we mean that the Companion's outputs are grounded in instructional materials and domain rules, aligned with instructor learning objectives, constrained to avoid prematurely giving direct answers, auditable through logs and retrieval traces, and safe with respect to exam integrity, student privacy, and appropriate use. The architecture uses layered processing that filters inputs, retrieves vetted resources, enforces domain constraints, and adapts scaffolding levels based on students' evolving mastery and strategy profiles, thereby supporting SRL-oriented personalization while maintaining cognitive engagement.

Six-component architecture of the STEM-AI Companion. The system diagrammed in Figure 4 integrates trustworthy data safeguards, retrieval- and reasoning-optimized tutoring, personalized learning trajectories, longitudinal refinement, and a scalable mixture-of-experts design, all orchestrated through the LASSO infrastructure. Together, these components provide vetted, adaptive, and course-aligned support for STEM learning at scale.

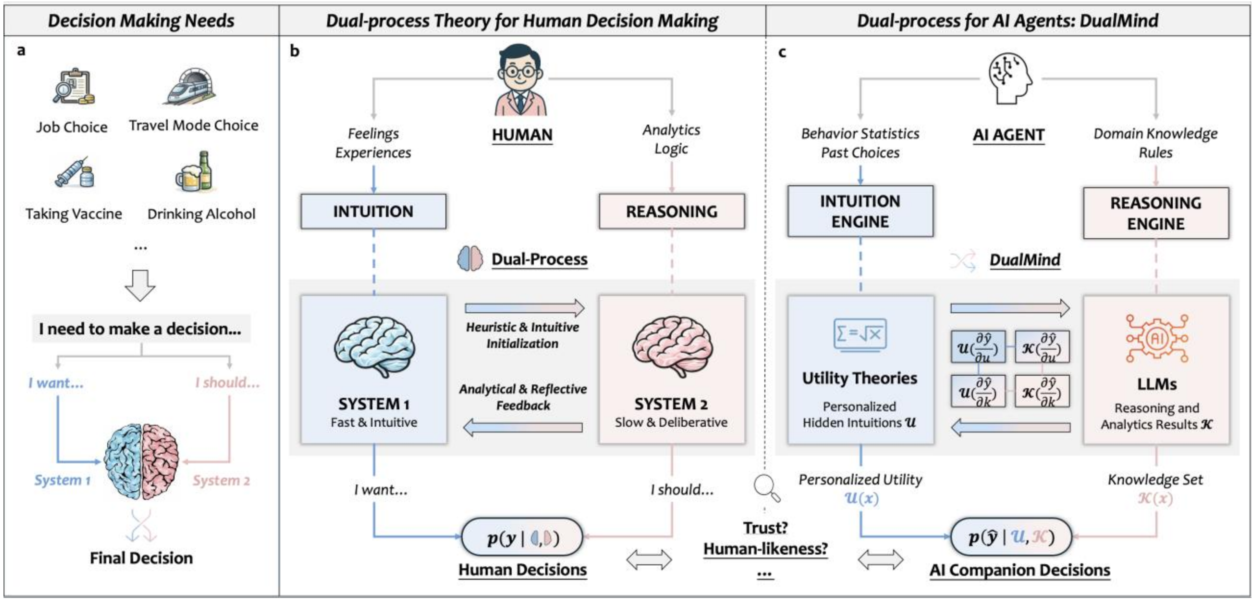
A first pillar of trustworthiness is the use of pre-training data detection to prevent contamination of instructional integrity. We adapt methods such as the Min-K%++ framework (Zhang et al., 2025) to educational settings to estimate whether sensitive instructional materials, such as assessment items or solution manuals, may be included in a foundation model's pre-training data. During deployment, similar monitoring logic flags responses that resemble protected materials and shifts the Companion toward vetted content or alternative scaffolding strategies. This approach helps protect exam integrity and reduces the risk of “memorized” outputs from pre-trained models.

Figure 4. AI Companion: Six-Component Architecture & STEM Case Example



A second pillar is retrieval-augmented, domain-aware reasoning. The Companion uses a RAG system (Lewis et al., 2020; Gao et al., 2023) that retrieves instructor-authored concept explanations, worked examples, multiple representations, and known misconceptions. The Companion’s EGO architecture lives within and powers this RAG pipeline. Specifically, EGO Prompting (Zhao et al., 2025) refines how retrieved materials are translated into optimized reasoning paths based on domain rules and prior student interactions. Retrieval supplies vetted content; EGO optimizes its explanatory form. Domain constraints—dimensional analysis, boundary conditions, and conservation laws—further prevent fluent but incorrect explanations, addressing a major limitation of general-purpose LLMs in STEM domains.

Figure 5. Proposed DualMind Framework.



Overview of the proposed DualMind. In Figure 5, personalized decision-making arises in diverse real-world scenarios. Dual-process theory, proposed by Kahneman (2011), models this process as the interaction between two cognitive systems: System 1, which processes intuitive signals such as "I want," and System 2, which performs deliberative reasoning associated with "I should." In human decision-making, these two systems interact to generate decisions. System 1 provides fast heuristic initialization based on feelings and experiences, while System 2 performs slower analytical reasoning based on logic and reflection. The iterative interaction between the two systems shapes observable decisions and how people learn from past choices. We propose DualMind to computationally model this decision mechanism. Utility functions are used to represent the intuitive system, while large language models (LLMs) capture the deliberative reasoning process. To integrate the two systems within a unified framework, we introduce DualMind forward reasoning for decision-making and DualMind gradient for learning from feedback. This integration allows DualMind to explore how AI agents' decisions can be aligned with human values.

A third pillar is DualMind-enabled prompt and reasoning optimization for STEM tasks. Prompt design will be refined through a dual-process optimization framework that combines structured learning-state evidence with course-aligned reasoning knowledge. The structured side, or utility mind, converts LASSO diagnostic profiles and interaction signals into estimates of which instructional action is most appropriate, such as asking the student to explain reasoning, providing a partial hint, shifting to a graph or symbolic derivation, reviewing a prerequisite concept, or prompting reflection. The reasoning side, or knowledge mind, converts instructor-authored concepts, prerequisite relations, common misconceptions, and solution strategies into optimized templates for physics, calculus, and chemistry tasks. In this way, DualMind strengthens the existing EGO-based prompt optimization pillar by giving it a principled mechanism for deciding which reasoning template, scaffold, representation, and reflection prompt should be used for a specific learner at a specific moment.

During each tutoring interaction, DualMind Forward Reasoning will generate a dynamic steering signal over candidate instructional moves. This signal will not replace the AI response or bypass retrieval safeguards. Instead, it will guide the Companion's retrieval, explanation strategy, and dialogue form so that responses remain grounded in vetted course materials while adapting to the learner's current mastery state, recent errors, help-seeking pattern, hint-to-answer ratio, time between attempts, and reflection quality. For example, two students may ask about the same physics problem, but one may need a prerequisite review of force components while another may need a metacognitive check about problem setup. DualMind allows the same RAG and EGO-based reasoning pipeline to select different prompts and representational forms for these learners without becoming a generic answer engine.

After the interaction, DualMind Gradient will refine both the utility mind and the knowledge mind using evidence from student responses, later mastery checks, reflection quality, and instructor review. If a scaffold improves subsequent performance, the associated symbolic instructional policy becomes stronger for similar contexts. If a student continues to make the same misconception after a certain explanation, the knowledge template can be revised toward a different representation, analogy, Socratic question, or prerequisite link. Thus, prompt optimization is not a one-time prompt-engineering activity; it becomes a learning process that improves the Companion's reasoning policies across Year 1 design-based research cycles, Year 2 pilot testing, and the larger randomized evaluation in Years 3-4.

Symbolic and graph outputs will make this third pillar auditable and trustworthy. Symbolic outputs translate the Companion’s personalization logic into readable rules, utility terms, or instructional policy summaries, such as selecting a targeted hint because mastery is low on a prerequisite objective, the same error appears across multiple attempts, and the student asks for an answer before articulating a plan. Graph outputs provide the relational context behind those symbolic rules by showing prerequisite links among concepts, misconception pathways, scaffold history, student progress across learning objectives, and learning-with-AI moves such as asking for hints, checking AI output, comparing representations, or reflecting on strategy. For students, the graph can show why a recommendation connects to earlier work and what pathway can lead to mastery; for instructors, it can reveal class-level bottlenecks and the AI-supported moves associated with improvement. Together, the symbolic and graph layers make the third pillar a glass-box reasoning system rather than an opaque prompt-selection process.

Merged into the third pillar, DualMind supports the full learning process. At the forethought phase, the Companion can use diagnostic mastery and prior interaction patterns to recommend process-focused goals and study strategies. During the performance phase, it can choose among hints, prerequisite reviews, worked examples, representation shifts, symbolic derivations, graph-based explanations, or explanation prompts while preserving productive struggle. During the self-reflection phase, it can guide students to interpret errors as controllable strategy issues and form concrete next-step plans. These decisions are generated through the same prompt and reasoning optimization pillar that governs STEM explanations, so personalization remains integrated with the course knowledge base rather than becoming a separate add-on.

The third pillar will therefore yield three concrete proposal products: (1) a symbolic instructional action policy library for selecting hints, prerequisite reviews, representation shifts, Socratic prompts, reflection prompts, and answer-withholding moves; (2) a STEM reasoning template bank for symbolic derivations, graph-based explanations, misconception contrasts, and SRL prompts in physics, calculus, and chemistry; and (3) a learning graph schema that connects concepts, prerequisites, misconceptions, scaffold histories, progress indicators, and learning-with-AI behaviors. These products will support instructor oversight, student trust, and evidence generation for the project evaluation.

A fourth pillar is evidence-based personalization grounded in utility modeling, adapted from the ATHENA framework for personalized decision modeling (Zhao et al., 2025). Rather than treating each interaction as isolated, the Companion models how different learners benefit from different instructional strategies—including hints, conceptual contrasts, partial solutions, reflection prompts, or prerequisite reviews—and adapts support accordingly. Students' "motivational constraints" are inferred from the time, effort, and strategy tendencies they reveal while working (for example, whether they give up early, repeatedly seek answers, or display persistent confusion). These signals shape the form, timing, and intensity of scaffolding. As a result, the level of personalization increases across conditions—minimal in LASSO alone, generic in the natural-AI condition, and fully tailored in the personalized Companion.

