

ICQE 2026 HIROSHIMA

Eighth International Conference on Quantitative Ethnography:

CONFERENCE PROCEEDINGS SUPPLEMENT

Edited by:
Ayano Ohsaki, Toshio Mochizuki, and Brendan Eagan

PROVISIONAL

9-13 November 2026

Hiroshima, Japan

INTERNATIONAL SOCIETY FOR QUANTITATIVE ETHNOGRAPHY

Preface

This volume comprises the supplementary proceedings of the Eighth International Conference on Quantitative Ethnography (ICQE26). To encourage an inclusive community, the conference has consistently provided both formal Springer-published proceedings and a conference proceedings supplement, allowing for a variety of submission types and offering authors several ways of sharing their work. The conference, along with the wider community supported by the International Society for Quantitative Ethnography (ISQE), serves as a forum for discussion on Quantitative Ethnography (QE), its data, methodologies, and applications, and welcomes both returning and new scholars from diverse fields.

This proceedings supplement features 11 Doctoral Consortium submissions, 24 Posters, 1 Japanese Poster, 6 Workshops, 1 Symposium, and 32 Research Agenda Development proposals. These contributions are part of the conference program presented at ICQE26 alongside the 46 accepted full papers.

The range of authors, contexts, data, and topics highlights the interdisciplinary nature of the QE community. As outlined in the Springer Proceedings, the submissions in this volume include three main themes: Theory and Methodology, Teaching and Learning, and Culture, Policy, and Society. Many submissions span more than one theme.

ICQE26 was made possible by the many people who contributed their time and effort. Our sincere thanks go to the Doctoral Consortium Chairs and the Program Committee, together with the Workshop Committee, the Conference Committee, the Additional Reviewers, and the wider QE community. Previous Program Committee Chairs generously shared materials and guidance, and we extend our heartfelt gratitude for their guidance and support. Vibha Srinivasan deserves special recognition for her outstanding contributions to editing and preparing these proceedings. We are grateful to the International Society for Quantitative Ethnography (ISQE), the National Science Foundation (NSF, DRL-2526678), and the Hiroshima Convention & Visitors Bureau for their support. Finally, we thank the local chairs, organizers and student volunteers, who planned the conference and welcomed ICQE to Hiroshima.

September 2026

[Ayano Ohsaki]
[Toshio Mochizuki]
[Brendan Eagan]

Organization

Program Committee Chairs

Ayano Ohsaki	Tohoku University, Japan
Toshio Mochizuki	Waseda University, Japan
Brendan Eagan	University of Wisconsin – Madison, USA

Program Committee

Amalia Daché	University of Pennsylvania, USA
Crina Damşa	University of Oslo, Norway
Daisuke Kaneko	Hokusei Gakuen University, Japan
Hanall Sung	University of Tennessee, Knoxville, USA
Hiroshi Kato	The Open University of Japan, Japan
Hiroyuki Masukawa	Aoyama Gakuin University, Japan
Mamta Shah	Elsevier, USA
Morten Misfeldt	University of Copenhagen, Denmark
Ryan Baker	Adelaide University, Australia
Seung Lee	Pepperdine University, USA
Sunny Lin	National Yang Ming Chiao Tung University, Taiwan
Ward Peeters	Monash University, Australia
Yotam Hod	University of Haifa, Israel
Zachari Swiecki	Monash University, Australia
Zhichun Liu	University of Hong Kong, Hong Kong SAR, China

Doctoral Consortium Chairs

Golnaz Arastoopour	Vanderbilt University, USA
Michael Phillips	Monash University, Australia

Workshop Committee

Yuanru Tan	University of Wisconsin-Madison, USA

Local Conference Committee Chairs

Jun Oshima	Shizuoka University, Japan
Ritsuko Oshima	Shizuoka University, Japan
Ryohei Ikejiri	Hiroshima University, Japan
Shunpei Yamaguchi	Hiroshima City University, Japan

Local Conference Committee

Miho Kawasaki	Shizuoka University, Japan
Bunichi Otaki	Shizuoka University, Japan
Jun Lu	Shizuoka University, Japan

Conference Committee

Golnaz Arastoopour	Vanderbilt University, USA
Brendan Eagan	University of Wisconsin-Madison, USA
César Hinojosa	University of Wisconsin-Madison, USA
Ward Peeters	Monash University, Australia
Vibha Srinivasan	University of Wisconsin-Madison, USA

Additional Reviewers

Adaurennaya Onyewuenyi
Adrienne Traxler
Akihito Takeuchi
Alif Syaiful Adam
Amit Paikrao
Andres Felipe Zambrano
Andrew Ruis
Azusa Nakata
Bunichi Otaki
Catherine Zhang
Chuanlei Guo
Curtis Isozaki
Danielle Aguilar
Danielle Espino
Dongpin Hu
Edith Bouton
Emma Morton
Eric Hamilton
Ethan Connolly
Geraldine Cochran
Go Shoji
Golnaz Arastoopour Irgens
Haiying Li
Hao Shen
Jade Pratt
Jai Oni Dane
James Mitchell
Jimpei Hitsuwari
Jingyi Chen
Jun Lu
Kaixiang Kang
Liane Weintraub
Lindi R. Shepard
Lingfei Cao
Lu Liang
Mao Morita
Marguerite Williamson
Mariah Autumn Knowles
Mercedes Diaz
Ming Ma
Mingxue Xu
Nastassia Jones
Neta Sheffer
Oindriza Reza Nodi
Pak Ho Chan
Raihan Munira Moh Sani
Ran Nakao
Rebecca Quintana
Salim George
Samuel Green
Seiyon Lee
Shan Zhang
Shiyao Wei
Shuntaro Iseri
Sinead Meehan
Siti Nabihah Md Sahak

Stine Jensen
Takamichi Asakura
Tianshi Hao
Ting-Yen Kuo
Tomohito Yamazaki
Tomoki Hirano
Vedant Khatri
Xin Li
Xiner Liu
Xuechun Shi
Yayoi Matsuhisa
Yeyu Wang
Yiqiu Zhou
Yixin Cheng
Yujeong Cho
Yuna Dai
Yung-Hsiang Huang
Yunseo Lee
Yusuke Hayashi
Zhanlan Wei
Zhenqi Xu
Zhiqiang Cai

Table of Contents

Keynote

What does Quantitative Ethnography do? — From a Methodological Perspective: A Dialogue with a Qualitative Research Methodologist and the QE Founder 1
Takashi Otani and David Williamson Shaffer

Panels

Ethnography of disasters: How we can approach the human practice through qualitative and quantitative methodologies 3
Mordecai G. Sheftall, Aleksandra Ross, Miki Mochizuki and Jun Oshima

Quantitative Ethnography in Narrative Analysis 4
Michaela Leslie-Rule and Ward Peeters

Doctoral Consortium

Analysis of Manzai Characteristics Based on Switching Pauses 5
Akihito Takeuchi

Building Cyber Confidence: Student Perspectives on How Out-of-School Cybersecurity Pre-Apprenticeship Influences Cybersecurity Skills, Conceptual Understanding, and Tech Career Interest 7
Luis Ortiz

Extending Shared Epistemic Agency in Game-Based Learning: An Ordered Network Analysis of Occultation 11
Lingfei Cao

Modeling the Dynamics of Learners' Complex Problem-Solving in Multi-Agent Supported Learning Environments 14
Mengyu Xia

Modeling Upward Influence in Hierarchical Organizations: A Quantitative Ethnography of Dual-Role Actors as Leaders and Followers 16
Paul McGachey

Olympic Identity and Value Networks 19
Thomas Briskey

Preparing Future Teachers to Enact Environmental Justice Education 23
Lindi Shepard

Supporting Novice Teachers in Responsive Mathematics Teaching: A Stages of Concern Analysis Across Three Phases of Preparation 27
Riku Sayuj

The Role of Culture and Identity Conflict on Leadership Identity Development in College Students 30
Catherine Zhang

Towards an Affordance-Based Taxonomy of GenAI-Assisted Writing Tools 33
Yixin Cheng

Research Agenda Development

Culture, Policy, and Society

A Quantitative Ethnographic Exploration of Epidermolysis Bullosa Patient Journeys in China 37
Yusong Zhang and Jiaqi Liu

Constructing Sustainability: Narratives of the Olympic Legacy 39
Thomas Briskey

Cultivating Self-Efficacy in Collaborative Repair: A Participatory Quantitative Ethnography Approach to Designing Spaces for Tinkering 42
Kohei Aida

Detecting Traces of Scaffolding in a Multi-Agent Simulation Environment 44
Williams Chidera Okpokiri, Cheng Jiu Yin and Juan Zhou

Evaluation of Machine Translation with Epistemic Network Analysis 47
Jianxing Yu, Tin Yan Koo and Santiago Alonso Sousa

Exploring Whether Labeling an LLM Interviewer as AI or Human Changes the Structure of Self-Disclosure 50
Satoshi Kadomatsu and Toru Nagahama

Facilitators and Barriers to STEM Persistence for African Diaspora Women in General Chemistry 52
Jade Pratt

From Pain Meaning to Clinical Action: Using Quantitative Ethnography to Identify Cultural-Linguistic Pain Phenotypes 55
Shinye Kim and Ji In Woo

Honoring Richness in Blackness at Work: A Quantitative-Ethnographic Exploration of Psychological Safety, Joy, and Thriving Among Black Women in Tech 57
Shelbi Chandler

Mapping “Responsibility” and “Deservingness” in State Truancy Laws: A Quantitative Ethnography Research Agenda 60
Ericka Weathers

Modeling Educators' Meaning-Making About Responsible AI Implementation: A Quantitative Ethnographic Comparison Across Educational Contexts and Professional Roles 64
Juhee Kim and Eunsung Park

Re-Examining Globalization: Definition and Biases 68
Letitia Lee and Andrea Matos Devesa

The Impact of Rapid Prototyping in the AI Era on Organizational Work Validation, and the Application of Quantitative Ethnography in Service Design Practice 70
Yuya Kashiwa

Theory and Methodology

A Conceptual Design of Researcher-Note-Mediated Coding Support in Quantitative Ethnography 72
Hideki Kondo, Sayaka Tohyama and Masayuki Yamada

An AI-Human Collaborative Content Analysis of University Functions in Chinese National Strategies 74
Xin Li and Kaixiang Kang

Teaching and Learning

Connecting Physics and Mathematics Reasoning Across Three Inventories: A Quantitative Ethnographic Approach 77
Kazi Aatish Imroz and Geraldine Cochran

Critical In-Depth Reading for Immigrant and Refugee Background Multilingual Learners 80
Hazel Vega, Jennifer C. Mann and Alison Turner

Defining Pharmaceutical Data Science: A Competence Model as an Emerging Artifact 83
Stine Marie Høgh Jensen, Adrienne Traxler, Hendra Agustian, Frederik Christiansen, Lykke Pedersen, Arnault-Quentin Vermillet and Morten Misfeldt

EFL Students' Agentic Engagement in AI-assisted Writing Revision: A Transmodal Analysis 86
Yangyang Yu

Examining Students' Experiences in Graduate Physics Programs 89
Reidyn Wingate and Geraldine Cochran

Experiencing Differentiation: A Quantitative Ethnography of Introductory Physics Pathways 91
Stella F. Nelson and Geraldine L. Cochran

Exploring IMPACT: Integrated Modeling Practices Advanced through Computational Thinking 94
Amanda Peel

Exploring Quantitative Ethnography Applications to Bridge Program Evaluation and Research in a Culturally-Sustaining Cascading Mentorship Pilot 96
Daranee Taychachaiwongse Teng

