Examining Linguistic Diversity and Data Science Curriculum in STEM Education with a Justice, Equity, Diversity, and Inclusion Lens

Manisha Sharan, Ph.D.
Vandana Janeja, Ph.D.

This material is based upon work supported by the National Science Foundation (NSF) under Grant No. DUE-1937267. Any opinions, findings, interpretations, conclusions or recommendations expressed in this material are those of its authors and do not represent the views of the AAAS Board of Directors, the Council of AAAS, AAAS' membership or the National Science Foundation.
This presentation is being recorded. The recording and slides will be available in the coming week at https://aaas-iuse.org

Please note: The discussion break-out groups following the presentations will NOT be recorded.
Closed Captioning:

To download the transcript or view the full transcript, please click the “CC” button at the bottom of your screen.
The AAAS-IUSE initiative supports faculty, students, and the greater undergraduate STEM education community by disseminating research and knowledge about STEM teaching, learning, equity and institutional transformation.

Check out our website to learn more and view: https://aaas-iuse.org

- Blog
- Workshops
- Summer Labs On-Demand
- Resources
- Lessons Learned During COVID
- NSF IUSE Proposal Preparation Toolkit

@IuseProgram

AAAS IUSE Initiative
Examining Linguistic Diversity and Data Science Curriculum in STEM Education with a Justice, Equity, Diversity, and Inclusion Lens

Manisha Sharan, Ph.D.
Vandana Janeja, Ph.D.
Teaching In a Linguistically Diverse Classroom

Dr. Manisha Sharan, PhD.
Professor, Baker College
22nd March 2023
OBJECTIVES

1. What is Linguistic Diversity?
2. Limited English Proficiency
3. Need/Strategies for Content-Based Language Support & Development
Teaching and Diversity

Classroom Diversity on the Rise

School Enrollment Percentages by Race and Hispanic Origin

<table>
<thead>
<tr>
<th>Race and Hispanic Origin</th>
<th>K-12</th>
<th>College</th>
</tr>
</thead>
<tbody>
<tr>
<td>White, non-Hispanic</td>
<td>57.6%</td>
<td>54.7%</td>
</tr>
<tr>
<td>Hispanic</td>
<td>19.9%</td>
<td>12.1%</td>
</tr>
<tr>
<td>Black, non-Hispanic</td>
<td>14.9%</td>
<td>14.5%</td>
</tr>
<tr>
<td>Asian, non-Hispanic</td>
<td>4.1%</td>
<td>6.1%</td>
</tr>
<tr>
<td>Other, non-Hispanic</td>
<td>3.6%</td>
<td>2.3%</td>
</tr>
</tbody>
</table>

www.census.gov/programs-surveys/cps.html
Sofa or Couch
or chesterfield or davenport or settee or divan?
Power of Vocabulary: Wrong/Different

- End Goal is teaching and learning
- What about GPA?
- Power lies with the authority/faculty/curriculum developer/publisher
What is Linguistic Diversity?

- Education: Business of sharing knowledge
- **Language**: Vocabulary is foundational for learning
- Language is an important navigational tool
- Speak many languages/not speak main language
- **Language variation**: how a word is used - Voice
The Math class!

Solve

1 gallon = 4 quart
2 gallon = ___ quart

Ratio/Proportion Method
Dimensional Analysis
Conversion Factor Method

➢ Respond in CHAT what method do you call it.
Cohort at First Entry

[Bar chart showing percentages for Asian, Black, Hispanic, and White cohorts in different age groups.]

https://nscresearchcenter.org/signaturereport12-supplement-2/
Language of the Online

- Non-verbal communication is a challenge
- Language intensive
- Lexicon driven

SMS vs Text
Linguistic Diversity Drivers

- ESL,
- Age,
- Geographical,
- Cultural
- Experiential
- What other metric can contribute to linguistic diversity?
Start of my Linguistic Journey

- ____ assumes shape of the container.