A fifth pillar is longitudinal optimization. De-identified interaction signals—including help-seeking patterns, error profiles, hint-to-answer ratios, and reflection content—are aggregated and linked to later cumulative assessments, course exams, and CD-CAT re-assessments. This linkage allows the system to learn which early scaffolds and behaviors predict durable conceptual understanding and course performance. Insights also inform instructors about recurring conceptual bottlenecks and effective supports, reinforcing continuous improvement at the course level.

A final pillar is scalable personalized instruction through a mixture-of-experts (MoE) design inspired by the EMOS system (Liu et al., 2025). In this architecture, specialized experts—LLM-based modules fine-tuned or configured for concept explanation, worked examples, metacognitive prompting, AI-literacy checking, representation switching, or constraint-based reasoning—are activated on demand. These experts are created by combining retrieved domain materials, optimized reasoning templates, and guardrail rules into task-specific modules. A lightweight routing mechanism selects only the necessary experts for each student query. This reduces computational load, maintains responsiveness under heavy usage, and makes it straightforward to scale the Companion to additional institutions or new STEM subjects by adding new expert modules rather than redesigning the core system.

Together, these components produce a trustworthy, course-integrated Companion that grounds its responses in vetted materials, aligns with CD-CAT diagnostic profiles, supports SRL-aligned scaffolding, and scales efficiently across diverse postsecondary institutions. The integration of RAG for vetted content retrieval, EGO for reasoning optimization, ATHENA for personalization, and EMOS-inspired routing for scalability ensures that the Companion operates not as a general-purpose chatbot but as an accountable instructional partner aligned with the goals of gateway STEM courses and the priorities of this DOE call.

The platform leverages LASSO's IRB-approved data governance model, including encrypted storage, role-based access controls, and FERPA-aligned protections (Van Dusen, 2018; Van Dusen et al., 2021). When assessment data are passed into the AI Companion, they are converted into anonymized, ephemeral representations to prevent exposure of identifiable information. LASSO's existing pipelines for automated scoring, analytics, usage logging, and instructor/course management will be adapted to support learning-objective reporting, CD administration, and integration with the new AI Companion.

A central contribution of the STEM-AI Companion is that it will not only provide personalized instructional support, but also function as a structured environment for developing students' awareness of disinformation, misinformation, and responsible AI use in authentic STEM learning contexts. Because students increasingly encounter AI-generated explanations, solution steps, graphs, summaries, and recommendations both inside and outside the classroom, the project treats AI literacy as an integral learning outcome rather than a secondary benefit. The Companion will be designed to help students recognize that AI responses can be useful but also incomplete, biased, overconfident, incorrectly reasoned, or misaligned with course concepts. Instead of positioning the AI as an unquestioned authority, the system will prompt students to evaluate, critique, compare, and revise AI-generated responses in relation to their own reasoning, course materials, instructor expectations, and diagnostic evidence from LASSO. For example, when the Companion provides a hint, explanation, symbolic representation, graph-based reasoning trace, or study recommendation, students may be asked to identify which parts are trustworthy, which assumptions are being made, what evidence supports the response, and whether the explanation is consistent with the target concept or procedure. In this way, engagement with the Companion becomes a repeated practice in epistemic monitoring: students learn not simply to ask AI for help, but to judge the quality, relevance, and reliability of AI-supported learning. This is especially important in gateway STEM courses, where students may lack the confidence or conceptual grounding needed to detect plausible but incorrect explanations.

The personalized design of the Companion strengthens this literacy goal because it can adapt critique prompts to students' current mastery level, help-seeking behavior, and prior interaction patterns. Students with fragile understanding may receive more structured comparison tasks, such as distinguishing between a correct and incorrect explanation, while students with stronger mastery may be asked to challenge the Companion's reasoning, identify missing constraints, or explain why a symbolic or graph-based output is valid. The evaluation will therefore examine not only whether students learn more STEM content, but also whether they become more capable of using AI responsibly: asking better questions, preferring hints over direct answers, checking AI output against evidence, identifying hallucinations or misleading claims, recognizing when human instructor support is needed, and reflecting on how AI shaped their reasoning. These outcomes will be measured through pre- and post-surveys of AI literacy and misinformation awareness, embedded critique tasks, analysis of interaction logs, student reflections, interviews, and comparisons across LASSO-only, general-purpose AI, and personalized STEM-AI Companion conditions. By integrating responsible AI critique directly into the learning cycle, the project will generate evidence on how personalized AI companions can improve both disciplinary learning and students' capacity to navigate AI-rich information environments with skepticism, agency, and informed trust.

Sampling, randomization, evaluation, and model finetuning in Years 2 - 4

Sampling Framework

This project features a three-arm, parallel-group stratified block randomization in which students, stratified by Pell, part-time, and transfer-in status, are randomly assigned within blocks to one of three conditions and nested within gateway courses for one 15-week semester. In semester two, those previously assigned to the non-AI Companion conditions will be assigned to the AI Companion treatment. Instructors of chemistry, calculus and physics will be the targets of recruitment for participation. The sampling of instructors will be facilitated in partnership with the National Science Foundation ECR Hub, which supports the Building Capacity in STEM Education Research (BCSER) Institutes. PI Odis Johnson leads one of the 11 BCSER institutes (ICQCM at JHU) and advises a second (SIARM at the University of Chicago). Each BCSER Institute has at least 20 STEM faculty researcher affiliates, ICQCM having the highest number of STEM faculty affiliates at 151 from different institutions, totaling a BCSER recruitment pool of approximately 350. Our second recruitment partner, the Learning Assistance Alliance, is a STEM focused instructional design network that has over 3000 members at 626 institutions. This proposal includes letters of support from the leadership of ECR Hub and the Learning Assistance Alliance.

A power analysis was conducted to determine the number of instructors to enroll and students for each of their courses. Setting a standard .80 level of power as our threshold, our calculations show we would need to recruit 16 sites and 45 students per site to approximately detect an effect size of at least. 20 at a significance of $p < .05$. Since these estimates are per course subject for three subjects, our target instructor recruitment is roughly 48 in years 3 and 4. It follows that the project would recruit 2160 students in years 3 and 4, for a combined sample of 4320. Adding the number of students involved in the JHU DBR year 1 work, and the expanded pilot in year 2, we estimate our total sample will include 48 – 54 instructors, and 4580 students.

Identifying Underserved Students through Administrative and Interview Data

This project will review publicly available **administrative data** about postsecondary institutions and the students they serve to ensure the project benefits all students, including the underserved. Institutional level data will include but is not limited to the institution's Pell Grant eligibility rates, first generation college status, part-time or transfer-in status, location within an EPSCoR (Estab-lished Program to Stimulate Competitive Research) state or territory, selectivity, and STEM degrees awarded, among other indicators. In addition, participating instructors will provide institutional data within the LASSO platform about their program structure, instructional history, practices, and beliefs about their teaching and STEM learners.

Using participating instructors' enrollment records, 30 students will be randomly selected for post-treatment **individual interviews** and invited via email (5 students in both years 1 and 2, and 10 in both years 3 and 4). Research staff will use both semi-structured questions to ensure theme coverage and open-ended questions to reveal their project experiences. We will code the qualitative data to identify key themes related to past access to AI, access to high-quality STEM courses, views about the effectiveness of the AI Companion or alternative AI tools, and their thoughts about self-regulation, drawing upon Miles et al's (Miles et al., 2018) systematic approach to qualitative analysis. Rather than following a predetermined set of codes, we will use an emergent strategy to illuminate dominant patterns across participants (Corbin & Strauss, 2008; Creswell, 2014). We will first organize data descriptively through an open-coding round, in which we will identify emergent features that describe the essence of students' learning and SRL. Next, these will be grouped into fewer categories, which we will use to perform a focused, second-order coding round to identify emergent themes and relationships between participants and the research questions.

A 3-arm parallel-group stratified block randomization design

Students are stratified by Pell, part-time, and transfer-in before assigning them to blocks. We define three experimental conditions. Arm 1 is a condition with the LMS/assessment system of LASSO, serving as the control. Arm 2 adds a natural generative AI model such as ChatGPT to LASSO, serving as a "low dose" of AI technology in a heterogeneous mix of student learning behaviors from passive to active. Arm 3 focuses on the proposed AI Companion that provides the "high dose" of trustworthy and personalized learning support with the LMS/assessment infrastructure from LASSO. Our three-arm design uses parallel-group stratified block randomization (Ratkowsky et al., 2020). Within each of the large courses (Calculus, Physics, Chemistry), participant students enrolled in the PLTL program are stratified by Pell, part-time, and transfer-in status and then randomly assigned, in equal proportion, to one of three groups (Groups 1, 2, and 3) corresponding to the three arm conditions, as shown in the following Table. Each student remains in the assigned condition across the 15 weeks of the fall semester, so that outcomes can be compared between arms while balancing student composition across conditions. Repeated measures collected over the term—at the beginning of the first half, at mid-semester, and near the end of the term (fall weeks 3–8, fall weeks 9–14)—support longitudinal modeling of within-student growth alongside the between-arm impact estimates.

Table 5 Randomization, Evaluation, and Model Finetuning at Test/Deployment Stages

Year 1	DBR and Instructor recruitment, technical consolidation, readiness of arm conditions for implementation				
Year 2	Montana State, Iowa State, community college (TBA), underserved (TBA)				
Testbeds: PLTL	JHU				
Courses	Physics, Chemistry, Calculus				
Instructors/TAs	3 courses				
Semester	Fall	Fall	Spring	Spring	Summer
Academic week	3-8	9-14	3-8	9-14	
Info/training	enrolled students				
Random assignment (parallel groups)					
Group 1	Arm 1	Arm 1	Arm 3	Arm 3	
Group 2	Arm 2	Arm 2	Arm 3	Arm 3	
Group 3	Arm 3	Arm 3	Arm 3	Arm 3	
Task	AI upskilling, mobilization	implement treatment, collect data			evaluation, model finetuning
Year 3	48 Instructors/16-18 Institutions				
Testbeds: PLTL					
Task	AI upskilling, mobilization	implement treatments, collect data			
Courses	Physics, Chemistry, Calculus				
Instructors/TAs	3 courses				
Semester	Fall	Fall	Spring	Spring	Summer
Academic week	3-8	9-14	3-8	9-14	Analysis
Info/training	enrolled students				
Random assignment (parallel groups)					
Group 1	Arm 1	Arm 1	Arm 3	Arm 3	
Group 2	Arm 2	Arm 2	Arm 3	Arm 3	
Group 3	Arm 3	Arm 3	Arm 3	Arm 3	
Task	AI upskilling, mobilization	implement treatments, collect data			evaluation, model finetuning
Year 4	48 Instructors/16-18 Institutions (same ones as in year 3)				
Deploy AI companion to 16-18 HE institutions					evaluation, model finetuning, report

Notes: Arm 1: LMS/assessment service of LASSO. Arm 2 adds to Arm1 natural AI support. Arm 3: AI companion.