Functions of Computational Writing in Pharmaceutical Data Science Teaching 99
Morten Misfeldt, Adrienne Traxler, Frederik Christiansen, Stine Jensen, Hendra Agustian, Lykke Pedersen and Arnault-Quentin Vermillet

Mapping Out Who Teaches Computer Science in New Mexico 102
Adaline De Chenne, Christabel Wayllace, Amanda Peel and Paige Prescott

Mapping the Mindset: Reframing Epistemic Network Analysis as a Hermeneutic Tool in Philosophical Doctoral Mentoring 104
Mercedes Diaz

Neurodiversity and a new science of learning 107
Hendra Agustian

Reimagining Achievement Goal Theory for Minoritized Engineering Students 110
Dina Verdin

Revealing Hidden Patterns in Peer Feedback: An Epistemic Network Analysis of Student Beliefs and Practices 113
Wanju Huang, Ryan Wynkoop and Janet Alsup

So much data, so many stories: Where should I start with Epistemic Network Analysis? 116
Wanju Huang

Students' Opportunities for Meaningful Inclusion as Measured by Classroom Observations and Teacher Interviews 119
Alexandra Shelton, Rebecca A. Cruz, Lindi R. Shepard, Katy Mullins, Irene Ramirez, Rachel S. McClam and Isun Malekghassemi

Understanding How the Physics Community Integrates Mathematical Methods in Undergraduate Curricula 122
Aiswarya Amalucottage, Lyra D Moscuzza and Geraldine Cochran

Workshops

A Workshop on Writing in Quantitative Ethnography 125
Zachari Swiecki, Andrew Ruis and David Williamson Shaffer

DIY QE: Demonstrating, Building, and Sustaining Community-Developed Tools for Quantitative Ethnography 129
Golnaz Arastoopour Irgens, Amanda Barany, Luc Paquette, Andrew Ruis, Yuanru Tan, Zhiqiang Cai and Cody Marquart

Learning SCAT as a Qualitative Data Analysis Method: A Hands-on Workshop from Coding to Theory Writing 133
Takashi Otani and Daisuke Kaneko

Modeling Complex Learning Interactions with HINA: A Hands-on Workshop on Heterogeneous Interaction Network Analysis 137
Shihui Feng, Baiyue He, Alec Kirkley and Dragan Gasevic

Open-Source Tools for Transcribing and Coding Multimodal Qualitative Data: An Interactive Tutorial 141
Kirk Vanacore, Bakhtawar Ahtisham, Zhuqian Zhou, Daryl Hedley, Doug Pietrzak, Joshua Marland, Rachel Slama and Rene Kizilcec

Workshop: Introduction to Teaching QE 2026 146
Ayano Ohsaki, Brendan Eagan, Zhichun Liu, Amalia Daché, Yujeong Cho and Amanda Barany

Symposium

Reconstructing Meaning Through Quantitative Ethnography: Critical Perspectives on Race, Immigration, and Identity 150
Adaurennaya Onyewuenyi, Danielle Espino, Golnaz Arastoopour Irgens, Nichole Garcia and Brendan Eagan

Posters

Culture, Policy, and Society

Analysis of Manzai Characteristics via Clustering Based on Switching Pauses . . 157
Akihito Takeuchi, Mineaki Ohishi and Takashi Mitsuishi

Exploring the Effects of Program Years has on Dog Training Programs for Formerly Incarcerated Folks 161
Jai Oni Dane

Mocking Gender: An Epistemic Network Analysis of Yinyang Guaiqi in Online Chinese Gender-Based Offensive Language 165
Zhenqi Xu and Ward Peeters

Tracing Interpretive Discourse in Art Appreciation through Discussion: A Quantitative Ethnographic Analysis of Two Sessions on a Japanese Artwork . . . 169
Tomoki Hirano

Theory and Methodology

Beyond Discourse Data: Positioning Observation Notes in Computational Quantitative Ethnography 173
Ayano Ohsaki and Daniel Spikol

Calibrated LLM Coding in Quantitative Ethnography: A Comparison Study . . . 177
James Mitchell, Curtis Isozaki, Amalfi Gayosso, Alison Lawton and Liane Weintraub

The Codebook Teaches Scope: Coding Student Experience with Retrieval-Augmented LLMs 192
Kaixiang Kang, Ran Nakao and Yizhou Fan

When Construct Complexity Extends Beyond Connections: Quantifying Recurrence in Affect Dynamics 196
Seiyon Lee, Anthony F. Botelho and Daryn A. Dever

Teaching and Learning

Comparing STEM Identity Connections Among Students in the USA and Kenya 200
Nastassia Jones, Anna Wilson, Danielle Espino, Seung Lee and Eric Hamilton

Disciplinary Differences in Conceptions of Laboratory Instruction among Preservice Science Teachers: An Epistemic Network Analysis with Multigroup Statistical Comparisons 205
Yung-Hsiang Huang and Chia-Ching Lin

Epistemic Network Analysis of Self-Concept Articulation in a Dialogue-Based Identity Exploration Program for Adolescents 209
Go Shoji, Naoya Yoshida, Kaoru Hatano and Keita Suzuki

Evaluating the Impact of a Global Collaborative STEM Project: A Pre–Post Analysis of Teen Participants 213
Lu Liang, Danielle Espino, Seung Lee, Julia Gibson, Anmol Kaur, Nastassia Jones and Eric Hamilton

Exploring Discourse Changes in a Facilitator-Supported Novice Learner Community 217
Mao Morita, Hideki Kondo and Hiroyuki Masukawa

How Does Ambiguity Tolerance Relate to Creativity-Fostering Teaching Practices among Art Teachers? An Exploration through Epistemic Network Analysis 221
Jimpei Hitsuwari, Yoshitaka Ito and Bunichi Otaki

Inclusion as Contested Discourse: Perspectives from State, District, and School-based Practitioners 226
Lindi R. Shepard, Rebecca A. Cruz, Alexandra Shelton, Rachel S. McClam, Sarah A. Caroleo, Irene Ramirez and Katy Mullins

Productive Failure in Generative-AI-Mediated App Development: A Quantitative Ethnography of Student Reflections and Test Score Change 232
Shuntaro Iseri, Hideki Kondo and Masaaki Ishii

Temporal Dynamics of Interactional Meaning-making of Topics in Interprofessional Clinical Education 236
Jun Lu, Crina Damsa, Rachelle Esterhazy, Jun Oshima and Ritsuko Oshima

The Changing Structure of Student Reflection in AI-Supported Inquiry Learning 241
Ran Nakao, Go Sorano, Yizhou Fan and Kaixiang Kang

The Social Dimension in Global Virtual STEM Teams: A Comparative Network Analysis 245
Lu Liang, Danielle Espino, Seung Lee, Julia Gibson, Anmol Kaur, Nastassia Jones and Eric Hamilton

Under Pressure to Excel: Epistemic Network Analysis of Academic Identity and Wellbeing in a Malaysian High-Achieving School 249
Raihan Munira Moh Sani, Aini Marina Ma'rof and Mohd Mokhtar Bin Muhamad

Understanding Social Presence in Hybrid Learning: Teachers' Perspectives Across Educational Levels 254
Azusa Nakata, Mervi Heikkinen, Hanna Järvenoja and Karel Kreijns

Using Epistemic Network Analysis to Examine How Identity Feedback Can Foster Transformative Learning 258
Neta Sheffer and Yotam Hod

When Does GenAI Mediate Knowledge Building? Tracing Uptake Patterns in Collaborative Learning Discourse 262
Bunichi Otaki, Ritsuko Oshima and Jun Oshima

"Who Are We to the Cosmos?" Epistemic Network Analysis of Identity and Global Citizenship in Middle-School Students' Civilisation-Reboot Discourse 266
Pak Ho Chan and Chen Chen

Japanese Poster

資格取得学習・専門知学習・推しの学習それぞれの学習中に生起する内言記録のENA比較分析 (A Comparative Analysis Using ENA of Inner Speech Records Occurring During Learning for Certification, Disciplinary Knowledge, and Oshi-katsu) 270
Nonoka Fujita and Hiroyuki Masukawa

What does Quantitative Ethnography do? — From a Methodological Perspective: A Dialogue with a Qualitative Research Methodologist and the QE Founder

Takashi Otani[1] and David Williamson Shaffer[2]

[1] Nagoya University of Economics, Japan
[2] University of Wisconsin-Madison, USA

Abstract. Can ethnography be quantitative? For Takashi Otani, a qualitative research methodologist and developer of SCAT (Steps for Coding and Theorization), the term "Quantitative Ethnography" (QE) initially sounded like an oxymoron. This apparent contradiction provides the starting point for his dialogue with David Williamson Shaffer, the founder of QE.

Drawing on their respective methodological perspectives, the two speakers explore the following three fundamental questions: What does "quantitative" mean in QE? What role does coding play in analysis and theorization? And how are quantitative and qualitative approaches related? Their dialogue evolves from empirical methodological questions to deeper epistemological issues concerning interpretation, transparency, and the justification of research claims. In QE, what is specifically quantified, and for what purpose? Is Epistemic Network Analysis (ENA) necessary for QE, and what kinds of qualitative data can it examine? How do researchers' interpretations shape analytical decisions, and how can these decisions be made transparent and open to scrutiny? In particular, the speakers consider how quantitative models and original qualitative data can inform one another, and what it means to "close the interpretive loop" while remaining open to alternative interpretations.

Beyond these methodological questions, the dialogue also addresses a tension shared by SCAT and QE in their efforts to support research inquiry. While SCAT focuses on qualitative inquiry and QE integrates qualitative and quantitative approaches, both aim to make analytical procedures more explicit and accessible. This support enables researchers to pursue inquiries that might otherwise have been difficult. At the same time, it risks encouraging superficial applications without a thorough understanding of the methodological and interpretive foundations involved. Pulling from their experiences as scholars of research methodology, the speakers reflect on how methodological tools shape research practice, how researchers engage with these tools, and what is required for responsible methodological practice.

By exploring where their approaches converge and where important differences remain, this keynote dialogue encourages researchers—from novices to experienced scholars—to reconsider the methodological assumptions,

interpretive responsibilities, and standards of evidence underlying their own research.

Ethnography of disasters: How we can approach the human practice through qualitative and quantitative methodologies

Mordecai G. Sheftall[1], Aleksandra Ross[2], Miki Mochizuki[1], and Jun Oshima[1]

[1] Shizuoka University, Japan
[2] Kveloce, Belgium

Abstract. Ethnography is an established sociological methodology for exploring unfamiliar cultural practices. While many high-quality studies are presented at the conference, few focus on sociological research themes or the identification of previously unknown cultural practices. This panel session will examine how quantitative approaches in ethnography can be grounded in sociological research through in-depth, constructive dialogue between qualitative and quantitative scholars.

We focus on a social practice: disaster preparation and resilience. The history of humankind is an iterative process of building resilience in the face of disasters. Human communities have experienced many types of disasters, including wars, earthquakes, unknown viruses, and nuclear power plant meltdowns, to name just a few. Such events will continue to occur; they will not simply fade away. How can we endure these unexpected disruptions in our lives while sustaining and developing our societies and communities? How can we learn to be resilient to disasters based on our past experiences? How might a quantitative approach to this social practice provide new insights that go beyond traditional qualitative approaches?

In this panel session, four panelists will present their qualitative or quantitative studies on disaster-related human practices. Based on their reports, the panelists will discuss why they chose their respective approaches (qualitative, quantitative, or mixed-methods) for their research programs, the challenges they perceive in adopting the alternative approach, and how they can continue to incorporate complementary methods in their fields.