✓ ____ assumes takes shape of the container.

“To assume is to suppose without proof.”
Content-Based Language Inclusion in my curriculum

“We don’t speak the way textbook authors and other authors write these days.”
- Nonie K. Lesaux

Solids
– have **rigid** shape

Solids
– have **rigid**, **fixed**, **definite** shape
Limited English Proficiency

1. LEP (purple line) ~25 million;
2. 9% US population.
3. 18% native born!

✓ Linguistically inclusive culture
Challenges for the LEP: NGSS

Next Generation Science Standards

How we do it

Pros:
- Students learn critical thinking skills
- Phenomenon based = **Language** intensive
- K-12: Support programs like ESL / IEP / WIDA

How NGSS does it
• Future undergraduate: NGSS educated
• K-12: have support structures for language/content skill gaps
• Higher Ed: Lack of formal structures or scaffolds
• Student suddenly “On Their OWN”
Linguistically Inclusive Class

• Building Inclusive culture:
  • Know your students
  • Every form of learning is welcome
  • Intentionally structured class

• Content Creation:
  • Limited English Proficiency awareness
  • Thoughtful Content Creation
  • Disciplinary Reading skill development
Linguistically Inclusive Class

- Assessment:
  - **Goal of assessment**: English language skills or concept-skill?
  - Multiple forms of response

- What other ways can you include in your content
• Institutional culture of inclusion and equity
• Curriculum Development: Lexicon Uniformity
• Inter-departmental collaboration for creating scaffolding curriculum for better student learning.
• Formal Support Structures for student and faculty.
Why Linguistically Inclusive matters..

And she cried....

3+1=?

- Feeling linguistically inadequate
- Students with high confidence in their language abilities (i.e., high language-efficacy), engage more
- Is the classroom a safe and inclusive space?

Joanna had six balloons. Two popped. How many does she have left?
My Proposal
Content Based Language Strategy

➢ Content-based language be part of Curriculum development and implementation
➢ Identification of core language skills in each curriculum with content expert. Eg. Voice in Writing
➢ Structures eg. Advising to ensure these core language skills are supported
➢ Inter-departmental collaboration E.g. ELA/Math faculty
Yay SpongeBob hey !*

*This is SpongeBob!
- in Hindi language

Just because we can read it
Doesn’t mean we can understand it!
Data Science Curriculum in STEM Education
Prioritizing Diversity, Equity, and Inclusion.

Vandana P. Janeja
University of Maryland, Baltimore County

Collaborators
Susan Wang, Mills College at Northeastern University
David Harding, Claudia von Vacano
University of California, Berkeley

National Science Foundation award 1915714
Data Science Ecosystem @UMBC

- Foundations of Data Science
- Access to Curriculum
- Peer mentoring and Professional Development

UMBC
Foundations of Data Science
Adoption and Adaptations

- Adapted Foundations course: University of California, Berkeley’s Data 8 Foundations of Data Science (Foundations)
- Translated Foundations to two other institutions with different student populations:
  - University of Maryland, Baltimore County
    IS 296 Foundations of Data Science
  - Mills College in Oakland, California
    Data 80A Data Science for Everyone
- Tailored the course to each institution’s student populations’ interests, backgrounds, and needs

S. Wang, V. Janeja, D. Harding, C. Von Vacano, D. Lobo, adapting data science curricula: a student-centric evaluation, 17th annual
## Comparative View of the Three Institutions