Horizontally by Groups 1, 2, and 3, the design yields between-arm contrasts on the confirmatory outcomes: the primary contrast is Arm 3 (Companion) versus Arm 1 (control), with Arm 2 (general-purpose AI) as a secondary contrast. This design can address central questions like "Does the use of natural AI effectively enhance student learning when compared with no AI involvement?" and "Does the AI companion effectively enhance student learning when compared with natural AI support?"

Design implementations in authentic learning environments

Testbeds. We use the peer-led team-learning (PLTL) programs or similar programs as our testbeds in JHU in Year 1 and then in four other higher education institutions in Year 2 of the project. The PLTL model employs the principle of proximal development (ZPD), the range of tasks that a learner can perform with the guidance and support of a more knowledgeable other such as a peer (Vygotsky & Cole, 1978). PLTL programs often adopt scaffolding procedures. PLTL emphasizes the social construction of knowledge where students overcome obstacles while moving from a ZPD to another in a learning community. PLTL has been a successful practical learning model for over two decades (Barrasso & Spilios, 2021; Wilson & Varma-Nelson, 2016). For example, the PILOT programs at JHU following the PLTL model serves nearly the entirety of undergraduate students enrolled in ~20 introductory STEM courses in a 90–120-minute weekly session. Using PLTL programs as testbeds has three unique advantages. First, the PLTL programs strictly follow the course-specific progress under the supervision of the instructor and led by carefully identified peer leaders from the older cohort who are trained with the PLTL practices. Second, the PLTL programs are explicitly uncoupled from the course grades such that students can ask any questions and any support without worries about reflecting them in the course grade. Third, the PLTL programs are mature with human peers' organizing and tutoring experiences. These properties are excellent for PLTL programs to serve as testbeds.

Student participants. The three courses we selected (Calculus, Physics, Chemistry) are foundational courses for many STEM majors, so the enrollment is large, and the enrollment in PLTL is also similarly large. We will mobilize all students enrolled in those courses' PLTL programs in weeks 3-7 of the fall semester. The mobilization activities include (1) one information section introducing the study (in the first weekly PLTL student meeting outside of PLTL sessions); (2) AI literacy basics and competence upskilling (depending on students' skill levels); and (3) Detailed introduction and usage of the Arms 2 and 3 conditions.

A customized interface. We will build a multi-functional interface. The functions include student logins and consents, 3-arm parallel-group random assignment, arm condition deliveries, participant compliance behaviors, timestamped student-LMS usage contents, student-AI Companion interaction contents, weekly problem set items and students' submitted solutions to each item (which measure individual weekly learning outcomes), and the data collection using relational database management to save at Amazon Work Spaces daily and delete after uploading the daily data to the project's OneDrive raw data folder.

Modeling Strategies

Educational experiments and causal inference. There has been a long tradition of field experimentation in educational research (Shadish et al., 2009; Cobb et al., 2003). An educational intervention such as a new pedagogical design or technical intervention can be randomly assigned to students, classrooms, schools, or school districts. Under the "potential outcome" framework (Rubin, 2005), statistical methods for estimating causal effects of treatment that is both randomly assigned and received (Little & Yau, 1998; Peugh et al., 2017) are used to estimate the causal effect of an intervention. Field experiments about human-machine interactions are relatively rare. Recently, researchers have attempted to estimate the causal effect of college students' use of ChatGPT. A review (Jin & Sercu, 2025) found that the research often focused on learning task assistance and general learning support in language and STEM domains. ChatGPT was generally effective for knowledge acquisition, but its impact on skills development varied. These studies also show that the duration of intervention matters–-medium-term interventions lead to better results than short-term intervention. However, the degree to which college students trust ChatGPT and the ways by which students interact with generative AI for a learning task have not been examined. In addition, these existing experiments have lacked large, diverse samples, rendering the robustness of the findings.

Addressing the limitations. To address the limitations in the literature, we set forth to both make causal inferences and analyze behavioral mechanisms to unpack the "black box" of experimentation. Primary impact estimates use a pre-specified, intent-to-treat multilevel model with students nested within sections, comparing the Companion arm to the control on each confirmatory outcome and adjusting for stratification variables and a pre-intervention baseline. We will establish and report baseline equivalence of the analytic sample and both overall and differential attrition against What Works Clearinghouse thresholds, and report standardized effect sizes. As a secondary, mechanism-focused analysis, we will model how interaction-log behaviors (help-seeking, hint-to-answer ratio, reflection quality) mediate effects on the confirmatory outcomes. We use natural language processing to encode the textual data of interaction contents using a series of supervised machine learning models to select the optimal model for the detection of behavioral mechanisms in Arms 2 and 3 leading to the measured outcomes.

A complementary research direction: interaction traces as instrumented self-regulation. Beyond the confirmatory impact analysis and the usability and qualitative coding described above, the team will pursue an exploratory line of work grounded in a data source the project already collects. Every student–Companion exchange is logged with timestamps, including help-seeking, retries, revisions, and responses to feedback. These traces carry direct signal about the process the intervention is designed to change. Because the Companion works by co-regulating self-regulated learning—planning, monitoring, and recovering (Zimmerman, 2002; Holstein et al., 2020)—and because self-regulation is a process expressed in the sequence and timing of student engagement rather than in a periodic survey, the interaction trace provides a high-frequency, in-situ window onto that mechanism that complements the project's administrative and survey measures.

Modeling the trace to enrich the core methods. The logging infrastructure that captures these traces is operational within LASSO, and the team brings established expertise in AI, data science, and large-scale data analysis—including methods well suited to interaction traces, such as sequence modeling and latent-state estimation. The team will apply that capability to model the trace as an evolving latent state—an estimate of how a student's self-regulation shifts over the course—rather than reducing it to counts. This strand enriches and contextualizes the core methods. It supplies the secondary mechanism analysis with a continuously measured mediator; it informs the adaptive-fading logic by giving the Companion a contingent read of student state rather than a fixed schedule; and it provides a principled place to detect learned dependence, where students offload rather than regulate (Rohilla, 2025). The confirmatory design is independent of this strand: its outcomes are the administrative measures—gateway-course completion, credit accumulation, and retention—and its validity does not rest on the latent-state model.

These mechanism questions guide the exploratory strand: (1) How do students engage with the LMS/assessment system and the Companion as they work through gateway STEM content in the

PLTL environment? (2) Which interaction patterns distinguish passive use from self-regulated use—planning, self-checking, and revising? (3) How does the Companion shift students from passive toward self-regulated patterns over the term, and does that shift track with stronger performance on the confirmatory outcomes? These are addressed through the interaction-trace modeling described above and are framed as questions the secondary analysis investigates, not as additional confirmatory effects.

Scope and validation. The data and the analytic capability are in hand; what remains a research problem is the validity of the constructed measure itself. Modeling self-regulation as a latent state from interaction traces is not a settled method, so the latent measure will be validated against established SRL instruments, expert-coded transcripts, and proximal performance during the Year 1 design-based research cycles before any mediation analysis relies on it. The strand is therefore positioned as an exploratory direction that extends the project's core methods, with no confirmatory claim resting on it.

MANAGEMENT PLAN

This project team reflects a strategic partnership between **Johns Hopkins University**, **Iowa State University/LASSO**, and **Quality Measures LLC.** Our team-based approach is designed around the three critical goals: *communication* (frequent, engaging communication and expectations alignment), *collaboration* (ensuring all partners have clarity about roles and responsibilities and work toward common goals), and *critical thinking* (examining our frameworks and tools frequently to ensure that we are adhering to our goals and standards and learning from testing and field work).

The **Core Team** includes **Odis Johnson, Ph.D. (JHU),** Bloomberg Distinguished Professor of Social Policy and STEM Equity at Johns Hopkins University. Dr. Johnson is core AI faculty in the JHU Data Science and Artificial Intelligence initiative, a six-time NSF Principal Investigator, and currently directs the NSF Institute in Critical Quantitative and Computational (ICQCM). Dr. Johnson will oversee project administration and serve as the chief science officer. **Hao (Frank) Yang, Ph.D. (JHU)** is an Assistant Professor focused on trustworthy and responsible AI, with award-winning work, including the AASHTO High-Value Research Award, in model safety and scalable AI systems. He leads development of the STEM-AI Companion's technical architecture to provide reliable and personalized instructional support. **Ben Van Dusen, Ph.D. (Iowa State University; LASSO)** is an Associate Professor of Science Education and the Director of the LASSO platform. Dr. Van Dusen studies modern measurement, large-scale digital assessment systems, and the integration of AI-supported tools to improve learning in undergraduate mathematics and science; he provides expertise in developing advanced assessment models and designing scalable technologies to support instructors and students. **James Diamond, Ph.D. (JHU)** is an Assistant Professor in the Johns Hopkins School of Education and program director for Learning Design and Technology. Dr. Diamond is a design-based researcher with expertise in the development and evaluation of educational technologies to support learning opportunities across K-16 domains. Dr. Diamond will oversee the qualitative research and assist with AI Companion implementation. **Lingxin Hao, Ph.D. (JHU)** is Benjamin H. Griswold III Professor in Public Policy and Professor of Sociology. Dr. Hao provides expertise in quantitative/computational methodologies and the sociology of education; she is the PI of an NIH Center on population research and has served as the PI of multiple NSF grants to study the processes and consequences of novel learning opportunities for K-16 students. **Dr. William Gray-Roncal, Ph.D.** is a Principal Research Engineer at the Johns Hopkins Applied Physics Laboratory, a Visiting Assistant Professor in the School of Education, and an Assistant Research Professor (courtesy) in Computer Science at the Johns Hopkins Whiting School of

Engineering. He is an expert in artificial intelligence, large-scale data systems, and graph-based analysis. He will support the development of AI-ready data representations, scalable analytical pipelines, and the application of modern machine learning methods and AI literacy measures. **Tak Igusa, Ph.D. (JHU)** is a Professor of Civil and Systems Engineering whose work focuses on engineering complex, data-driven interventions in public health and related domains, including the development of educational products. He provides systems-integration expertise to coordinate assessment and AI support. **Gwen Lee-Thomas, Ph.D. (QM)**, CEO of **Quality Measures LLC**, has provided evaluative services since 2009, including federal projects and RCTs that aligned with WWC research and evaluation standards. Dr. Lee-Thomas and her staff will provide formative and summative evaluation feedback to the team to advise management and key decisions throughout the project.