Quantitative Ethnography in Narrative Analysis

Michaela Leslie-Rule[1] and Ward Peeters[2]

[1] Gates Foundation, USA
[2] Monash University, Australia

Abstract. Narratives are the cornerstone of the human experience, offering profound insights into culture, identity, and social systems. However, scaling qualitative narrative analysis to capture broader population patterns without flattening the rich, complex nuance of individual lived experiences remains a persistent methodological challenge. This special session explores the powerful intersection of Quantitative Ethnography (QE) and Narrative Analysis, demonstrating how mixed-methods frameworks and epistemic network analysis can uncover the deep structural connections, and elucidate the narrative landscape within human stories.

Led by researcher, narrative strategist, and grantmaker Michaela Leslie-Rule, this session draws upon her decades long career exploring young people's relationship to education and work, including the landmark Striving to Thriving study and her book, How We See Us: Young People Imagining a Path to Their Futures. Participants will discover how QE can be applied to large-scale qualitative datasets—such as focus groups, deep-dive interviews, and digital media—to rigorously map narrative landscapes, how individuals connect concepts and beliefs, navigate institutional barriers, and form identities. By bridging the micro-level analysis of personal discourse with the macro-level analysis of systemic cultural narratives, this session offers a transformative toolkit for researchers striving to back community stories with empirical, actionable weight.

Analysis of Manzai Characteristics Based on Switching Pauses

Akihito Takeuchi[1]

Graduate School of Information Sciences, Tohoku University,
Sendai, Miyagi 980-8577, Japan
akihito.takeuchi.r5@dc.tohoku.ac.jp

1 Abstract

My doctoral dissertation aims to capture characteristics of Manzai from utterances. These characteristics include changes in the pace of interaction, performers' persona, and humor. It also aims to clarify the skills of expert performers and the challenges faced by novice performers. At present, I have analyzed switching pauses in expert Manzai performers. The results could be interpreted as reflecting the maintenance of fast-paced interaction, which is characteristic of Manzai. They could also be interpreted as reflecting performers' persona. In future work, I will examine humor and differences according to proficiency level.

2 Goals of the research

My doctoral dissertation focuses on Manzai. I examine switching pauses, which refer to the timing of utterances during speaker transitions. Through switching pauses, I aim to clarify characteristics of Manzai, such as changes in the pace of interaction, performers' persona, and humor. Furthermore, by comparing expert and novice performers, I aim to clarify the skills of expert performers and the challenges faced by novice performers.

3 Background of the project

Manzai is a form of performing art in which mainly two performers take on the roles of boke and tsukkomi. Based on a script, they engage in humorous interaction and aim to make the audience laugh. Manzai is performed on stage. Therefore, audience laughter is thought to be related not only to the content of utterances but also to nonverbal elements such as utterance timing, posture, and gaze. This study focuses on switching pauses. This is because switching pauses are known to influence the impression of speakers and the naturalness of conversation [1]. Manzai is a form of conversation that intentionally aims to elicit laughter from the audience. Therefore, in Manzai, changes in the pace of interaction, performers' persona, humor, and proficiency level are expected to be reflected in switching pauses.

4 Methodology

At present, this study uses switching pauses produced by expert Manzai performers as the target of analysis. In future work, the same analysis will be applied to novice Manzai performers. Manzai performers use various utterance timings depending on the scene. Therefore, it is necessary to compare the distributions of switching pauses for each performer. For this reason, hierarchical clustering is applied to the dissimilarities between distributions. The clustering is conducted separately for each role. This makes it possible to classify switching pauses within each role. For each obtained cluster, the characteristics of the cluster are interpreted based on its representative distribution.

5 Preliminary or expected findings

At present, I have conducted clustering based on the distributions of switching pauses for each role, using data from 25 expert Manzai performances. For tsukkomi, three types were observed. The first was a major type that responds immediately. The second was a sub-major type that responds after a short pause. The third was a singular type that includes long positive switching pauses. These types could be interpreted as reflecting the maintenance of fast-paced interaction, thinking, and confusion, which are characteristic of Manzai. On the other hand, for boke, three types were also observed. The first was a major type that starts speaking around the end of the other performer's utterance. The second was a sub-major type that starts speaking before the other performer finishes speaking. The third was a singular type that starts speaking after a pause. These types could be interpreted as forms of performers' persona, such as maintaining fast-paced interaction, self-centeredness, and performing at one's own pace. In future work, I will also examine humor. I will also apply the same analysis to novice Manzai performers. Through this comparison, I will examine the skills of expert performers and the challenges faced by novice performers.

6 Expected contributions

At present, I have shown that quantitative distributions of switching pauses can be interpreted in relation to changes in the pace of interaction and performers' persona. In future work, I will relate switching pauses to the humor of Manzai. This will make it possible to understand characteristics of Manzai through quantitative switching pauses. I will also examine qualitative skills and challenges based on differences in proficiency level.

References

1. Nagaoka, C., Draguna, M., Komori, M., Nakamura, T.: The influence of switching pauses on interpersonal perception in dialogues. In: Proceedings of Human Interface Symposium 2002. pp. 171–174 (2002), (in Japanese)

Building Cyber Confidence: Student Perspectives on How Out-of-School Cybersecurity Pre-Apprenticeship Influences Cybersecurity Skills, Conceptual Understanding, and Tech Career Interest

Luis Manuel Ortiz

Drexel University, School of Education, Philadelphia, PA, USA

lmo58@drexel.edu

Abstract. This exploratory qualitative study examines how participation in a state-registered out-of-school cybersecurity pre-apprenticeship program influences high school students' confidence in cybersecurity skills, conceptual understanding, and interest in technology careers. Conducted at One Art Community Center in Philadelphia, Pennsylvania, United States with 11 urban underrepresented high school students, the study captures authentic student perspectives on problem-solving strategies and collaborative learning experiences. Using qualitative group interviews analyzed through Ordered Network Analysis (ONA), the research models relationships among confidence, understanding, career interest, collaboration, and problem-solving to reveal meaningful structural connections within student discourse. Findings aim to inform equitable approaches to cybersecurity education and highlight how targeted out-of-school interventions can bridge opportunity gaps for underserved youth in STEM.

Keywords: Ordered Network Analysis, Quantitative Ethnography, Cybersecurity Education, Out-of-School Learning, Career Interest, Underrepresented Youth

1 Goals of the Research

This study explores how participation in an out-of-school cybersecurity pre-apprenticeship program influences high school students' confidence in cybersecurity skills, understanding of cybersecurity concepts, interest in tech careers, problem-solving strategies, and collaboration experiences during the initial phase of a 24-week program. Grounded in equity-centered perspectives on STEM education, the research expects that students will demonstrate emerging confidence, developing conceptual understanding, and evolving career interest as they progress through the program, while articulating the problem-solving and collaborative learning experiences they are currently developing. The study is guided by two research questions: (1) How does participation influence students' confidence, understanding, career interest, problem-solving, and collaboration? (2) What aspects of the program do students identify as most supportive of their educational and career aspirations in technology?

2 Background of the Project

The cybersecurity workforce faces a critical shortage, with an estimated 1.8 million unfilled positions globally, exacerbated by the underrepresentation of women and people of color in technology fields [1]. Research demonstrates that exposure to cybersecurity careers must begin during middle and high school, as early engagement is strongly associated with sustained STEM participation and career success [2]. Yet traditional Career and Technical Education programs are often inaccessible or misaligned with industry needs, limiting career exposure for underrepresented students and focusing disproportionately on technical skills while overlooking competencies such as systems thinking, communication, and civic engagement [3]. Game-based learning and out-of-school programs have emerged as promising pathways, with programs like GenCyber camps demonstrating success in enhancing cybersecurity interest among underrepresented youth [4]. Despite these advances, significant gaps remain in understanding how such programs influence student confidence, conceptual understanding, career interest, problem-solving, and collaboration from the students' own perspectives. This study addresses that gap by examining student perspectives during a state-registered cybersecurity pre-apprenticeship designed specifically for underserved Black and Brown youth in Philadelphia.

3 Methodology

This one-month exploratory qualitative study was conducted during the active pre-apprenticeship program at One Art Community Center. Participants engaged in their regular program activities with no additional interventions. Data collection involved two 15-minute online surveys administered via Qualtrics at weeks 1-2 and week 4, and six 1-hour in-person group interviews conducted during weeks 2-3 and 3-4. Each survey contains 25 five-point Likert-style items measuring confidence, self-concept, and career goals. Group interviews follow semi-structured protocols across five key domains: confidence in cybersecurity skills (CONF), understanding of cybersecurity concepts (UNDER), interest in technology careers (CAREER), problem-solving strategies (PROB), and collaboration and peer learning (COLL).

Group interview transcripts will first undergo inductive thematic analysis to identify recurring meanings within student narratives, followed by Ordered Network Analysis (ONA) to model directed relationships among the five analytic codes: confidence in cybersecurity skills (CONF), understanding of cybersecurity concepts (UNDER), interest in technology careers (CAREER), problem-solving strategies (PROB), and collaboration/peer learning (COLL). This combined approach enables both thematic interpretation and structural modeling of how students connect and sequence these constructs in their discourse.

ONA moves beyond isolated themes and standard co-occurrence by visualizing and comparing the directionality and structure of connections between codes. This reveals not only what outcomes students describe, but the sequential pathways of how learning processes, peer interactions, and problem-solving experiences lead to

growing cybersecurity confidence, understanding, and career interest across program timepoints.

4 Expected Findings

ONA models are expected to reveal meaningful sequential patterns and directed connections between students' expressions of confidence, conceptual understanding, and career interest. Furthermore, by interpreting the position of these codes within the network space, ONA is expected to reveal the specific roles that features like peer collaboration and hands-on challenges play such as whether they act as primary drivers or supportive factors in the students' overall aspirations. Comparisons between early and later group interview data are expected to reveal shifts in how students sequentially articulate and integrate these constructs as the program progresses. Qualitative themes are anticipated to highlight specific program features that students identify as most supportive of their aspirations.

5 Expected Contributions

Findings will offer a grounded, student-centered account of how out-of-school cybersecurity programming shapes learning identity and career readiness for underrepresented youth. This study produces empirical findings and design principles for practitioners serving diverse youth through nontraditional cybersecurity career pathways. By applying ONA to student discourse, it advances methodological approaches for examining how confidence, understanding, career interest, collaboration, and problem-solving interconnect in informal learning environments. Furthermore, findings will inform program design recommendations for We Love Philly and similar initiatives, offering evidence-based strategies for how targeted pre-apprenticeship programs can bridge opportunity gaps and prepare underserved youth for careers in cybersecurity and technology.

References

1. Bowen, D., Juarez, J., Jones, N., Reid, W., & Simpson, C. (2022). Cybersecurity educational resources for K–12. Journal of Cybersecurity Education, Research and Practice, 2022(1), Article 6.
2. Clark, D. B., Tanner-Smith, E. E., & Killingsworth, S. (2016). Digital games, design, and learning: A systematic review and meta-analysis. Review of Educational Research, 86(1), 79–122.
3. Coenraad, M., Pellicone, A., Ketelhut, D. J., Cukier, M., Plane, J., & Weintrop, D. (2020). Experiencing cybersecurity one game at a time: A systematic review of cybersecurity digital games. Simulation & Gaming, 51(5), 586–611.
4. Connolly, T. M., Boyle, E. A., MacArthur, E., Hainey, T., & Boyle, J. M. (2012). A systematic literature review of empirical evidence on computer games and serious games. Computers & Education, 59(2), 661–686.