<table>
<thead>
<tr>
<th>Institution</th>
<th>UC Berkeley</th>
<th>UMBC</th>
<th>Mills</th>
</tr>
</thead>
<tbody>
<tr>
<td>Carnegie classification</td>
<td>R1</td>
<td>R1</td>
<td>Liberal arts</td>
</tr>
<tr>
<td>Public/Private</td>
<td>Public</td>
<td>Public</td>
<td>Private</td>
</tr>
<tr>
<td>Total undergrad student body</td>
<td>31,800</td>
<td>10,835</td>
<td>660</td>
</tr>
<tr>
<td>Average Foundations class size</td>
<td>1500</td>
<td>18</td>
<td>20</td>
</tr>
<tr>
<td>Minority serving status</td>
<td>[no]</td>
<td>MSI</td>
<td>HSI</td>
</tr>
<tr>
<td>URM population</td>
<td>22%</td>
<td>52%</td>
<td>65%</td>
</tr>
<tr>
<td>First-gen population</td>
<td>23%</td>
<td>25%</td>
<td>44%</td>
</tr>
<tr>
<td>Female identifying undergrads</td>
<td>54%</td>
<td>44%</td>
<td>100%</td>
</tr>
</tbody>
</table>

**UMBC**
- Maryland’s public research university
- Diverse student body; Minority Serving Institution
UMBC Adaptation

- The UMBC adaptation took into consideration
  - the student body and their backgrounds,
  - the number of credits/hours they can dedicate to studying per week with additional workload that they might have to balance.
- Lowered the number of credits to three with limited contact hours to fit the course into existing frameworks
- Included it on a pathway to existing degree requirements
  - Fulfil a programming requirement for BTA major
  - Became part of the X+CS effort
  - new Business Analytics certificate
  - Gen Ed course.
Infrastructure

• Setting up the JupyterHub infrastructure was non-trivial
• The first run-through had challenges with versioning and interdependencies of Python libraries as well as security and authentication issues.
• Once the infrastructure was set up and running, the students’ experiences with authentication accessing files, and working in Jupyter notebooks, were smooth, easy, and seamless
• The trick to having a strong working infrastructure was getting the right expertise at the right time, and the right amount of time from experts
UMBC Implementation and Lessons Learned

- Developed, refined and offered IS 296 Foundations of Data Science course, since Spring 2020, 6th offering in Spring 2023
- The course attracts majors across the three colleges at UMBC, smaller class with enrollment typically maxed at 25 students
- IS 296 adaptations: 3 credit with reduced and adapted material, introduced individualized projects, ethics module,
- Implemented an ethics module embedded across multiple semesters of IS 296: https://mdata.umbc.edu/ethical-thinking-in-data-science/
- Smaller classes allowed for refining the course, with more individualized work attuned to student interests, peer mentoring through fellows and scholars supported a small, well connected data science ecosystem.
- Additional variations (Dr. Karen Chen):
  - Connections to community college via digital data storytelling project (presented at the National Data Science Workshop in summer 2022), pilot alternative assessment guided by “learning-by-teaching” principal in Fall 2022, designed new assignment and activity with UMBC student data
  - K-12 outreach: Pilot high school student data science introductory training with a subset of IS 296 materials, taught by IS 296 student (Spring 2022 cohort) in summer 2022
Peer mentoring and Professional Development
Transdisciplinary Ecosystem

Cross Disciplinary Collaborations
- Information Systems
- Physics
- Computer Science
- Geography & Environmental Systems
- Language Literature and Culture
- Goddard Earth Sciences Technology and Research (GESTAR) II @UMBC
- Anthropology

Data Science Scholars  [https://datasciencescholars.umbc.edu/](https://datasciencescholars.umbc.edu/)
Data Science Scholars

- DSS Program (launched Spring 2021, now funded through multiple NSF grants)
- Creating a Data Science Ecosystem across Majors
- Collaborations with CWIT and CS3 at UMBC
- Types of scholarly work
  - Research
  - Teaching fellows
  - Peer Mentoring
  - Professional development and mentoring
• Program page:  [https://datasciencescholars.umbc.edu/](https://datasciencescholars.umbc.edu/)

• Working as teaching fellows
  – Provide time slots to meet with IS 296 students to help with programming tasks
  – First meeting/introduction with IS 296
  – Attend class for labs

• Working on research projects
  – Develop a research idea and meet with your research mentor regularly
  – First meeting/introduction with IS 296
  – Share lessons learned with IS 296