Project Staff. In addition to the above identified scholars, a team of post-docs, interns, and financial and administrative support professionals at the partner institutions will work with the PI and Co-PIs to ensure that project objectives are achieved on time and within budget. Regular virtual all-team staff meetings will check progress against key milestones, with quarterly reflections to identify challenges and collaboratively create solutions.

Table 6 Timelines and Milestones

Year/Quarter	**Milestone/Activity**	**Responsible Parties**	**Deliverables/Outcomes**
Year 1 Q1	Project launch, team onboarding, finalize partnerships, IRB approvals	PI, Co-PIs, Admin Team	Signed MOUs, IRB approval letters, kickoff meeting notes
Year 1 Q2	Develop and pilot minimal viable STEM-AI Companion features in LASSO; instructor orientation	Tech Team, Faculty Partners	Pilot Companion deployed in select courses; instructor feedback collected
Year 1 Q3	Design-Based Research Cycle 1: Feasibility & Usability Testing	Research Team, Instructors	Usability data, student/instructor surveys, technical integration report

Year 1 Q4	Refine Companion features and interaction signals; begin Cycle 2	Tech Team, Research Team	Revised feature set, updated logs, focus group summaries
Year 2 Q1-Q2	Design-Based Research Cycle 2: Scaffold & Signal Refinement; expand pilot to additional courses/institutions	Research Team, Faculty Partners	Mixed-methods analysis, refined scaffolding, instructor dashboards
Year 2 Q3-Q4	Consolidate features for scale-up; prepare for impact study	Tech Team, Evaluation Team	Stable Companion version, deployment plan, instructor training materials
Year 3 Q1-Q2	Impact Study Launch: Randomization across multiple institutions; data collection	Research Team, Third-Party Evaluator	Baseline data, randomization logs, compliance tracking
Year 3 Q3-Q4	Ongoing data collection, formative evaluation, continuous improvement	All Teams	Interim evaluation report, system refinements, stakeholder feedback
Year 4 Q1-Q2	Full deployment to 16-18 institutions; summative evaluation	PI, Co-PIs, Third-Party Evaluator	Final performance data, summative evaluation report
Year 4 Q3-Q4	Dissemination of findings, sustainability planning, final reporting	PI, Co-PIs, Advisory Board	Annual and final reports, publications, sustainability plan

EVALUATION PLAN

Project External Evaluation Plan Overview: The overall evaluation will reflect an ***Impact Evaluation*** (Bamberger, Mabry, & Rugh, 2020) with a Stratified Block Randomization (SBR) to compare three groups of students randomly assigned to (a) the STEM-AI Companion, (b) a general-purpose AI model, or (c) no AI assistance (LASSO only). The primary purpose of the comparison is to determine if there were any statistically significant differences in the change in students' non-cognitive factors and cognitive factors as related to retention and success in "make-or-break" gateway courses in STEM. In addition to the impact evaluation, a *Process Evaluation* will occur simultaneously to monitor the implementation of activities (recruitment of faculty and their students to test STEM-AI Companion, cross-institutional context (i.e., similarities and differences in gateway courses, instruction, in-class assessment, and academic leadership support), and data monitoring (i.e., retention tracking, student success indicators, etc.)) These two forms of evaluation—*Impact Evaluation* and *Process Evaluation*—will inform the project team of what worked well and what needed improvement in real-time as well as document the extent to which the STEM-AI Companion project resulted in increasing student retention and content mastery in gateway courses and the contribution of motivation, self-efficacy, perceived support, and sense of belonging.

Research Design within the Evaluation: For the SBR, the independent variables will include the non-cognitive profile of the students (student major or minor, transfer statuses, Pell recipient, first generation college student, self-efficacy, motivation, perceived support, and sense of belonging), and cognitive profile (GPA in pre-requisite courses, course repeat) across all groups. Confirmatory outcomes are: (a) gateway-course completion (Pass A/B/C vs. DFWI; primary), (b) first-year credit accumulation (credits earned; primary), and (c) term-to-term retention/persistence (secondary). Analysis uses a pre-specified, covariate-adjusted multilevel model (students within sections), reporting intent-to-treat effects and standardized effect sizes. Because this is an early-phase project, the impact study is designed to meet What Works Clearinghouse standards with reservations: random assignment to condition, a measured pre-intervention baseline with reported equivalence, administrative outcomes collected identically across arms, and attrition kept within WWC bounds.

Evaluation Questions: To guide the evaluation strategies, the overarching evaluation question is *"To what extent did a personalized AI companion result in retention, content mastery in gateway courses, and improvements in institutional readiness for promoting student AI literacy regardless of students' cognitive and non-cognitive profiles?"*

To provide documented evidence that can provide a response to the overarching evaluation question, the following evaluation sub-questions are provided.

1. To what extent did the STEM-AI Companion increase student retention in STEM and student' content mastery in gateway courses across the 16–18 participating institutions? *(Impact Evaluation)*
2. To what extent did implementation of the STEM-AI Companion provide iterative knowledge regarding the support of generating AI literacy among students in STEM degree programs *(Process Evaluation)*
3. To what extent did faculty adoption of the STEM-AI Companion prepare them for supporting student learning with an influence on retention and completion of gateway courses? *(Impact Evaluation)*
4. To What extent did the results from the project's examining questions support understanding of how personalized AI operates in various instructional settings? *(Impact Evaluation)*
5. To what extent can the combined results from the study and evaluation of the STEM-AI Companion be replicated and scaled to other programs and institutions? *(Sustainability)*

Data Collection Activities: The triangulation method of data collection for each evaluation question will be used to confirm and disconfirm findings across instruments and groups (Patton & Campbell-Patton, 2022). The following data collection sources and activities are proposed.

- <u>Student Participants</u>: In addition to the research component of measuring the influence of the STEM-AI Companion on cognitive and non-cognitive profiles as independent variables with two dichotomous dependent variables, students will also complete **pre- and post-surveys** that measure their cognitive profiles with adaptations from the following validated instruments: (a) Motivated Strategies for Learning Questionnaire (MLSQ), Academic Self-Efficacy Scale (ASES), Validation of Perceived Academic

Support Questionnaire, and The Development of the Student Belonging Scale SBS). The pre-post surveys will also measure students' level of comfort using/not using AI as well as their clarity regarding ethical and supportive use of AI. Finally, a representative sample of students will participate in two **focus groups** midway and at the end of the project to gauge the extent to which the iterations resulted in perceived helpfulness and influence on cognitive and non-cognitive factors.

- <u>Faculty Adopters</u>: Faculty who implemented the STEM-AI Companion will participate in two customized surveys at the beginning **(pre-survey)** and at the end **(post-survey)** or their participation to measure their level of comfort and satisfaction with using personalized AI in their teaching and learning environments. Cross-comparisons of subject matter/content, teacher experience, and level of comfort with using AI in the learning process will also inform pedagogical implications and outlook on AI's contribution to student learning.
- <u>Project Leaders</u>: Project Leaders include the PI, Co-PI, and primary contacts who guide and lead the work. The leaders will participate in a **semi-structured interview** at the end of Y1 to discuss lessons learned, challenges resolved or unresolved, and successful and unsuccessful implemented activities as well as unimplemented activities. The interviews will also address the challenges and benefits of scaling the STEM-AI Companion experiences and activities, recruiting and onboarding faculty, and navigating multi-institutional cultures and bureaucracies. A second interview will occur at the beginning of Y3 to recap the Y1 responses and identify any changes that occurred, identify newly resolved and unresolved challenges, and share new lessons learned. The final interview will be at the end of Y4 to discuss their perspectives regarding what made the project

successful/unsuccessful, and what could have had a greater influence on the outcome, and share lessons learned that can impact replicability, scalability, and sustainability.

Data Analysis: Quantitative data will be collected via metrics (i.e., project team reporting of the tracking metrics by the project, number and descriptions of course modifications, and pass-rates in gatekeeping courses) and non-metrics (i.e., satisfaction with experience, impact of experiences on retention, completing the gatekeeping course, future perceived use of AI in the teaching and learning process, etc.). Quantitative impact analysis uses the pre-specified, covariate-adjusted, intent-to-treat multilevel model described in the Research Design—students nested within sections—estimating the effect of random assignment to the Companion arm versus control on the confirmatory outcomes (gateway-course completion, first-year credit accumulation, and term-to-term retention) and reporting standardized effect sizes. Descriptive statistics summarize each arm.

Qualitative data will include interviews, focus groups, open-ended questions on the surveys, and potentially observations. Qual Analysis will include a phenomenological approach to address the context in which the STEM-AI Companion related experiences and activities contributed to increasing students' non-cognitive experiences, content mastery, AI literacy, and ethical use of AI; and sustainable efforts. Together these designs will form a mixed-methods approach to the evaluation.

Reporting: The reporting will include three types. The first will include semester updates of the evaluation activities, challenges faced, summary of results/findings, and identification of any support that may be needed to enhance/improve response rates of participants for the various data collection activities for just-in-time decision- making (Patton and Campbell-Patton, 2022). The second type of reporting will include an annual report (Y1 – Y3) as a Formative Evaluation of the year's evaluation activities and findings operationalized by the evaluation questions. The annual report will reflect a continuous improvement approach and include both recommendations for the project's work and evaluation strategies. Finally, at the close of the project in Y4, a Summative Evaluation Report will provide documented evidence from the findings and the previous reports to inform the overarching evaluation question that aligns with the primary purpose of the project.

AI-Companion and Assessment System Formative Evaluation. In addition to evaluation of the project, formal evaluation activities will examine the quality, reliability, usability, and instructional value of the assessment system and AI Companion through coordinated psychometric, technical, and experiential measures. The evaluation framework is structured across four integrated domains: psychometric validity, AI-tutor quality, system performance, and student experience. The effort will be handled by a third-party evaluator, who will conduct a high-touch suite of evaluations and engage in regular feedback to support continuous improvement throughout the life of the grant.