5. Dawson, J., & Thomson, R. (2018). The future cybersecurity workforce: Going beyond technical skills for successful cyber performance. Frontiers in Psychology, 9, Article 744.
6. Jin, G., Tu, M., Kim, T.-H., Heffron, J., & White, J. (2018). Evaluation of game-based learning in cybersecurity education for high school students. Journal of Education and Learning (EduLearn), 12(1), 150–158.

Extending Shared Epistemic Agency in Game-Based Learning: An Ordered Network Analysis of Occultation

Lingfei Cao[1 [0009−0001−5438−5239]]

The University of Hong Kong, Hong Kong SAR, China lingfei.cao@connect.hku.hk

Abstract. This study examines students' shared epistemic agency in Occultation, a role-based multiplayer game. Building on Damşa et al.'s[3] shared epistemic agency framework, the study analyzes gameplay records from Sessions 2 to 5 through Ordered Network Analysis. The analysis hopes to trace the sequence and structure of students' epistemic and process-related actions, including which actions tend to appear first and which actions follow during gameplay. It also explores whether role changes, failed strategies, prompt-testing outcomes, or peer challenges function as possible trigger factors that shape how shared epistemic agency emerges in a highly interactive game-based learning context.

Keywords: Shared Epistemic Agency; Game-Based Learning; Collaborative learning.

1 Background and Goals of the Research

Collaborative learning can support knowledge creation because, when students interact around shared tasks or shared knowledge objects, they are possible to create and refine understandings that were not fully formed by any individual learner before the collaboration[2, 5].However, this possibility depends on shared epistemic agency, because students need to take collective responsibility for developing, evaluating, regulating, and improving ideas rather than simply being placed in groups or exchanging information[1, 6–8].

Epistemic agency provides a useful lens for examining how students take responsibility for this process of knowledge creation. Epistemic agency refers to students' intentional engagement in advancing, revising, and improving their knowledge and understanding [8, 3]. In collaborative learning, this agency becomes shared when students jointly create, regulate, and improve knowledge through interaction with others [3, 10] Damşa et al.[3] conceptualize shared epistemic agency as the capacity that enables groups to deliberately engage in knowledge-driven collaborative activities and create shared knowledge objects. Their framework identifies knowledge-related actions, such as creating awareness, creating shared understanding, and generative collaborative actions, as well as process-related actions, such as projective, regulative, and Relational actions.

Game-based learning provides a meaningful context for examining shared epistemic agency because games can organize students' actions through goals,

roles, rules, feedback, abilities, and constraints [4]. In collaborative game-based learning environments, students may need to question peers' decisions, respond to others' ideas, negotiate strategies, and adjust their actions. These features make game-based learning a useful context for observing how students shared epistemic agency is enacted in fast, interactive, and highly collaborative activities. Building on Damşa and colleagues' [3]framework, this study aims to explore a possible framework that can better capture shared epistemic agency in the collaborative game-based learning context. In addition, how the game-settings like role-play infludence students' enactment of shared epistemic agency

2 Methods

This study will draw on data from Occultation, a role-based multiplayer game designed to create uncertainty between human and AI participation. During each round, students interact in anonymous chat rooms and make judgments about which participant is the hidden AI, a design inspired by the logic of the Turing test [9]. The game includes two main roles. Interrogators attempt to identify the AI through questioning and observation, while traitors support the AI by designing prompts that help it remain unnoticed during the conversation.

The analysis will focus on gameplay records from Sessions 2 to 5. Because the data were collected in China and all participants are native Mandarin speakers, the videos will be transcribed by Mandarin-speaking researchers. Ordered Network Analysis will be used to model students' shared epistemic agency as ordered action patterns during gameplay. Students' roles as interrogators or traitors will be considered as a possible factor which may influence the pattern of students' shared epistemic agency.

3 Expecting Findings

Through Ordered Network Analysis, this study seeks to extend the current shared epistemic agency framework by capturing how students enact shared epistemic agency in a highly collaborative game-based learning context. The analysis may show which epistemic and process-related actions tend to occur before or after others during gameplay, and how these ordered connections form broader patterns of shared agency. By tracing these patterns across rounds, the study may also identify possible trigger factors, such as role changes, failed strategies, prompt-testing outcomes, or peer challenges, that shape when and how shared epistemic agency emerges.

4 Expected Contributions

This study aims to extend Damşa et al.'s[3] shared epistemic agency framework to capture students' agency in highly interactive game-based learning contexts.

This is because role-based interaction, immediate feedback, and repeated strategy testing may stimulate forms of shared epistemic agency that differ from those observed in Damşa et al.'s original collaborative learning context. In addition, this study may provide practical implications for designing game-based learning environments. By identifying how shared epistemic agency is enacted in gameplay, the findings highlight the importance of recognizing diverse forms of epistemic agency that students may enact during learning.

References

1. Barron, B.: Achieving coordination in collaborative problem-solving groups **9**(4), 403–436. https://doi.org/10.1207/S15327809JLS0904_2, _eprint: https://doi.org/10.1207/S15327809JLS0904_2
2. Bereiter, C.: Education and Mind in the Knowledge Age. Taylor & Francis Group, http://ebookcentral.proquest.com/lib/hkuhk/detail.action?docID=227481
3. Damşa, C.I., Kirschner, P.A., Andriessen, J.E.B., Erkens, G., Sins, P.H.M.: Shared epistemic agency: An empirical study of an emergent construct **19**(2), 143–186. https://doi.org/10.1080/10508401003708381
4. Nguyen, C.T.: Games and the art of agency **128**(4), 423–462. https://doi.org/10.1215/00318108-7697863
5. Paavola, S., Hakkarainen, K.: The knowledge creation metaphor – an emergent epistemological approach to learning **14**, 535–557. https://doi.org/10.1007/s11191-004-5157-0
6. Perkins, D.N.: King Arthur's round table : how collaborative conversations create smart organizations. New York : Wiley, http://archive.org/details/kingarthursround0000perk
7. Salas, E., Sims, D.E., Burke, C.S.: Is there a Big Five in teamwork? Small Group Research **36**(5), 555–599 (2005). https://doi.org/10.1177/1046496405277134
8. Scardamalia, M.: Collective cognitive responsibility for the advancement of knowledge. In: Smith, B. (ed.) Liberal Education in a Knowledge Society, pp. 67–98. Open Court, Chicago, IL (2002)
9. Turing, A.M.: Computing machinery and intelligence. In: Epstein, R., Roberts, G., Beber, G. (eds.) Parsing the Turing Test: Philosophical and Methodological Issues in the Quest for the Thinking Computer, pp. 23–65. Springer Netherlands. https://doi.org/10.1007/978-1-4020-6710-5_3
10. Zhang, J., Tian, Y., Yuan, G., Tao, D.: Epistemic agency for costructuring expansive knowledge-building practices **106**(4), 890–923. https://doi.org/10.1002/sce.21717, _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/sce.21717

Modeling the Dynamics of Learners' Complex Problem-Solving in Multi-Agent Supported Learning Environments

Mengyu Xia[0009-0001-8496-8611]

Department of Educational Technology, Graduate School of Education, Peking University, Beijing, China
marine_xia@stu.pku.edu.cn

Abstract. This study examines complex problem solving in a multi-agent AI-supported diagnostic environment. Multimodal process data are coded for metacognition, reasoning, and task execution, then modeled using Epistemic Network Analysis and trajectory analysis. Preliminary analyses identify material-driven reasoning, AI-anchored reasoning, and multi-resource coordination, alongside individual differences across tasks. The study informs understanding of learner regulation and the design of AI-supported learning environments.

Keywords: Complex problem solving · ENA · Metacognition · Reasoning.

Goals of the research. Multi-agent systems may support learning and motivation while creating additional regulatory demands [2, 6]. How learners coordinate multiple AI agents during complex problem solving remains insufficiently understood. This study models hypothesis generation, evidence organization, and metacognitive regulation, examining differences among learners and changes across tasks.

Background of the project. Learners complete simulated medical diagnosis tasks with a virtual patient, physical examination and diagnostic test reporters, an AI assistant providing medical knowledge, and an AI tutor supporting reasoning and reflection without giving direct answers. Learners decide what information to collect and how to use these agents.

Data for Tasks 2, 4, and 8 include platform logs, dialogues, behavioral traces, video, eye tracking, and think-aloud recordings.

Methodology. Drawing on self-regulated learning and complex reasoning [3, 4], the coding framework covers metacognition, reasoning, and task execution. Eight coders independently coded sample data and checked agreement. ENA models connections among coded activities using learner-by-task units [5]. Trajectory analysis projects task networks into a common ENA space to examine change across tasks [1].

Preliminary findings. Analysis of four learners across three tasks suggests three patterns: material-driven reasoning, using learning materials to generate hypotheses and organize evidence; AI-anchored reasoning, using AI feedback to generate or verify hypotheses; and multi-resource coordination, such as using the AI assistant for hypothesis generation and the AI tutor for evidence organization.

Trajectories differ: Student 026 shows a large displacement, including a downward shift at Task 4, whereas Student 022 remains closely clustered across tasks.

Expected contributions. Combining ENA and trajectory analysis characterizes differences and changes in problem-solving structures. The study aims to clarify how learners regulate reasoning and coordinate resources, informing support tailored to their resource-use patterns.

Acknowledgments. This work was supported by the National Natural Science Foundation of China (No. 62407001).

Disclosure of Interests. The authors have no competing interests to declare that are relevant to the content of this article.

References

1. Brohinsky, J., Marquart, C., Wang, J., Ruis, A.R., Shaffer, D.W.: Trajectories in epistemic network analysis. In: Ruis, A.R., Lee, S.B. (eds.) Advances in Quantitative Ethnography, vol. 1312, pp. 106–121. Springer International Publishing, Cham (2021). https://doi.org/10.1007/978-3-030-67788-6_8
2. Dinçer, S., Doğanay, A.: The effects of multiple-pedagogical agents on learners' academic success, motivation, and cognitive load. Computers & Education 111, 74–100 (2017). https://doi.org/10.1016/j.compedu.2017.04.005
3. Li, S., Huang, X., Wang, T., Pan, Z., Lajoie, S.P.: Examining the interplay between self-regulated learning activities and types of knowledge within a computer-simulated environment. Journal of Learning Analytics 9(3), 152–168 (2022)
4. Magnani, L.: Basic science reasoning and clinical reasoning intertwined: Epistemological analysis and consequences for medical education. Advances in Health Sciences Education 2(2), 115–130 (1997). https://doi.org/10.1023/A:1009736414434
5. Shaffer, D.W., Collier, W., Ruis, A.R.: A tutorial on epistemic network analysis: Analyzing the structure of connections in cognitive, social, and interaction data. Journal of Learning Analytics 3(3), 9–45 (2016)
6. Zheng, L., Shi, Z., Gao, L.: A generative artificial intelligence-enhanced multiagent approach to empowering collaborative problem solving across different learning domains. Computers & Education 241, 105489 (2026). https://doi.org/10.1016/j.compedu.2025.105489

Modeling Upward Influence in Hierarchical Organizations: A Quantitative Ethnography of Dual-Role Actors as Leaders and Followers

Ernest Paul McGachey

Colorado State University, Fort Collins, CO 80523, USA

paul.mcgachey@colostate.edu

Abstract. This dissertation investigates how mid- to senior-level leaders who operate as both leaders and followers enact upward influence towards more senior-level leaders. Specifically, the dissertation will interview members who have served as U.S. military senior enlisted leaders and been senior enlisted advisors to commanders of military organizations. This study uses quantitative ethnography and epistemic network analysis to develop a model of integrated followership behaviors and influence tactics as a system of co-occurring behaviors. These findings aim to contribute to followership and leadership theories by providing conceptual models that can inform a framework for mid- to senior-level leaders who must effectively advise and influence more senior-level decision-makers.