• Serve as advisors to IS 296 students as they develop their own projects

• Participate in professional development and mentoring activities, such as through CWIT, CS3 and external events
Center Partnerships: Dr. Carolyn Seaman, Director CWIT; Dr. Christine Mallinson (Director CS3), Dr. Felipe Filomeno in (Associate Director CS3), Ms. Morgan Lovell, CWIT

Mentoring partners: Dr. Karen Chen (IS), Dr. Dillon Mahmoudi (GES), Dr. Brian Soller (Sociology, Anthropology, and Public
● Fellows participate in the **CWIT Affiliates program**
  ○ Allows students to participate in CWIT activities
  ○ Open to all students studying computing or engineering
  ○ Networking and Mentorship opportunities
  ○ Academic support
  ○ Community connection

● **CWIT Mission**: Enable success for all women and other underrepresented groups in technology fields

● **Center for Social Science Scholarship (CS3)**
  ○ Hub for social science research at UMBC
  ○ External grant administration support
  ○ Internal seed grants for faculty and students
  ○ Events, trainings, and webinars
  ○ Social Sciences Forum Distinguished Lectures
  ○ [Podcast](#) (Retrieving the Social Sciences)
Past Talks

- Exploring physical and Machine Learning approaches for stochastic modeling and ensemble prediction of weather and climate by: Dr. Aneesh Subramanian
- What is special about GEO-AI and Spatial Data Science by Dr. Shashi Shakhar

Upcoming Talks

- March 14: POLARIS: The Pursuing Opportunities for Long-term Arctic Resilience for Infrastructure and Society By: Dr. Guangqing Chi
- April 18: Talk by Dr. Lauren Andrews from NASA Goddard Space Flight Center

Watch past events: https://iharp.umbc.edu/talk-tuesday-videos/
Access to Curriculum
Certificate in Business Analytics

Curriculum (15 credits total)

Core required courses
- IS 296 Introduction to Data Science
- IS/MGMT 496 Business Analytics

Electives (choose three – 9 credits)
- IS 425 Decision Support Systems
- IS 438 Project Management
- IS 474 Legal Aspects of Information Systems OR MGMT 360 Business Law
- MGMT 385 Business Ethics
- IS 480 Data Visualization course
- IS 497 Internship in Business Analytics
- MGMT 490 Independent Study in Business Analytics

Augment our existing undergraduate curriculum in with an additional focus on business analytics skills and topics through an upper division certificate program.

Certificate open to students in other departments or students outside UMBC.
IS 296 accepted as GEC for Science/non-lab GEP starting in the fall of 2023

IS 296 part of Computing Minor at UMBC, open to all departments (for non computing students)
Emerging Themes

- Transdisciplinarity in Education
- Adopting Notebooks for new Modules (across disciplines and career levels)
- Support Diversity at all levels
- Intentional Engagement
- Meeting Students where they are
- Developing Community supports
- Continued Mentoring/Career Support and Advancement
1. What are the essential elements of a data science ecosystem, which supports diversity, equity, inclusion and accessibility?
2. What are some inventive ways to support data science ecosystems in environments with limited resources?
3. Can we or How do we support a data science ecosystem at scale from a DEI lens?
Examining Linguistic Diversity and Data Science Curriculum in STEM Education with a Justice, Equity, Diversity, and Inclusion Lens

Manisha Sharan, Ph.D.
Vandana Janeja, Ph.D.

This material is based upon work supported by the National Science Foundation (NSF) under Grant No. DUE-1937267. Any opinions, findings, interpretations, conclusions or recommendations expressed in this material are those of its authors and do not represent the views of the AAAS Board of Directors, the Council of AAAS, AAAS’ membership or the National Science Foundation.
Thank you for attending!

Slides and recording will be available in the coming days.

We value your feedback, please take a few minutes to complete the survey.

@IuseProgram

https://www.aaas-iuse.org

AAAS IUSE Initiative