1. **Psychometric Evaluation**. Psychometric analyses will assess item performance, model fit, score precision, and classification reliability across physics, calculus, and chemistry, with explicit attention to the transition from fixed-form assessments to CD. Item functioning will be examined through difficulty, discrimination, and guessing parameters, along with differential performance across administrations. Model fit will be assessed using established criteria for IRT fit indices (Hooper et al., 2008; Hu & Bentler, 1999), and diagnostic components will be evaluated using classification consistency and accuracy procedures for DINA modeling (Cui et al., 2012; Ravand & Robitzsch, 2015).

For the CD-CAT system, evaluation will examine routing efficiency, item–skill alignment accuracy, and reductions in test length achieved through adaptive item selection. The project aims for at least 80% of items to achieve discrimination values of 0.30 or higher; IRT model-fit indices to meet conventional thresholds ($CFI \geq .95$, $RMSEA \leq .06$); score stability across administrations of 0.85 or higher; DINA RMSEA2 and SRMSR <0.06 classification accuracy >0.85 across learning objectives.

2. **AI Companion Quality and Safety**. Companion output will be evaluated through expert-coded samples of tutoring transcripts that assess factual accuracy, grounding, alignment to learning objectives, and safety. Coders will examine whether responses draw on retrieved knowledge-based documents, remain coherent with assessment profiles generated from the CD engine, and follow productive scaffolding patterns (Cai et al., 2021; Gupta et al., 2020). Evaluation metrics include achieving alignment accuracy of at least 85%, grounding citation rates of 90% or higher, grounding similarity values of at least 0.12 (Radeva et al., 2024), and inter-rater reliability of 0.70 or higher (Warren, 2015). AI behavior will also be reviewed for instructional challenge and support, informed by research showing that properly scaffolded confusion can deepen learning (D'Mello et al., 2014; Graesser et al., 2014).

3. **Student Experience and Usability**. Student experience will be assessed through post-interaction surveys, log-based analytics, and validated usability measures. Students will complete the System Usability Scale (Freeman et al., 2014) and the Chatbot Usability Scale, with a performance target of achieving mean scores of 70 or higher on both instruments by the end of Year 2. Interaction logs—capturing number of turns, time on task, and task completion—will be used to identify pacing issues, conversational drop-off points, and improve the flow of the tutoring experience.

Evaluation Outputs. Across the four-year project, evaluation findings will guide iterative refinements to the assessment banks, CD engine, tutoring architecture, item calibration, and student-facing reporting. Annual summaries will document improvements to psychometric performance, AI-tutor grounding and coherence, system stability, and student usability, ensuring that the platform remains robust, instructionally aligned, and scalable for widespread use.

REFERENCES

Anders, A. D., & Dux Speltz, E. (2025). Developing generative AI literacies through self-regulated learning: A human-centered approach. *Computers and Education: Artificial Intelligence, 9*, 100482. https://doi.org/10.1016/j.caeai.2025.100482

Arogundade, O. R. (2023). Structuring knowledge bases with AI and machine learning. *Information and Knowledge Management, 13*, 31–39.

Bamberger, M., Mabry, L., & Rugh, J. (2020). *RealWorld evaluation: Working under budget, time, data, and political constraints* (3rd ed.). SAGE.

Barrasso, A. P., & Spilios, K. E. (2021). A scoping review of literature assessing the impact of the learning assistant model. *International Journal of STEM Education, 8*(1), Article 12. https://doi.org/10.1186/s40594-020-00267-4

Black, P., & Wiliam, D. (1998). Assessment and classroom learning. *Assessment in Education: Principles, Policy & Practice, 5*(1), 7–74. https://doi.org/10.1080/0969595980050102

Buncher, J. B., Nissen, J. M., Van Dusen, B., & Talbot, R. M. (2025). Is the Force Concept Inventory biased across the intersections of gender and race? *Physical Review Physics Education Research, 21*(1), 010137. https://doi.org/10.1103/PhysRevPhysEducRes.21.010137

Cai, H., Chen, H., Song, Y., Zhao, X., & Yin, D. (2021). Exemplar-guided neural dialogue generation. In *Proceedings of the Thirtieth International Joint Conference on Artificial Intelligence* (pp. 3601–3607).

Cai, L., Msafiri, M. M., & Kangwa, D. (2024). Exploring the impact of integrating AI tools in higher education using the zone of proximal development. *Education and Information Technologies, 30*(6), 7191–7264. https://doi.org/10.1007/s10639-024-13112-0

Chan, J. Y. K., & Bauer, C. F. (2015). Effect of peer-led team learning (PLTL) on student achievement, attitude, and self-concept in college general chemistry in randomized and quasi-experimental designs. *Journal of Research in Science Teaching, 52*(3), 319–346. https://doi.org/10.1002/tea.21197

Chen, X. (2013). *STEM attrition: College students' paths into and out of STEM fields* (NCES 2014-001). National Center for Education Statistics. https://nces.ed.gov/pubs2014/2014001rev.pdf

Cobb, P., Confrey, J., diSessa, A., Lehrer, R., & Schauble, L. (2003). Design experiments in educational research. *Educational Researcher, 32*(1), 9–13. https://doi.org/10.3102/0013189X032001009

Corbin, J., & Strauss, A. (2008). *Basics of qualitative research: Techniques and procedures for developing grounded theory* (3rd ed.). SAGE. https://doi.org/10.4135/9781452230153

Creswell, J. W. (2014). *Research design: Qualitative, quantitative, and mixed methods approaches* (4th ed.). SAGE.

Cui, Y., Gierl, M. J., & Chang, H.-H. (2012). Estimating classification consistency and accuracy for cognitive diagnostic assessment. *Journal of Educational Measurement, 49*(1), 19–38.

Dai, T., & Cromley, J. G. (2014). Changes in implicit theories of ability in biology and dropout from STEM majors: A latent growth curve approach. *Contemporary Educational Psychology, 39*(3), 233–247. https://doi.org/10.1016/j.cedpsych.2014.06.003

D'Mello, S., Lehman, B., Pekrun, R., & Graesser, A. (2014). Confusion can be beneficial for learning. *Learning and Instruction, 29*, 153–170.

Eagan, K., Hurtado, S., Figueroa, T., & Hughes, B. E. (2014). *Examining STEM pathways among students who begin college at four-year institutions*. National Academy of Sciences, Engineering, and Medicine.

Freeman, S., Eddy, S. L., McDonough, M., Smith, M. K., Okoroafor, N., Jordt, H., & Wenderoth, M. P. (2014). Active learning increases student performance in science, engineering, and mathematics. *Proceedings of the National Academy of Sciences, 111*(23), 8410–8415. https://doi.org/10.1073/pnas.1319030111

Gao, Y., Xiong, Y., Gao, X., Jia, K., Pan, J., Bi, Y., Dai, Y., Sun, J., Wang, M., & Wang, H. (2023). *Retrieval-augmented generation for large language models: A survey* [Preprint]. arXiv. https://doi.org/10.48550/arXiv.2312.10997

Graesser, A. C., D'Mello, S. K., & Strain, A. C. (2014). Emotions in advanced learning technologies. In R. Pekrun & L. Linnenbrink-Garcia (Eds.), *International handbook of emotions in education* (pp. 473–493). Routledge.

Greene, J. A., & Azevedo, R. (2007). A theoretical review of Winne and Hadwin's model of self-regulated learning: New perspectives and directions. *Review of Educational Research, 77*(3), 334–372. https://doi.org/10.3102/003465430303953

Gupta, P., Bigham, J. P., Tsvetkov, Y., & Pavel, A. (2020). *Controlling dialogue generation with semantic exemplars*[Preprint]. arXiv. https://doi.org/10.48550/arXiv.2008.09075

Hattie, J., & Timperley, H. (2007). The power of feedback. *Review of Educational Research, 77*(1), 81–112. https://doi.org/10.3102/003465430298487

Hawkins, W., Goddard, K., & Favero, C. (2021). A cocurricular program that encourages specific study skills and habits improves academic performance and retention of first-year undergraduates in introductory biology. *CBE—Life Sciences Education, 20*(1), ar4. https://doi.org/10.1187/cbe.20-06-0117

Holstein, K., Aleven, V., & Rummel, N. (2020). A conceptual framework for human–AI hybrid adaptivity in education. In *Artificial intelligence in education* (Vol. 12163, pp. 240–254). Springer. https://doi.org/10.1007/978-3-030-52237-7_20

Hooper, D., Coughlan, J., & Mullen, M. (2008). Evaluating model fit. In *Proceedings of the European Conference on Research Methodology* (pp. 195–200).

Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling, 6*(1), 1–55. https://doi.org/10.1080/10705519909540118

Huang, J., Zhang, Y., Morphew, J., Nissen, J., Van Dusen, B., & Chang, H.-H. (2025). Two-phase content-balancing CD-CAT online item calibration. *Journal of Educational Measurement*.

Huang, W., Jabbari, J., Johnson, O., Jr., & Chun, Y. (2025). Can certificate programs solve the skills and spatial mismatch problem? Job portability and residential mobility in a coding and apprenticeship program. *Urban Education, 61*(3). https://doi.org/10.1177/00420859251315480

Jabbari, J., Huang, W., & Johnson, O., Jr. (2023). Broadening participation in STEM through alternative preparation programs. *Journal of Women and Minorities in Science and Engineering, 29*(6), 1–47. https://doi.org/10.1615/JWomenMinorScienEng.2022041267

Jin, Y., & Sercu, L. (2025). ChatGPT interventions in higher education: A systematic review of experimental studies.*Journal of Computer Assisted Learning, 41*(4), e70072.

Kahneman, D. (2011). *Thinking, fast and slow*. Farrar, Straus and Giroux.

Kestin, G., Miller, K., Klales, A., Milbourne, T., & Ponti, G. (2025). AI tutoring outperforms in-class active learning: An RCT introducing a novel research-based design in an authentic educational setting. *Scientific Reports, 15*(1), Article 17458.

LASSO. (2025). *Learning About STEM Student Outcomes (LASSO)*. https://lassoeducation.org/

Le, V., Nissen, J., Tang, X., Zhang, Y., Mehrabi, A., Morphew, J. W., Chang, H.-H., & Van Dusen, B. (2025). Applying cognitive diagnostic models to mechanics concept inventories. *Physical Review Physics Education Research, 21*(1), 010103. https://doi.org/10.1103/PhysRevPhysEducRes.21.010103

Le, V., Van Dusen, B., Nissen, J., Tang, X., Zhang, Y., Chang, H.-H., & Morphew, J. (2024). Mechanics Cognitive Diagnostic: Mathematics skills tested in introductory physics courses. In *Physics Education Research Conference 2024*.