Keywords: Upward Influence, Upward Leadership, Followership, Organizational Leadership, Military Organizations, Quantitative Ethnography, Epistemic Network Analysis, Senior Enlisted Leaders, Enlisted Development

1 Goal of Research

The goal of the dissertation is to examine how mid- to senior-level leaders, who operate as both organizational leaders and followers within hierarchical organizations, enact upward influence using follower behaviors targeted toward a more senior leader. The research questions driving the study are:

1. What follower behaviors do senior enlisted leaders (SELs) use to enact upward influence tactics when ethically influencing their commanders in hierarchical military organizations?
2. Do the followership behaviors employed by SELs significantly differ between military branches?
3. How do the epistemic network models of SELs' behaviors inform existing followership theories?

2 Background of the Project

Our bias toward leadership has led to underdeveloped followership identities. Additionally, research indicates that followership increases organizational

performance [1], job satisfaction [2], organizational citizenship behaviors, work engagement, and job performance [3]. Furthermore, poor followership has been shown to contribute to unethical cultures in organizations [4]. Followership is a phenomenon that happens at all levels of an organization. This study explores senior-level leaders in military organizations that serve in the role of SEL. These leaders are the senior enlisted advisors to a military commander. Their job is to advise on personnel, operations, and strategic decisions. While SELs are senior leaders within their organization, they have no command authority but have a lot of influence due to their role and proximity to the decision-maker, the commander. This study explores how they use their follower behaviors in the role as advisors to influence the decisions and behaviors of military commanders.

3 Methodology

This study employs a quantitative ethnography (QE) methodology. The data consist of semi-structured interviews with SELs from different branches of the U.S. military who have served as SEL advisors to commanders at the grade of O-6 or higher. The codebook will be developed through a two-phase process to define the nodes used in the epistemic network analysis (ENA).

Phase 1 of coding consists of thematic coding using both deductive and inductive codes. A literature review of followership and influence theories has resulted in the identification of eighteen deductive codes. Seven followership behavior codes were deductively derived from the courageous follower theory [5], Kelley's followership theory [6], and the patterns of subordinacy theory [7]. Eleven influence tactic codes were deductively developed from Yukl's eleven proactive influence tactics [8]. In addition, inductive codes will be allowed to emerge within the interview transcripts to capture behaviors not represented in existing frameworks. These deductive and inductive codes will then be reviewed and refined to determine the nodes used in the second phase of coding.

Phase 2 of coding involves binary coding to format the data for ENA, enabling the modeling of co-occurrence relationships among codes. ENA will be used to generate network models of the interaction of SEL followership behaviors and influence tactics. This will allow analysis of the study's research questions.

4 Expected Findings

The first expected product is ENA models depicting the co-occurrence of followership behaviors and influence tactics. These models are expected to reveal patterns in how combined behaviors and tactics are used during upward influence attempts. Differences in these patterns of behaviors may emerge between SELs in different military branches, reflecting differences in cultural norms. Finally, new follower behaviors and/or new influence tactics could be discovered. Additionally, the study may reveal new or underdeveloped followership behaviors.

5 Expected Contributions

This study contributes to theory by modeling the integration of followership behaviors and influence tactics as a system of co-occurring behaviors. These network-based models provide an empirical foundation for extending followership theories such as the courageous follower, which have been conceptually developed but lack extensive empirical research. Additionally, the insights may inform leadership and followership training in both corporate and military contexts. The conceptual models can inform a framework of reference for mid- to senior-level leaders who must effectively advise and influence more senior-level decision-makers.

References

1. Alameh, K. Effective followership and successful leadership in small and medium enterprises in the Middle East. Doctoral dissertation, Walden University (2025).
2. Ribbat, M., Nohe, C., & Huffmeier, J. Followership styles scrutinized: temporal consistency and relationships with job attitudes and self-efficacy. PeerJ, 11, 1–38 (2023).
3. Liu, S. W. The effects of interaction of leadership and followership based on followers' perceptions: Relationships among leader-member exchange, organizational citizenship behavior, work engagement, and job performance. Journal of Neurology, Neurological Science and Disorders, 9(1), 22–32 (2023).
4. Carsten, M. K., & Uhl-Bien, M. Ethical followership: An examination of followership beliefs and crimes of obedience. Journal of Leadership & Organizational Studies, 20(1), 49–61 (2013).
5. Chaleff, I. The courageous follower: Standing up to and for our leaders (3rd ed.). Berrett-Koehler (2009).
6. Kelley, R. E. The power of followership: How to create leaders people want to follow and followers who lead themselves. Doubleday (1992).
7. Zaleznik, A. The dynamics of subordinacy. Harvard Business Review, 43(3), 119-131 (1965).
8. Yukl, G. Leadership in organizations (7th ed.). Prentice Hall (2011).

Olympic Identity and Value Networks

Thomas Briskey [0009-0006-1707-506X]

Pepperdine University, Malibu, CA, USA
thomas.briskey@pepperdine.edu

Abstract. This study explores how United States Olympic athletes construct and narrate their identity within the context of elite sport. While existing research on athletic identity focuses on major life events like retirement and injury, less is known about how athletes make sense of their experiences through personal narratives. Grounded in Narrative Identity Theory, this study examines how athletes articulate their athlete role and their "who I am" through stories of training, competition, and life beyond sport. This research uses semi-structured interviews with U.S.-based athletes across multiple Olympic sports, using Quantitative Ethnography (QE), and Epistemic Network Analysis (ENA) to examine patterns within athlete narratives. The research also investigates personal values of the athletes that emerge within the narratives and how they are structured and connected across stories. By centering narrative identity and personal values, this research hopes to contribute to a deeper understanding of the lived experiences of elite athletes. Findings are intended to inform coaching and development approaches in sport, as well a broader application to organizational and governance practices in athletics.

Keywords: Olympics, Identity, Values, Athletes.

1 Goals of the Research

The goal of this dissertation research is to explore how Olympic athletes construct and narrate identity through storytelling about their Olympic experience using Narrative Identity as the principal framework, while also mapping personal values within these stories. Specifically, it will explore how narratives construct and reflect athletes' lived experiences, what types of narrative themes emerge, and how they influence the structure of their personal value systems. Building on previous identity research, this study will broaden the application of a narrative approach to meaning-making by using Quantitative Ethnography and Epistemic Network Analysis to map relational connections between narrative identities and the values that define the Olympic athletes.

This study focuses on these questions:

1. How do Olympic athletes construct/narrate their identity and meaning within the context of elite sport?
   a. What personal value emerge within athletes' narratives, and how do these values relate across their stories?

2 Background

The Olympic Games is a global sport phenomenon that brings the best athletes around the world into competition to represent their country and sport. Most research tends to focus factors of the event that largely ignore the roles of the Olympians themselves, such as hosting and sustainability challenges [2, 6, 7, 8]. Missing from Olympic research are the lived experiences and values of the participants, how athletes construct their identity as elite athletes, and what values drive their training, performance, and competitive ethos. Traditional research on narrative identity has focused on measuring the exclusivity or strength of an athlete role, which results in a "snapshot" of identity rather than a dynamic process of development [5]. Furthermore, this "snapshot" approach ignores the relational nature of identity and its effect on decision-making and other psycho-social processes an athlete experiences [4, 5]. It also suggests that identity is static, whereas Olympians typically manage multiple identities (mother, son, competitor) beyond their status as an athlete.

This study addresses these gaps by exploring identity narratives and mapping their relationship to personal values and self-perceptions. By analyzing athletes' stories, this research moves beyond static snapshots to deepen the understanding of identity as an internal and evolving process in the athlete's journey. This narrative platform offers an opportunity for Olympians to articulate their experiences, philosophies, and values in relation to their athlete role and broader roles in society.

3 Methodology

The study will use a qualitative and quantitative design grounded in Quantitative Ethnography (QE) and Epistemic Network Analysis (ENA) to examine how identity and values emerge in athletes' narratives [10]. Data will be collected through semi-structured interviews with fourteen (N=14) current and former U.S. Olympic athletes (2000 – present) and recruited using purposive and snowball sampling to ensure a high-yield of appropriate data [3].

Interviews with focus on athletes' lived experiences, identity construction, and meaning-making in Olympic Sport. Each interview will be preceded by a brief demographic survey to provide contextual background. The interviews will be recorded and transcribed, and analyzed as first-person narratives to capture the nuanced or "thick description" accounts of the Olympic athlete experience [1,10].

Data will be segmented and open-coded to identify emergent constructs/codes within the narrative, focusing on narrative identities, meaningful experiences, and personal values. The constructs will be developed inductively through an iterative coding process, giving space to elicit higher- and lower-order themes. ENA will model the relationships and co-occurrences amongst these constructs to further analyze and visual the connections across narratives.

4 Expected Findings

The analysis of interview narratives and the subsequent modeling is expected to reveal how Olympic athletes construct identity and meaning through storytelling and what narrative themes athletes use to make sense of their experiences. These themes may relate to performance, resilience, conflict, or purpose and are expected to reflect the individual's mean-making processes as well as confirm or challenge the master narratives that already exist in elite sports. The study will also uncover the underlying values within the narratives, and how they relate to each other through different contexts and experiences. Given the relational analysis, the study may uncover the interplay between identity and values, and how they are expressed in personal narratives.

5 Expected Contributions

This research aims to advance the study of athlete identity by moving beyond role-based conceptualizations of identity towards a more dynamic definition of evolving narratives. This study also contributes a narrative identity perspective that captures how athletes actively construct meaning and identities through storytelling. It is expected to contribute to the breadth of Olympic literature by expanding the understanding of the Olympic athlete by exploring their stories and values through empirically grounded narratives. Insights may inform coaching education, athlete development programs, and future United States Olympic and Paralympic Committee work around identity and values.

By integrating Narrative Identity Theory with QE/ENA, the study offers a methodological contribution demonstrating how rich narratives can be analyzed relationally across experiences, contexts, and values. It also broadens the application Narrative Identity Theory to include psycho-social components that are often ignored in identity and sport research [5,9].

References

1. Arastoopour Irgens, G., & Eagan, B. (2023). The foundations and fundamentals of Quantitative Ethnography. *Communications in Computer and Information Science*, 3–16. https://doi.org/10.1007/978-3-031-31726-2_1
2. Boykoff, J. (2024a). *What are the Olympics for?* Bristol University Press.
3. Campbell, S., Greenwood, M., Prior, S., Shearer, T., Walkem, K., Young, S., Bywaters, D., & Walker, K. (2020). Purposive sampling: Complex or simple? research case examples. *Journal of Research in Nursing*, *25*(8), 652–661. https://doi.org/10.1177/1744987120927206
4. Carless, D., & Douglas, K. (2013). "in the boat" but "Selling myself short": Stories, narratives, and identity development in Elite Sport. *The Sport Psychologist*, *27*(1), 27–39. https://doi.org/10.1123/tsp.27.1.27

5. Chun, Y., Wendling, E., & Sagas, M. (2023). Identity work in athletes: A systematic review of the literature. *Sports*, *11*(10), 203. https://doi.org/10.3390/sports11100203
6. Grix, J., Gallant, D., Brannagan, P. M., & Jones, C. (2020). Olympians' attitudes toward olympic values: A "sporting" life history approach. *Journal of Olympic Studies*, *1*(2), 72–92. https://doi.org/10.5406/jofolympstud.1.2.0072
7. IOC - International Olympic Committee | olympics.com. (n.d.). https://www.olympics.com/ioc
8. Kretchmar, S. R., Dyreson, M., Llewellyn, M. P., & Gleaves, J. (2017). Introduction. History and Philosophy of Sport and Physical Activity. https://doi.org/10.5040/9781718221017.0004
9. McAdams, D. P., & McLean, K. C. (2013). Narrative identity. *Current Directions in Psychological Science*, *22*(3), 233–238. https://doi.org/10.1177/0963721413475622
10. Shaffer, D. W. (2018a). Quantitative ethnography. Cathcart Press.

Preparing Future Teachers to Enact Environmental Justice Education

Lindi R. Shepard[0009-0003-5629-575X]

Johns Hopkins University, Baltimore MD 21218, USA

Abstract. In times defined by polycrisis, education plays an essential role in cultivating the knowledge, dispositions, skills, and collective capacities needed to build more just and sustainable futures. These goals are especially important in the United States, where consumption and policy contribute disproportionately to global emissions, while harms fall unequally on marginalized communities—particularly in the Global South. Despite broad public support and expanding state-level mandates for climate and environmental literacy in K-12 standards, educators in the U.S. rarely receive professional preparation for this type of instruction. This gap is particularly consequential because teaching about environmental justice presents challenges beyond those of traditional subject areas (e.g., addressing topics with existential consequence). Yet little is known about how teachers learn to navigate these challenges or what conditions best support their preparation for this work. This study examines a transdisciplinary teacher preparation course designed for secondary science, technology, engineering and math (STEM) educators that positioned environmental justice as an integrating context for teaching and learning. Using a quantitative ethnographic approach, this study examines how participation in the course shapes pre-service teachers' learning, sense-making, and intended enactment across the overlapping activity systems of course, classroom, home, and community, as conceptualized through fourth-generation cultural-historical activity theory (CHAT). I use reflexive thematic analysis to identify key challenges, tensions, and catalysts related to enacting environmental justice education (EJE), while Epistemic Network Analysis (ENA) models relationships among coded elements to visualize how teacher candidates connect EJE to their professional practice, their communities, and their own lives.

Keywords: Environmental Justice Education, Teacher Education, Cultural-Historical Activity Theory, Epistemic Network Analysis

1 Background

Education is increasingly recognized as essential to building socially just and climate-resilient futures. Yet, despite growing support for climate and EJE [1], most teacher preparation programs in the United States provide little guidance on how educators might meaningfully integrate these issues into instruction [2-3]. Existing research suggests that many teachers support the inclusion of climate-related topics in schools [4] but feel underprepared to teach them [5], particularly in ways that address social

inequities, interdisciplinary complexity, and the emotional dimensions of ecological crises.

Although prior scholarship has highlighted the potential of EJE to connect content to students' lived experiences [6], empower youth through action in their communities [7-8], and meet or surpass curricular benchmarks [9], opportunities for teachers to develop the knowledge, skills, and dispositions needed to enact EJE remain limited. In response, a team of researchers and teacher educators (including the author) developed a course centered on building competencies for implementing EJE and integrated it as a required component of a secondary STEM teacher preparation program. The course was designed around three pedagogical innovations: (1) interdisciplinary learning across STEM subject areas, (2) explicit attention to social-emotional learning and educator well-being, and (3) place-based and outdoor learning experiences connected to local ecosystems and communities.

2 Goals of the Research

This research investigates the impact of a course designed to prepare future STEM teachers to weave environmental justice into instruction within their respective disciplines. Course outcomes will be examined with three primary goals; First, it seeks to inform ongoing development of high-quality teacher preparation for EJE—including improvements for subsequent iterations of the target course—by identifying effective pedagogical strategies, potential barriers to enactment, and catalysts that support transformative learning among pre-service teachers. Second, this work seeks to understand how learning about environmental justice issues impacts teacher candidates across multifaceted roles as learners, teachers, and community members. Third, the study explores the potential of quantitative ethnographic approaches to study complex transformations in teacher attitudes, learning, and behaviors, particularly, how ENA can help visualize differences in the way teacher candidates reflect on EJE, based on their respective subject area specializations and prior teaching experience.

3 Methodology

This study employs a quantitative ethnography design [10] using data collected across three cohorts of per-service secondary STEM teachers enrolled in a required EJE course. I will triangulate qualitative and quantitative data across pre- and post-course surveys, interviews, and artifacts including lesson plans, teaching journals, and teaching philosophy statements. Course surveys are designed to measure change across six domains: knowledge of environmental topics, pedagogical content knowledge, self-efficacy in teaching EJE, nature-connectedness, emotions, and pro-environmental behavior. Likert-type items within surveys will be converted to composite scores by taking the mean across all items in a given scale, and pre- to post-course change will be assessed using paired t-tests.

Interviews and artifacts will be analyzed through a process of reflexive thematic analysis [11] informed by fourth-generation CHAT [12]. CHAT provides a framework for examining processes of transformation as teacher candidates engage in sense-making and navigate tensions and challenges across four overlapping activity systems: course, classroom, home, and community activity. ENA will then be used to model co-occurrences among coded elements within interviews and artifacts. ENA will support analysis of how teacher candidates connect ideas related to environmental justice, pedagogy, identity, and action across these activity systems. In particular, ENA will be useful in determining whether the course differentially impacts teacher candidates from different subject area specializations and across the three cohorts. In the final stage of analysis, survey data will be used to identify high- and low-transformation cases, and ENA will be used to visualize differences in the discourse between these groups. These network structures are interpreted alongside thematic analysis in order to reconnect quantitative representations with the thick description and contextual nuance of participants'experiences.

4 Preliminary Findings

Preliminary findings draw on survey data from the first two cohorts and interview data from the first cohort. Initial thematic analysis of interview data revealed that the most frequently coded teacher outcomes were uncovering connections between EJ concepts and teacher candidates' focal subjects, realizing the interconnectedness of human societies and natural systems (i.e., development of socioecological consciousness), and increasing environmental and pedagogical knowledge. Teacher candidates most frequently noted that integrating EJ content would benefit students by connecting learning to real-world challenges and students' lived experiences. Further analysis using ENA is needed to understand which course components are most closely linked with outcomes, and how different groups of teacher candidates (i.e. by subject area, experience, EJ knowledge, etc.) differ in their understandings of EJE and the outcomes they identify from engaging with this work. Understanding how the course differentially impacted these groups is essential, as much of our current knowledge around preparing teachers for environmental education is solely related to science education.

5 Expected Contribution

Empirical research examining teacher education for EJE remains limited [13]. This study contributes to environmental and teacher education research by examining how pre-service teachers develop the knowledge, dispositions, and pedagogical capacities needed to integrate environmental justice into their future practice. By investigating a teacher preparation course grounded in local environmental issues and interdisciplinary learning, the study will provide insight into the challenges, tensions, and catalysts that shape teacher learning and intended enactment of EJE. In addition, the research extends understanding of how EJE-focused professional learning may

influence teachers' lives beyond the classroom by examining changes in pro-environmental behaviors, nature connectedness, and socio-ecological consciousness.

References

1. Pizmony-Levy, O. (2023). Two worldviews, one planet: U.S. public views on climate change education in a time of political polarization.
2. Will, M., Prothero, A. (2023, March 7). Students want climate change education. Most teachers don't get enough training. Education Week. https://www.edweek.org/teaching-learning/students-want-climate-change-education-most-teachers-dont-get-enough-training/2023/01
3. Yates, K., Reefer, A., Robertson, D., Hubbard-Sanchez, J., Huss, J., & Wilder, M. (2019). Educators' perceptions of environmental education and professional development in teacher preparation programs. Applied Environmental Education & Communication, 18(3), 207–218. https://doi.org/10.1080/1533015X.2018.1451411
4. Newall, M., & Patino, T. (2020, July 7). Teachers agree that climate change is real and should be taught in schools. Ipsos. https://www.ipsos.com/en-us/news-polls/teachers-agree-climate-change-real-and-should-be-taught-schools-04-22-2019
5. North American Association for Environmental Education. (2022). The state of climate change education: Findings from a national survey of educators. In naaee.org. EDGE Research. Retrieved March 28, 2024, from https://naaee.org/sites/default/files/2023-02/NAAEE_State%20of%20Climate%20Change%20Education%20Report_SUBMITTED%2012_12_22%5B1%5D.pdf
6. Peloso, J. (2024). Environmental justice education: Empowering students to become environmental citizens. Penn GSE Perspectives on Urban Education. https://urbanedjournal.gse.upenn.edu/archive/volume-5-issue-1-spring-2007/environmental-justice-educationempowering-students-become-envi
7. Hecht, M., & Jadallah, C. C. (2023). Learning with water: Centering more-than-human interactions in science learning for more just social-ecological futures. Digital Culture & Education, 14(5).
8. Stewart, K. D., Burke, C. J. F., & Askari, E. (2024). Change day: How a high school environmental justice class inspired student agency and prompted civic action. Race Ethnicity and Education, 27(7), 970–988. https://doi.org/10.1080/13613324.2022.2069744
9. Morales-Doyle, D. (2017). Justice-centered science pedagogy: A catalyst for academic achievement and social transformation. Science Education, 101(6), 1034–1060. https://doi.org/10.1002/sce.21305
10. Shaffer, D. W. (2017). *Quantitative ethnography*. Cathcart Press.
11. Braun, V., & Clarke, V. (2019). Reflecting on reflexive thematic analysis. Qualitative Research in Sport, Exercise and Health, 11(4), 589–597. https://doi.org/10.1080/2159676x.2019.1628806
12. Grimalt-Álvaro, C., & Ametller, J. (2021). A cultural-historical activity theory approach for the design of a qualitative methodology in science educational research. International Journal of Qualitative Methods, 20, 16094069211060664. https://doi.org/10.1177/16094069211060664
13. Miles, M., Schindel, A., Haq, K., & Aziz, T. (2025). Advancing environmental justice education: A critical review of research and practice. Environmental Education Research, 31(7), 1461–1480. https://doi.org/10.1080/13504622.2025.2483443

Supporting Novice Teachers in Responsive Mathematics Teaching: A Stages of Concern Analysis Across Three Phases of Preparation

Riku Sayuj*

University of Pennsylvania, Graduate School of Education
*Corresponding author: sayuj7@upenn.edu

Abstract. Novice mathematics teachers, as they enter the profession, face a long road between learning about a more responsive style of teaching in a university methods course and actually carrying it out in a classroom of their own. This study uses the Stages of Concern construct, part of the Concerns-Based Adoption Model from research on how people respond to change, to trace how ten novice teachers in a US teacher preparation program experience that road. I compare each teacher's stance toward learning while they are in a mathematics methods course against their stance once they are placed in a classroom alongside an experienced mentor teacher, still connected to their university coursework. Half of the ten teachers received an additional, more intensive layer of mentoring support (via the MPACT project intervention) and half received the program's standard mentoring only. Using epistemic network analysis on classroom observation field notes, written reflections, and interviews with the novice teachers, their methods instructors, and their mentor teachers, I ask whether and how each teacher's learning stance changes between the methods course setting and the mentored classroom setting, and whether the additional mentoring support is associated with a different pattern of change. This comparison is part of a larger dissertation that follows novice teachers through three settings of preparation as they take up Responsive Math Teaching, an approach to mathematics instruction that asks teachers to respond in the moment to what students already know rather than follow a fixed script.