Le, V., Van Dusen, B., Nissen, J., Tang, X., Zhang, Y., Chang, H.-H., & Morphew, J. (n.d.). *Mechanics Cognitive Diagnostic: Testing fine-grained learning objectives in introductory physics* [Manuscript submitted for publication]. https://doi.org/10.13140/RG.2.2.21044.21129

Leoni, D., Fleming, T., & McFarland, J. L. (2023). Cultivating a science, technology, engineering, and mathematics (STEM) community for two-year college student success and persistence. *PLOS ONE, 18*(9), e0290958. https://doi.org/10.1371/journal.pone.0290958

Létourneau, A., Deslandes Martineau, M., Charland, P., Karran, J. A., Boasen, J., & Léger, P. M. (2025). A systematic review of AI-driven intelligent tutoring systems (ITS) in K–12 education. *npj Science of Learning, 10*, Article 29. https://doi.org/10.1038/s41539-025-00320-7

Lewis, P., Perez, E., Piktus, A., Petroni, F., Karpukhin, V., Goyal, N., Küttler, H., Lewis, M., Yih, W.-T., Rocktäschel, T., Riedel, S., & Kiela, D. (2020). Retrieval-augmented generation for knowledge-intensive NLP tasks. *Advances in Neural Information Processing Systems, 33*, 9459–9474.

Limeri, L. B., Carter, N. T., Lyra, F., Martin, J., Mastronardo, H., Patel, J., & Dolan, E. L. (2023). Undergraduate lay theories of abilities: Mindset, universality, and brilliance beliefs uniquely predict undergraduate educational outcomes. *CBE—Life Sciences Education, 22*(4), ar40. https://doi.org/10.1187/cbe.22-12-0250

Little, R. J., & Yau, L. H. Y. (1998). Statistical techniques for analyzing data from prevention trials: Treatment of no-shows using Rubin's causal model. *Psychological Methods, 3*(2), 147–159. https://doi.org/10.1037/1082-989X.3.2.147

Liu, L., Wang, P., Wu, G., Jiang, J., & Yang, H. (2025). Towards optimal mixture of experts system for 3D object detection: A game of accuracy, efficiency, and adaptivity. *IEEE Transactions on Pattern Analysis and Machine Intelligence*.

Ma, W., Adesope, O. O., Nesbit, J. C., & Liu, Q. (2014). Intelligent tutoring systems and learning outcomes: A meta-analysis. *Journal of Educational Psychology, 106*(4), 901–918. https://doi.org/10.1037/a0037123

Maas, L., Brinkhuis, M. J., Kester, L., & Wijngaards-de Meij, L. (2022). Cognitive diagnostic assessment in university statistics education: Valid and reliable skill measurement for actionable feedback using learning dashboards. *Applied Sciences, 12*(10), 4809. https://doi.org/10.3390/app12104809

McKenzie, L. (2024, October 18). Science leaders prepare pitch for National Defense Education Act 2.0. *American Institute of Physics*. https://www.aip.org/fyi/science-leaders-prepare-pitch-for-national-defense-education-act-2-0

Miles, M. B., Huberman, A. M., & Saldaña, J. (2018). *Qualitative data analysis: A methods sourcebook*. SAGE.

Morley, A., Nissen, J., & Van Dusen, B. (2023). Measurement invariance across race and gender for the Force Concept Inventory. *Physical Review Physics Education Research, 19*, 020102. https://doi.org/10.1103/PhysRevPhysEducRes.19.020102

Morphew, J., Mehrabi, A., Van Dusen, B., Nissen, J., & Chang, H. (2024). AI-based concept inventories: Using cognitive diagnostic computer adaptive testing in LASSO for classroom assessment. In *2024 ASEE Annual Conference & Exposition*.

Morris, R., Perry, T., & Wardle, L. (2021). Formative assessment and feedback for learning in higher education: A systematic review. *Review of Education, 9*(3), e3292. https://doi.org/10.1002/rev3.3292

Mousavinasab, E., Zarifsanaiey, N., Niakan Kalhori, S. R., Rakhshan, M., Keikha, L., & Ghazi Saeedi, M. (2021). Intelligent tutoring systems: A systematic review of characteristics, applications, and evaluation methods. *Interactive Learning Environments, 29*(1), 142–163. https://doi.org/10.1080/10494820.2018.1558257

National Academies of Sciences, Engineering, and Medicine. (2024). *International talent programs in the changing global environment*. The National Academies Press.

National Research Council. (2012). *Discipline-based education research: Understanding and improving learning in undergraduate science and engineering*. The National Academies Press. https://doi.org/10.17226/13362

National Science Board. (2018). *Science and engineering indicators 2018* (NSB-2018-1). National Science Foundation. https://www.nsf.gov/statistics/2018/nsb20181/

National Science Board. (2020). *Vision 2030* (NSB-2020-15). https://www.nsf.gov/nsb/publications/2020/nsb202015.pdf

National Science Board. (2021). *The STEM labor force of today: Scientists, engineers, and skilled technical workers*(NSB-2021-2). National Science Foundation. https://ncses.nsf.gov/pubs/nsb20212

Nguyen, N. N., & Barbieri, W. (2025). Mentorship in the age of generative AI: ChatGPT to support self-regulated learning of pre-service teachers before and during placements. *Education Sciences, 15*(6), 642. https://doi.org/10.3390/educsci15060642

Nickow, A., Oreopoulos, P., & Quan, V. (2020). *The impressive effects of tutoring on PreK–12 learning: A systematic review and meta-analysis of the experimental evidence* (NBER Working Paper No. 27476). National Bureau of Economic Research. https://doi.org/10.3386/w27476

Nissen, J. M., Jariwala, M., Close, E. W., & Van Dusen, B. (2018). Participation and performance on paper- and computer-based low-stakes assessments. *International Journal of STEM Education, 5*(1), Article 21.

Oliveira, A. W., Musah, R. A., Zacharia, Z. C., Pavlou, Y., Sgro, C. M., Lathrop, R., Kim, Y., Espinosa, K., Japashov, N., & Park, J. (2025). The persistent problem of undergraduate student attrition in STEM. *International Educational Review, 3*(1), 85–111. https://doi.org/10.58693/ier.316

Panadero, E. (2017). A review of self-regulated learning: Six models and four directions for research. *Frontiers in Psychology, 8*, Article 422. https://doi.org/10.3389/fpsyg.2017.00422

Park, C. L., Williams, M., Hernandez, P., Agocha, V. B., Carney, L. M., DePetris, A., & Lee, S. Y. (2019). Self-regulation and STEM persistence in minority and non-minority students across the first year of college. *Social Psychology of Education, 22*(1), 91–112. https://doi.org/10.1007/s11218-018-9465-7

Patton, M. Q., & Campbell-Patton, C. E. (2022). *Utilization-focused evaluation* (5th ed.). SAGE.

Peugh, J. L., Strotman, D., McGrady, M., Rausch, J., & Kashikar-Zuck, S. (2017). Beyond intent to treat (ITT): A complier average causal effect (CACE) estimation primer. *Journal of School Psychology, 60*, 7–24. https://doi.org/10.1016/j.jsp.2015.12.006

Radeva, I., Popchev, I., & Dimitrova, M. (2024). Similarity thresholds in RAG. In *IEEE Intelligent Systems Conference*.

Ratkowsky, D., Alldredge, R., & Evans, M. A. (2020). *Cross-over experiments: Design, analysis, and application*. CRC Press.

Ravand, H., & Robitzsch, A. (2015). Cognitive diagnostic modeling using R. *Practical Assessment, Research & Evaluation, 20*, Article 11.

Ritter, S., Anderson, J. R., Koedinger, K. R., & Corbett, A. (2007). Cognitive Tutor: Applied research in mathematics education. *Psychonomic Bulletin & Review, 14*, 249–255. https://doi.org/10.3758/BF03194060

Roberge, K., Le, V., Van Dusen, B., & Nissen, J. (2025). Calculus Cognitive Diagnostic: Mathematics skills tested in first-semester calculus courses. In *Physics Education Research Conference 2025* (pp. 368–373). https://doi.org/10.1119/perc.2025.pr.Roberge

Rohilla, D. A. (2025). Impact of excessive AI tool usage on the cognitive abilities of undergraduate students: A mixed methods study. *Advance Social Science Archive Journal, 4*(1), 2131–2143. https://doi.org/10.55966/assaj.2025.4.1.0115

Rubin, D. B. (2005). Causal inference using potential outcomes: Design, modeling, decisions. *Journal of the American Statistical Association, 100*, 322–331. https://doi.org/10.1198/016214504000001880

Salazar Morales, A., Nissen, J., Van Dusen, B., & Baran, E. (n.d.). *Open-ended graph literacy assessment for educators: Developing and validating an open-response instrument to measure educators' graph literacy* [Manuscript submitted for publication]. https://www.researchgate.net/publication/406994321_Open-Ended_Graph_Literacy_Assessment_For_Educators_Developing_And_Validating_An_Open_Response_Instrument_To_Measure_Educators%27_Graph_Literacy

Sandoval, W. A. (2014). Conjecture mapping: An approach to systematic educational design research. *Journal of the Learning Sciences, 23*(1), 18–36. https://doi.org/10.1080/10508406.2013.778204

Schaffer, H. E., Young, K. R., Ligon, E. W., & Chapman, D. D. (2017). Automating individualized formative feedback in large classes based on a directed concept graph. *Frontiers in Psychology, 8*, Article 260. https://doi.org/10.3389/fpsyg.2017.00260

Schunk, D. H., & Zimmerman, B. J. (2013). Self-regulation and learning. In W. M. Reynolds, G. E. Miller, & I. B. Weiner (Eds.), *Handbook of psychology: Educational psychology* (2nd ed., Vol. 7, pp. 45–68). Wiley.

Sebesta, A. J., & Bray Speth, E. (2017). How should I study for the exam? Self-regulated learning strategies and achievement in introductory biology. *CBE—Life Sciences Education, 16*(2), ar30. https://doi.org/10.1187/cbe.16-09-0269

Shadish, W. R., Cook, T. D., & Campbell, D. T. (2002). *Experimental and quasi-experimental designs for generalized causal inference*. Houghton Mifflin.