Keywords: Stages of Concern, epistemic network analysis, teacher preparation, responsive mathematics teaching, mentoring.

1 Goals of the Research

My broader dissertation asks how novice mathematics teachers develop beliefs, skills, and confidence to teach responsively across three phases of preparation, using Stages of Concern to track development over time. The ICQE study examines whether each teacher's stance toward learning to teach responsively changes from the methods course to the mentored classroom, and whether an additional layer of mentoring support is associated with a different pattern of change.

2 Background of the Project

Responsive Math Teaching (RMT) asks teachers to listen closely to what students already know and adjust instruction accordingly rather than follow a fixed lesson. Penn GSE researchers developed and studied RMT in partnership with the School District of Philadelphia [1]. Because RMT can displace the more scripted forms of instruction novices experienced in their own schooling and may encounter in placements, I treat it as an innovation in the Concerns-Based Adoption Model (CBAM) sense [2].

Stages of Concern describes seven concerns grouped as unrelated, self, task, and impact. Teachers do not always move through them in strict order [5]. I examine the methods-course-to-mentored-classroom transition, where novices first put RMT into practice with real students and the practical demands of teaching may reorganize their concerns.

3 Methodology

This study is part of MPACT (Mentoring Practice for Advancing Classroom Teaching), which examines mentor-novice conversations and novice teacher development. The sample includes ten novice teachers in an urban teacher preparation program: five received an additional intensive layer of mentoring support and five received standard mentoring only. I track the same teachers across the methods-course and mentored-classroom phases.

Methods-course data include observation field notes, written reflections, and interviews with novices and methods instructors; mentored-classroom data include observation field notes and interviews with novices and mentors. I use epistemic network analysis (ENA) [3, 4] to model co-occurrences of Stages of Concern and learning-stance codes, visualizing how each teacher's network shifts across phases and whether patterns differ between mentoring groups. The developing codebook combines Stages of Concern with relational and authority-oriented codes from earlier mentor-novice analyses.

4 Preliminary or Expected Findings

Data collection and coding are still in progress, so I do not yet have findings. I examine whether concerns shift from self and task toward impact with experience, or whether entry into a real classroom temporarily returns teachers to self and task concerns before impact concerns develop. The latter would align with nonlinear patterns documented in recent CBAM research [5]; patterns may also differ between mentoring groups.

5 Expected Contributions

This comparison uses ENA to follow the same small group across two structurally different settings rather than separate cohorts. It examines how concerns about RMT reorganize across settings and tests whether combining Stages of Concern with relational and authority-oriented learning-stance codes yields interpretable within-subject networks.

References

1. Ebby, C.B., Hess, B., Pecora, L., Valerio, J.: "Teaching them how to fish": Learning to learn and teach responsively. CPRE Working Papers, University of Pennsylvania Graduate School of Education, Philadelphia (2021)
2. Hall, G.E., Hord, S.M.: Implementing Change: Patterns, Principles, and Potholes, 4th edn. Pearson, Boston (2015)
3. Shaffer, D.W.: Quantitative Ethnography. Cathcart Press, Madison, WI (2017)
4. Shaffer, D.W., Collier, W., Ruis, A.R.: A tutorial on epistemic network analysis: analyzing the structure of connections in cognitive, social, and interaction data. J. Learn. Anal. 3(3), 9-45 (2016). https://doi.org/10.18608/jla.2016.33.3
5. Zhao, G., Fan, B., Wang, Y., Aliyev, M., Sun, X.: Investigating Chinese in-service teachers' concerns about teaching thinking using the Concern-Based Adoption Model (CBAM). J. Educ. Dev. 8(2), 1-20 (2024). https://doi.org/10.20849/jed.v8i2.1426

The Role of Culture and Identity Conflict on Leadership Identity Development in College Students

Catherine Zhang[1]

[1] University of Pennsylvania, Philadelphia, PA 19104, USA
clzhang@upenn.edu

Abstract. This study examines how undergraduate student leaders navigate identity conflicts and how their navigation, social identities, and organizational and cultural contexts shape their leadership identity development. The study employs an ethnographically informed design with college students in the U.S. and China, proposing Quantitative Ethnography (QE) and Epistemic Network Analysis (ENA) to enable cross-contextual comparisons among sites, organizations, and participants. Findings will contribute to a culturally responsive understanding of the influences that shape student leadership development.

Keywords: leadership identity development, identity conflict, cultural influence

1 Introduction and Background

Higher education represents a critical context for students' identity formation and leadership development. While colleges frequently frame their missions around developing the next generation of leaders, the processes of developing a leadership identity remain obscured [2]. Student leaders navigate multiple stakeholders and expectations, and often may encounter identity conflict, or the tension between the values, beliefs, norms, and expectations attached to multiple identities [3]. Identity conflict -- and the roles of social identity, organizational, and cultural context -- remain undertheorized as mechanisms for understanding leadership identity development. Thus, this study asks: (1) What identity conflicts do student leaders encounter, and how does the process of navigating them shape students' leadership identity development? (2) How do students' social identities and organizational and cultural contexts shape how these conflicts are perceived and negotiated, and how leadership identity is developed?

Leadership identity development refers to the dynamic, socially constructed process through which individuals come to understand and claim a sense of themselves as leaders, progressing from positional to relational understandings of leadership [4]. Critical scholars have challenged leadership identity developmental frameworks to further account for how power and social identity shape their trajectories [2]. Thus, this study utilizes the Critical Campus Ecology Model (CCEM) [6] to situate leadership identity development within the microsystem, mesosystem, exosystem, macrosystem, and chronosystem, and to name how oppression is embedded across all levels. This provides

scaffolding for examining how identity conflicts emerge across different levels of students' ecological contexts.

2 Methodology

This study employs an ethnographically informed design, where data collection includes semi-structured ethnographic interviews with student leaders in both the U.S. and China, additional contextual interviews, participant observation across organizational settings, document analysis, and analytical memos. [7] Participants are selected based on having completed most of their K-12 schooling within that national context, grounded in the recognition that students' earliest leadership memories are shaped by formative educational environments. Quantitative Ethnography is integrated as an analytical layer, utilizing ENA to model and visualize connections between key constructs for cross-cultural comparisons, within-group comparisons, and longitudinal shifts [9].

3 Expected Findings & Contributions

Initially drawing from 23 interviews, preliminary findings have emerged, namely, the tension between personal and professional relationships and between belonging to the group and leading it. Students have shared that, especially in peer leadership roles where there is little formal authority, it can be difficult to do what is "right," especially if it feels ostracizing. These conflicts have shaped how students view interpersonal communication and their own self-awareness within their leadership identity. Additionally, across both U.S. and Chinese contexts, students' earliest leadership memories consistently reference early school or family experiences, often mentioning a teacher or family member asking them to help other students or siblings. This suggests that students' first encounters with leadership are relational and initiated by authority.

Theoretically, this study expects to advance understandings of identity conflict as a developmental mechanism for leadership identity development, extending critical leadership scholarship by centering prior socialization and cultural context as constitutive factors. Practically, the findings will aim to equip educators and leadership program designers with frameworks for supporting students, particularly those from marginalized communities and in diverse cultural contexts, as they navigate the tensions inherent in leadership development.

References

1. Chunoo, V. S., & Torres, M. (2023). Critical perspectives on leadership identity development. *New Directions for Student Leadership*, 2023(178), 55–64. https://doi.org/10.1002/yd.20554
2. Horton, K. E., Bayerl, P. S., & Jacobs, G. (2014). Identity conflicts at work: An integrative framework. *Journal of Organizational Behavior*, 35(S1), S6-22.

3. Komives, S. R., Longerbeam, S. D., Owen, J. E., Mainella, F. C., & Osteen, L. (2006). A Leadership Identity Development Model: Applications from a Grounded Theory. *Journal of College Student Development*, 47(4), 401–418. https://doi.org/10.1353/csd.2006.0048
4. Noble, C. E. (2023). Self, context, and power: Introducing the critical campus ecology model. *New Directions for Higher Education*, 2023(204), 23-33.
5. Rinaldo, R., & Guhin, J. (2022). How and why interviews work: Ethnographic interviews and meso-level public culture. *Sociological Methods & Research*, 51(1), 34-67.
6. Shaffer, D. W. (2017). Quantitative ethnography. Madison, WI: Cathcart Press.

Towards an Affordance-Based Taxonomy of GenAI-Assisted Writing Tools

Yixin Cheng[0000−0002−8596−7944]

Monash University, Australia
yixin.cheng@monash.edu

Abstract. Generative AI (GenAI) tools are widely used by learners, and this trend is poised to continue. These tools are transforming various activities, including writing. Writing is a critical skill in many domains; however, little is known about the specific skills that should be assessed when learners use GenAI to write. This thesis addresses this gap by (a) examining the affordances of widely used GenAI tools to understand how these tools enhance or hinder the writing process; (b) examining the authentic use of GenAI tools within communities of practice to identify the skills that should be assessed; and (c) conducting a larger-scale observational study to provide a more generalisable and rigorous investigation of the skills inherent to writing with GenAI. Together, this thesis aims to lay the foundation for developing a valid assessment framework for writing with GenAI.

Keywords: Assessment · Writing Skills · Generative AI.

1 Goals of the Research

The goal of this thesis is to identify the key skills involved in writing with GenAI and to lay the foundation for a valid assessment framework for GenAI-mediated writing. Writing is treated as a broad and socially important practice that spans educational, professional, and everyday contexts. As GenAI tools become embedded in writing activities, it becomes increasingly important to understand not only whether learners can produce high-quality texts, but also how they use AI to plan, generate, evaluate, revise, and take responsibility for their writing. The research is guided by three research questions:

RQ1: What are the affordances of GenAI tools for writing?

RQ2: How does writing with GenAI compare to writing without it?

RQ3: How are GenAI tools used in authentic writing practices over time?

2 Background of the Project

Writing assessment traditionally relies heavily on written products as evidence of learners' writing ability. However, GenAI complicates this assumption because

human and AI contributions may be intertwined throughout the writing process. When a learner uses GenAI to generate ideas, summarise information, draft text, revise arguments, improve language, or provide feedback, the final product alone may not show what the learner knew, decided, evaluated, or controlled[4]. This obfuscation creates a problem for assessment validity: if the evidence no longer clearly supports claims about learners' skills, then assessment practices need to be reconsidered. This thesis is informed by evidence-centred design[2], which emphasises the need to clearly define the claims an assessment intends to make, the tasks that elicit evidence, and the evidence used to support those claims. The thesis focuses particularly on the student model: the skills that should be assessed when students write with GenAI. It is also informed by affordance theory[1] and mediated action[3], which frame GenAI tools not as neutral additions to writing but as mediators that shape what writers can do, how they act, and what skills become necessary.

3 Methodology

The thesis includes three connected studies. Study 1 examines GenAI writing affordances through a systematic literature review, a review of 28 widely used GenAI writing tools, and a case study with eight adult writers. Data include published studies, tool interface features, documentation, screen recordings, interviews, and survey responses. Analysis combines affordance coding, logistic principal component analysis, and qualitative analysis of tool use.