Smith, J., Wilson, S. B., Banks, J., Zhu, L., & Varma-Nelson, P. (2014). Replicating peer-led team learning in cyberspace: Research, opportunities, and challenges. *Journal of Research in Science Teaching, 51*(6), 714–740. https://doi.org/10.1002/tea.21163

Steinert, S., Avila, K. E., Ruzika, S., Kuhn, J., & Küchemann, S. (2024). Harnessing large language models to develop research-based learning assistants for formative feedback. *Smart Learning Environments, 11*(1), Article 62. https://doi.org/10.1186/s40561-024-00354-1

Tang, F., & Zhan, P. (2021). Does diagnostic feedback promote learning? Evidence from a longitudinal cognitive diagnostic assessment. *AERA Open, 7*, 23328584211060804. https://doi.org/10.1177/23328584211060804

Trujillo, G., & Tanner, K. D. (2014). Considering the role of affect in learning: Monitoring students' self-efficacy, sense of belonging, and science identity. *CBE—Life Sciences Education, 13*(1), 6–15. https://doi.org/10.1187/cbe.13-12-0241

U.S. Department of Education, Office of Educational Technology. (2023). *Artificial intelligence and the future of teaching and learning: Insights and recommendations*. https://tech.ed.gov/files/2023/05/ai-future-of-teaching-and-learning-report.pdf

Van Dusen, B. (2018). *LASSO: A new tool to support instructors and researchers* [Preprint]. arXiv. https://doi.org/10.48550/arXiv.1812.02299

Van Dusen, B., Shultz, M., Nissen, J. M., Wilcox, B. R., Holmes, N. G., Jariwala, M., Close, E. W., Lewandowski, H. J., & Pollock, S. (2021). Online administration of research-based assessments. *American Journal of Physics, 89*(1), 7–8.

VanLehn, K. (2011). The relative effectiveness of human tutoring, intelligent tutoring systems, and other tutoring systems. *Educational Psychologist, 46*(4), 197–221. https://doi.org/10.1080/00461520.2011.611369

Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological processes* (M. Cole, V. John-Steiner, S. Scribner, & E. Souberman, Eds.). Harvard University Press.

Warrens, M. J. (2015). Five ways to look at Cohen's kappa. *Journal of Psychology & Psychotherapy, 5*(4), Article 197. https://doi.org/10.4172/2161-0487.1000197

Wilson, S. B., & Varma-Nelson, P. (2016). Small groups, significant impact: A review of peer-led team learning research with implications for STEM education researchers and faculty. *Journal of Chemical Education, 93*(10), 1686–1702.

Winne, P. H. (2021). Open learner models working in symbiosis with self-regulating learners: A research agenda. *International Journal of Artificial Intelligence in Education, 31*(3), 446–459. https://doi.org/10.1007/s40593-020-00212-4

Wood, D., Bruner, J. S., & Ross, G. (1976). The role of tutoring in problem solving. *Journal of Child Psychology and Psychiatry, 17*(2), 89–100. https://doi.org/10.1111/j.1469-7610.1976.tb00381.x

Xu, Z., Zhao, Y., Zhang, B., Liew, J., & Kogut, A. (2023). A meta-analysis of the efficacy of self-regulated learning interventions on academic achievement in online and blended environments in K–12 and higher education. *Behaviour & Information Technology, 42*(16), 2911–2931. https://doi.org/10.1080/0144929X.2022.2151935

Zhang, J., Sun, J., Yeats, E., Ouyang, Y., Kuo, M., Zhang, J., Yang, H. F., & Li, H. (2025). Min-k%++: Improved baseline for pre-training data detection from large language models. In *International Conference on Learning Representations*.

Zhao, Y., Wang, P., & Yang, H. F. (2025). How to auto-optimize prompts for domain tasks? Adaptive prompting and reasoning through evolutionary domain knowledge adaptation. In *Advances in Neural Information Processing Systems*.

Zhao, Y., Zhao, Y., Du, H., & Yang, H. F. (2025). Personalized decision modeling: Utility optimization or textualized-symbolic reasoning. In *Advances in Neural Information Processing Systems* [Conference spotlight].

Zimmerman, B. J. (2002). Becoming a self-regulated learner: An overview. *Theory Into Practice, 41*(2), 64–70. https://doi.org/10.1207/S15430421TIP4102_2

Data Management Plan

This project will follow the U.S. Department of Education's expectations for responsible data stewardship, protection of student information, transparent documentation, and long-term accessibility of de-identified datasets. All activities will comply with FERPA, IRB protocols at Iowa State University (ISU) and partner institutions, and the established governance model of the LASSO platform. No personally identifiable information (PII) will be shared outside ISU, and only de-identified datasets will be used for research analyses.

1. Types of Data Collected

The project will generate three categories of data:

1. Student-Level Data

- Assessment responses from fixed-form and adaptive instruments
- IRT proficiency estimates, item-level metadata, and skill mastery profiles
- AI Companion interaction logs (text, turn-by-turn timestamps)
- AI-generated learning analytics Metadata (e.g., interaction patterns, hint–answer ratios, error profiles, strategy-use indicators), produced by the Companion's personalization algorithms
- Time-on-task, feature usage, and session-level engagement data
- Student-provided demographics (e.g., race/ethnicity, gender, academic background), consistent with LASSO's standard FERPA-aligned consent

2. Instructor-Level and Course Data

- Course structure, learning objectives, and assignment metadata
- Instructor onboarding surveys and optional interviews

- Classroom implementation logs, dashboard usage, and feedback
- Aggregated section-wide summaries of student performance

3. System Metadata

- Platform activity logs (assessment delivery, CAT item selection, scoring events)
- AI Companion system logs (retrieval history, metadata generation timestamps, routing decisions)
- Server performance logs used for reliability and reproducibility

All AI Companion data will be stored only in ISU-secured environments, and model API calls will be generated from ISU servers. Only anonymized, ephemeral representations of student data will be included in the API request.

2. Data Formats and Documentation

Data will be stored in structured, documented formats:

- **Quantitative data**: CSV, JSON, SQL database exports
- **Qualitative data**: TXT/DOCX transcripts
- **Metadata**: data dictionaries, README files, variable schemas, version notes
- **Derived analytics**: structured JSON and SQL tables describing AI-generated metadata

Documentation will follow existing standards used by LASSO assessment datasets, including consistent naming conventions, embedded metadata, and open data dictionaries used in prior NSF-funded projects.

3. Storage, Security, and Access Controls

All raw and identifiable data will be stored on encrypted ISU servers managed within the LASSO infrastructure. Security protections include:

- AES-256 encryption at rest and in transit

- Role-based access control (RBAC)
- Two-factor authentication
- Secure VPN and VPC isolation
- Audit logs of all data access

The AI Companion will run entirely through ISU-orchestrated API calls, ensuring that no student identifiers or sensitive materials are transmitted to external providers. The system will send only anonymized and minimal context required for generating responses. No AI interaction data will be used to train external models.

Access to identifiable data is limited to IRB-approved research personnel. Collaborators outside ISU will receive only de-identified datasets approved for analysis.

4. Data Sharing, Reuse, and Licensing

Because this project includes proprietary tutoring algorithms, prompt structures, and AI-generated analytics that are integrated into LASSO's infrastructure, data sharing must be controlled.

Accordingly:

- Only **de-identified datasets** will be eligible for sharing with external researchers
- Sharing will occur **on request**, under a controlled-access agreement specifying research scope, permitted use, security expectations, and citation requirements
- No identifiable data, interaction logs, or institutional metadata will be publicly released
- Proprietary components—including prompt orchestration logic, metadata-generation algorithms, and routing strategies—will not be shared
- No datasets will be posted to public repositories

This controlled-access model balances transparency with student privacy, institutional requirements, and protection of project intellectual property.

5. Ethical Compliance

All data collection and use will be reviewed and approved by ISU's Institutional Review Board. Students will provide:

1. **Standard FERPA-aligned LASSO consent**, and
2. **An additional AI-specific disclosure** describing the collection of interaction logs and AI-generated metadata.

All datasets used for research and dissemination will be de-identified before analysis.

6. Long-Term Preservation

All de-identified datasets—assessment responses, derived IRT variables, AI-generated metadata, and logs suitable for replication—will be archived annually and stored within the **permanent LASSO de-identified archive at Iowa State University**. ISU will serve as the long-term steward of the data.

Data retained for replication or future analyses will follow the LASSO platform's permanent archival procedures, ensuring stable access and documentation long after the project's completion.

This plan ensures that data are collected, stored, analyzed, and shared responsibly while preserving student privacy, meeting FERPA and IRB requirements, and protecting the project's intellectual property.

U.S. Department of Education
Evidence Form

OMB Number: 1894-0001
Expiration Date: 08/31/2028

1. Level of Evidence

Select the level of evidence of effectiveness for which you are applying. See the Notice Inviting Applications for the relevant definitions and requirements.

☒ Demonstrates a Rationale ☐ Promising Evidence ☐ Moderate Evidence ☐ Strong Evidence

2. Citation and Relevance

Fill in the chart below with the appropriate information about the studies that support your application.

A. Research/Citation	B. Relevant Outcome(s)/Relevant Finding(s)	C. Project Component(s)/Overlap of Populations and/or Settings
Nickow, A., Oreopoulos, P., & Quan, V. (2020). The impressive effects of tutoring on PreK–12 learning: A systematic review and meta-analysis of the experimental evidence (NBER Working Paper No. 27476). National Bureau of Economic Research. https://doi.org/10.3386/w27476; Ma, W., Adesope, O. O., Nesbit, J. C., & Liu, Q. (2014). Intelligent tutoring systems and learning outcomes: A meta-analysis. Journal of Educational Psychology, 106(4), 901–918. https://doi.org/10.1037/a0037123; Kestin, G., Miller, K., Klales, A., Milbourne, T., & Ponti, G. (2025). AI tutoring outperforms in-class active learning: An RCT introducing a novel research-based design in an authentic educational setting. Scientific Reports, 15(1), Article 17458	More students pass gateway courses, greater retention and completion	AI tutoring that guides rather than answers
Black, P., & Wiliam, D. (1998). Assessment and classroom learning. Assessment in Education: Principles, Policy & Practice, 5(1), 7–74. https://doi.org/10.1080/0969595980050102; Hattie, J., & Timperley, H. (2007). The power of feedback. Review of Educational Research, 77(1), 81–112. https://doi.org/10.3102/003465430298487; Freeman, S., Eddy, S. L., McDonough, M., Smith, M. K., Okoroafor, N., Jordt, H., & Wenderoth, M. P. (2014). Active learning increases student performance in science, engineering, and mathematics. Proceedings of the National Academy of Sciences, 111(23), 8410–8415. https://doi.org/10.1073/pnas.1319030111	Fewer failures and withdrawals, greater credit accumulation and retention.	Frequent feedback on student work
Tang, F., & Zhan, P. (2021). Does diagnostic feedback promote learning? Evidence from a longitudinal cognitive diagnostic assessment. AERA Open, 7, 23328584211060804. https://doi.org/10.1177/23328584211060804; Maas, L., Brinkhuis, M. J., Kester, L., & Wijngaards-de	Strong mastery, greater course completion and progression	Skill-by-skill diagnostic feedback

Meij, L. (2022). Cognitive diagnostic assessment in university statistics education: Valid and reliable skill measurement for actionable feedback using learning dashboards. Applied Sciences, 12(10), 4809. https://doi.org/10.3390/app12104809		
Zimmerman, B. J. (2002). Becoming a self-regulated learner: An overview. Theory Into Practice, 41(2), 64–70. https://doi.org/10.1207/S15430421TIP4102_2; Schunk, D. H., & Zimmerman, B. J. (2013). Self-regulation and learning. In W. M. Reynolds, G. E. Miller, & I. B. Weiner (Eds.), Handbook of psychology: Educational psychology (2nd ed., Vol. 7, pp. 45–68). Wiley.	Persistence and completion, especially for at-risk students	Support that builds self-direction and fades over time
Nguyen, N. N., & Barbieri, W. (2025). Mentorship in the age of generative AI: ChatGPT to support self-regulated learning of pre-service teachers before and during placements. Education Sciences, 15(6), 642. https://doi.org/10.3390/educsci15060642;Anders, A. D., & Dux Speltz, E. (2025). Developing generative AI literacies through self-regulated learning: A human-centered approach. Computers and Education: Artificial Intelligence, 9, 100482. https://doi.org/10.1016/j.caeai.2025.100482	Productive AI use, greater progression and retention	Building AI literacy and responsible use

Instructions for Evidence Form

1. **Level of Evidence.** Check the box next to the level of evidence for which you are applying. See the Notice Inviting Applications for the evidence definitions.

2. **Citation and Relevance.** Fill in the chart for each of the studies you are submitting to meet the evidence standards. If allowable under the program you are applying for, you may add additional rows to include more than four citations. (See below for an example citation.)

   a. **Research/Citation.** For Demonstrates a Rationale, provide the citation or link for the research or evaluation findings. For Promising, Moderate, and Strong Evidence, provide the full citation for each study or WWC publication you are using as evidence. If the study has been reviewed by the WWC, please include the rating it received, the WWC review standards version, and the URL link to the description of that finding in the WWC reviewed studies database. Include a copy of the study or a URL link to the study, if available. Note that, to provide promising, moderate, or strong evidence, you must cite either a specific recommendation from a WWC practice guide, a WWC intervention report, or a publicly available, original study of the effectiveness of a component of your proposed project on a student outcome or other relevant outcome.

   b. **Relevant Outcome(s)/Relevant Finding(s).** For Demonstrates a Rationale, describe how the research or evaluation findings suggest that the project component included in the logic model is likely to improve relevant outcomes. For Promising, Moderate and Strong Evidence, describe: 1) the project component included in the study (or WWC practice guide or intervention report) that is also a component of your proposed project, 2) the student outcome(s) or other relevant outcome(s) that are included in both the study (or WWC practice guide or intervention report) and in the logic model (theory of action) for your proposed project, and 3) the study (or WWC intervention report) finding(s) or WWC practice guide recommendations supporting a favorable relationship between a project component and a relevant outcome. Cite page and table numbers from the study (or WWC practice guide or intervention report), where applicable.

   c. **Project Component(s)/Overlap of Population and/or Settings.** For Demonstrates a Rationale, explain how the project component(s) is informed by the research or evaluation findings. For Promising, Moderate, and Strong Evidence, explain how the population and/or setting in your proposed project are similar to the populations and settings included in the relevant finding(s). Cite page numbers from the study or WWC publication, where applicable.

EXAMPLES: For Demonstration Purposes Only (the three examples are not assumed to be cited by the same applicant)

A. Research/Citation	B. Relevant Outcome(s)/Relevant Finding(s)	C. Project Component(s)/Overlap of Populations and/or Settings
Graham, S., Bruch, J., Fitzgerald, J., Friedrich, L., Furgeson, J., Greene, K., Kim, J., Lyskawa, J., Olson, C.B., & Smither Wulsin, C. (2016). Teaching secondary students to write effectively (NCEE 2017-4002). Washington, DC: National Center for Education Evaluation and Regional Assistance (NCEE), Institute of Education Sciences, U.S. Department of Education. Retrieved from the NCEE website: https://ies.ed.gov/ncee/wwc/PracticeGuide/22. This report was prepared under Version 3.0 of the WWC Handbook (p. 72).	*(Table 1, p. 4) Recommendation 1 ("Explicitly teach appropriate strategies using a Model – Practice – Reflect instructional cycle") is characterized as backed by "strong evidence."* *(Appendix D, Table D.2, pp. 70-72) Studies contributing to the "strong evidence" supporting the effectiveness of Recommendation 1 reported statistically significant and positive impacts of this practice on genre elements, organization, writing output, and overall writing quality.*	*(Appendix D, Table D.2, pp. 70-72) Studies contributing to the "strong evidence" supporting the effectiveness of Recommendation 1 were conducted on students in grades 6 through 12 in urban and suburban school districts in California and in the Mid-Atlantic region of the U.S. These study samples overlap with both the populations and settings proposed for the project.*

A. Research/Citation	B. Relevant Outcome(s)/Relevant Finding(s)	C. Project Component(s)/Overlap of Populations and/or Settings
U.S. Department of Education, Institute of Education Sciences, What Works Clearinghouse. (2017, February). Transition to College intervention report: Dual Enrollment Programs. Retrieved from https://ies.ed.gov/ncee/wwc/Intervention/1043. This report was prepared under Version 3.0 of the WWC Handbook (p. 1).	*(Table 1, p. 2) Dual enrollment programs were found to have positive effects on students' high school completion, general academic achievement in high school, college access and enrollment, credit accumulation in college, and degree attainment in college, and these findings were characterized by a "medium to large" extent of evidence.*	*(pp. 1, 19, 22) Studies contributing to the effectiveness rating of dual enrollment programs in the high school completion, general academic achievement in high school, college access and enrollment, credit accumulation in college, and degree attainment in college domains were conducted in high schools with minority students representing between 32 and 54 percent of the student population and first generation college students representing between 31 and 41 percent of the student population. These study samples overlap with both the populations and settings proposed for the project.*
Bettinger, E.P., & Baker, R. (2011). The effects of student coaching in college: An evaluation of a randomized experiment in student mentoring. Stanford, CA: Stanford University School of Education. Available at https://ed.stanford.edu/sites/default/files/bettinger_baker_030711.pdf *Meets WWC Group Design Standards without Reservations under review standards 2.1 (http://ies.ed.gov/ncee/wwc/Study/72030).*	*The intervention in the study is a form of college mentoring called student coaching. Coaches helped with a number of issues, including prioritizing student activities and identifying barriers and ways to overcome them. Coaches were encouraged to contact their assignees by either phone, email, text messaging, or social networking sites (pp. 8-10). The proposed project for Alpha Beta Community College students will train professional staff and faculty coaches on the most effective way(s) to communicate with their mentees, suggest topics for mentors to talk to their mentees, and be aware of signals to prevent withdrawal or academic failure.* *The relevant outcomes in the study are student persistence and degree completion (Table 3, p. 27), which are also included in the logic model for the proposed project.* *This study found that students assigned to receive coaching and mentoring were significantly more likely than students in the comparison group to remain enrolled at their institutions (pp. 15-16, and Table 3, p. 27).*	*The full study sample consisted of "13,555 students across eight different higher education institutions, including two- and four-year schools and public, private not-for-profit, and proprietary colleges." (p. 10) The number of students examined for purposes of retention varied by outcome (Table 3, p. 27). The study sample overlaps with Alpha Beta Community College in terms of both postsecondary students and postsecondary settings.*

DISCLOSURE OF LOBBYING ACTIVITIES

Complete this form to disclose lobbying activities pursuant to 31 U.S.C.1352

OMB Number: 4040-0013
Expiration Date: 06/30/2028

1. * Type of Federal Action:

- [ ] a. contract
- [x] b. grant
- [ ] c. cooperative agreement
- [ ] d. loan
- [ ] e. loan guarantee
- [ ] f. loan insurance

2. * Status of Federal Action:

- [ ] a. bid/offer/application
- [x] b. initial award
- [ ] c. post-award

3. * Report Type:

- [x] a. initial filing
- [ ] b. material change

4. Name and Address of Reporting Entity:

- [x] Prime
- [ ] SubAwardee

*Name: Johns Hopkins University

*Street 1: 3400 North Charles Street

Street 2:

*City: Baltimore

State: MD: Maryland

Zip: 212182625

Congressional District, if known: MD-007

5. If Reporting Entity in No.4 is Subawardee, Enter Name and Address of Prime:

6. * Federal Department/Agency:

US Dept of Labor on behalf of Dept of Ed

7. * Federal Program Name/Description:

Fund for the Improvement of Postsecondary Education

Assistance Listing Number, *if applicable:* 84.116

8. Federal Action Number, *if known:*

9. Award Amount, *if known:*

$

10. a. Name and Address of Lobbying Registrant:

Prefix:

*First Name: na

Middle Name:

*Last Name: na

Suffix:

*Street 1: na

Street 2:

*City: na

State:

Zip:

b. Individual Performing Services (including address if different from No. 10a)

Prefix:

*First Name: na

Middle Name:

*Last Name: na

Suffix:

*Street 1: na

Street 2:

*City: na

State:

Zip:

11. Information requested through this form is authorized by title 31 U.S.C. section 1352. This disclosure of lobbying activities is a material representation of fact upon which reliance was placed by the tier above when the transaction was made or entered into. This disclosure is required pursuant to 31 U.S.C. 1352. This information will be reported to the Congress semi-annually and will be available for public inspection. Any person who fails to file the required disclosure shall be subject to a civil penalty of not less than $10,000 and not more than $100,000 for each such failure.

*** Signature:** Completed on submission to Grants.gov

***Name:**

Prefix:

*First Name: Denise

Middle Name:

*Last Name: Sparks

Suffix:

Title: Senior Grants Associate

Telephone No.: 667-208-8806

Date: Completed on submission to Grants.gov

Federal Use Only:

Authorized for Local Reproduction
Standard Form - LLL (Rev. 7-97)