Study 2 examines how GenAI-supported writing compares with other writing and support contexts through a set of empirical studies. One study compares learners' question asking when interacting with GenAI versus a human tutor during a writing task, using quantitative ethnographic methods to analyse question types, conversational patterns, and links to performance. A second study compares writing with and without GenAI across four Monash workshops using Trace, a web-based writing environment that captures fine-grained writing interactions. A third study collects longitudinal data in a Monash Law context to examine how GenAI support is used in a disciplinary writing setting over time. Across these studies, data include trace logs, written products, chat histories, and interaction records. Analysis combines qualitative coding, regression, mediation analysis, and epistemic network analysis.

Study 3 investigates authentic GenAI-supported writing over time using the Trace browser extension. Participants will include academics and doctoral students who regularly engage in academic or professional writing. Data will include trace logs, selected screen recordings, interviews, and reflection logs. Analysis will examine how GenAI-supported writing skills emerge, stabilise, and change across authentic contexts.

4 Preliminary or Expected Findings

Preliminary findings from Study 1 show that GenAI writing tools can be characterised by a coherent affordance space that remained substantially stable across 2025 and 2026. The first dimension contrasted editorial refinement with planning and external knowledge access, while the second dimension contrasted guidance-oriented support with elaboration and drafting. These findings suggest that GenAI tools do not support writing in a uniform way, but instead mediate different writing activities, including revising existing text, planning, sourcing information, guiding decisions, and generating extended drafts.

Findings from the question-asking study show that learners asked substantially more questions when interacting with GenAI than with a human tutor. GenAI interactions tended to involve one-off, pragmatic questions focused on conceptual and procedural knowledge, whereas human tutor interactions involved longer conversational exchanges and more indirect requests for feedback. Epistemic network analysis showed that question-asking patterns mediated the relationship between condition and performance, suggesting that pragmatic question asking may be an important skill in GenAI-supported writing.

For the workshop study, we have developed a codebook for analysing events in GenAIassisted writing. The codebook assigns one or more Process codes to each event to capture what action is being performed, and, when applicable, one or more Product codes to capture what aspect of the written artefact is being attended to. It is designed for web-based writing environments that include GenAI chat interfaces, text editors, search engines, and web pages. The codebook is currently being applied to the workshop data to identify how students plan, seek information, interact with GenAI, evaluate outputs, revise text, and manage the written product across AI-supported and non-AI conditions.

Data collection is also underway for the Monash Law study, which will provide longitudinal evidence of how students use GenAI in a disciplinary writing context. Together, these studies are expected to identify skills that are shared across writing contexts, such as task understanding, information seeking, planning, monitoring, evaluation, and revision, as well as skills that become more visible in GenAI-supported writing, such as prompting, evaluating AI output, verifying information, integrating AIgenerated content, and maintaining authorial control. Study 3 will extend this work by examining how these skills develop in authentic writing practice over time.

5 Expected Contributions

This thesis will contribute to research on writing assessment, GenAI in education, and quantitative ethnography. Substantively, it will provide an empirically grounded framework of skills involved in writing with GenAI. This framework can inform future assessments that move beyond evaluating final written products and instead attend to the processes through which writing is produced.

Methodologically, the thesis contributes by combining multiple forms of process evidence, including trace logs, screen recordings, interviews, reflections, written products, and GenAI conversation histories. By using quantitative ethnographic approaches such as epistemic network analysis, the thesis aims to make complex human–AI writing processes visible and interpretable. More broadly, the work responds to a pressing educational challenge: how to assess writing validly when writing is increasingly mediated by GenAI tools.

References

1. Gibson, J.J.: The Theory of Affordances: (1979). In: The People, Place, and Space Reader. Routledge (2014), num Pages: 5
2. Mislevy, R.J., Haertel, G.D.: Implications of evidence-centered design for educational testing. Educational measurement: issues and practice **25**(4), 6–20 (2006)
3. Wertsch, J.V.: Mind As Action. Oxford University Press, Incorporated, Oxford, UNITED STATES (1998), http://ebookcentral.proquest.com/lib/monash/detail.action?docID=1591253
4. Yancey, K.B.: Looking back as we look forward: Historicizing writing assessment. College Composition & Communication **50**(3), 483–503 (1999)

A Quantitative Ethnographic Exploration of Epidermolysis Bullosa Patient Journeys in China

Yusong Zhang[1][0009-0001-3110-9194] and Jiaqi Liu[1][0009-0007-1668-8363]

[1] JC School of Public Health and Primary Care, The Chinese University of Hong Kong, Shatin NT, Hong Kong SAR

Abstract. This study aims to explore how QE can be applied to examine how medical, familial, social, and other multidimensional factors interact to shape diverse disease trajectories among individuals living with epidermolysis bullosa (EB) in China. Existing qualitative data have been collected through in-depth interviews with patients and caregivers and analyzed using thematic analysis. We aim to explore how QE approaches can provide additional perspectives for understanding the complex relationships among different dimensions of rare disease patient journeys.

Keywords: Epidermolysis Bullosa, Patient Experience, Patient Journey, Rare Disease Management

1 Area and Domain of Research

This proposed research study is in the domain of Health Sciences and Healthcare Improvement.

2 Background

Epidermolysis bullosa (EB) is a group of rare genetic skin disorders characterized by skin fragility, recurrent blistering, chronic wounds, and multisystem complications. The lifelong physical and psychosocial burden of EB substantially affects patients' daily lives. Existing EB research has primarily focused on genetic mechanisms, clinical management, and therapeutic or supportive interventions. However, despite China's large population and the considerable number of individuals affected by EB, patients remain widely dispersed and socially isolated, and current understanding of their experiences and needs remains largely fragmented. Medical-focused research captures only part of patients' challenges, underscoring the need to explore the multidimensional factors shaping their disease journeys.

Patient journey mapping has gained prominence as a framework for understanding experiences among individuals with rare diseases. It provides a patient-centered approach to capture long-term and multifaceted experiences while enabling in-depth exploration of challenges and needs within small and geographically dispersed rare disease populations. Patient journey frameworks typically consider key elements shaping patients' experiences, including multiple stakeholders involved in care [1],

different stages of disease progression [2], and broader social and contextual factors such as family resources, healthcare accessibility, and social support systems [3]. However, further research is needed to understand how these diverse elements intersect and interact to shape individual patient journeys.

3 Research Objectives

This research aims to explore how medical, familial, social, and other factors interact and collectively shape diverse patient journeys among individuals with EB in China. The expected outcomes are to improve understanding of EB patients' experiences, develop better support strategies with key stakeholders, and inform healthcare insurance and social support policies. Potential research questions include: (1) How do medical, familial, social, and broader factors interact to shape different patient journey trajectories among individuals with EB in China? (2) What similarities and differences exist among EB patient journey patterns, and what mechanisms underline these patterns? We have collected online and offline observations and 33 in-depth interviews with patients and caregivers, and completed qualitative coding and thematic analysis, which will serve as the foundation and reference for future QE exploration.

4 Prior Experience and Expectations

We have reviewed introductory concepts and applications of QE, but have not yet conducted QE research. We believe our study could benefit from QE to better understand how multiple factors interact and shape diverse patient journeys. We are interested in learning how to use QE to reanalyze and organize existing data, or guide future data collection, as well as how to visualize relationships among complex factors beyond traditional qualitative text-based presentations.

References

1. Benson, M., Albanese, A., Bhatia, K. P., Cavillon, P., Cuffe, L., König, K., Reinhard, C., & Graessner, H. Development of a patient journey map for people living with cervical dystonia. Orphanet Journal of Rare Diseases 17(1), 130 (2022).
2. Davies, E. L., Bulto, L. N., Walsh, A., Pollock, D., Langton, V. M., Laing, R. E., Graham, A., Arnold-Chamney, M., & Kelly, J. Reporting and conducting patient journey mapping research in healthcare: A scoping review. Journal of Advanced Nursing 79(1), 83–100 (2023).
3. Hausmann, J. S., Lomax, K. G., Shapiro, A., & Durrant, K. The patient journey to diagnosis and treatment of autoinflammatory diseases. Orphanet Journal of Rare Diseases 13(1), 156 (2018).

Constructing Sustainability: Narratives of the Olympic Legacy

Thomas Briskey[0009-0006-1707-506X]

Pepperdine University, Malibu, CA, USA
thomas.briskey@pepperdine.edu

Abstract. This research proposal explores how sustainability is conceptually constructed and communicated within Olympic host city reporting and International Olympic Committee (IOC) rhetoric. Sustainability has become a priority for the IOC, but research tends to focus on performance and outcomes of sustainable actions rather than how it is framed in narratives. Grounded in Institutional Theory, this study will use Quantitative Ethnography and Epistemic Network Analysis to explore relationships between sustainability-related concepts within Olympic literature and seeks to identify how they connect to themes such as governance, legacy, and community impact, contributing to both Olympic Studies, Institutional Theory, and Quantitative Ethnography.

Keywords: Organizational Communication, Olympic Studies, Institutional Theory, Sustainability

1 Area and Domain of Research

The proposed research study is in Organizational Studies, focusing on organizational communication, sustainability, and legitimacy. Using Quantitative Ethnography (QE), this study hopes to examine relationships among concepts and visualize the structure of sustainability narratives across Olympic host cities.

2 Background

Sustainability has become a central focus of the International Olympic Committee (IOC), with Olympic Charter reforms such as the Agenda 2020 positioning it as an organizational priority, while also mandating host cities to report on performance during the Games [1]. The IOC has developed strategic initiatives and guiding principles for host cities and organizing committees to implement around Environmental, Social, and Governance (ESG) to help bolster this commitment.

Despite the increase of sustainability rhetoric from the IOC and organizing cities, the material outcomes of "sustainable games" has been questioned [4]. Most Olympic research in this area has focused on the discrepancy between sustainability commitments and their outcomes and the realities of hosting a mega-event, while less attention is given to the institutional rhetoric and narratives within the IOC. The goal of this study is to understand how sustainability is constructed and legitimized through

institutional reporting, and how sustainable concepts connect and relate within the narratives.

Institutional Theory provides the foundation for how institutions adopt and communicate socially valuable concepts such as sustainability [2]. DiMaggio & Powell and Suchman [2,3] suggested that organizations conform to social expectations to maintain legitimacy, where legitimacy comes from social perceptions of actions. From this perspective, sustainability reports are the narratives that communicate organizational legitimacy.

3 Research Objectives

The purpose of this study is to explore how host cities construct and communicate sustainability through their post-game legacy and impact reports. It focuses on two research questions:

1. How is sustainability framed within Olympic host city reports?
   a. What conceptual networks emerge around sustainable discourse?

The study anticipates using host city game reports from the 2012 – 2024 Olympic Games (via IOC Library Archives), focusing on the sections that describe sustainability progress and outcomes. The goal for data collection is to reach saturation allowing for meaningful interpretation and results, which may require more, or less, Olympic reports for analysis [5]. This study is currently in the proposal development phase.

4 Prior Experience and Expectations

My previous experience in QE has been in a classroom environment where we were taught QE basics and executed a "mock" study. I am interested in the different ways veteran researchers use QE/ENA to tell stories and leverage ENA's visual outputs. I also hope that by attending ICQE I can integrate feedback and suggestions for my study from QE experts, specifically on study design, segmenting, and coding. As I enter my dissertation phase, I hope to learn about data analysis as well, and best practices for organizing and understanding the data

Specific questions for the experts include:

1. Best practice - segmentations be for lengthy reports
2. Connecting Institutional Theory into a QE/ENA analysis
3. Design of a QE study that elicits good visual outputs

This proposed study on Olympic Sustainability would contribute to the growing literature around greenwashing and corporate responsibility, while also highlighting existing challenges within the IOC and other NGO's that operate on the global stage.

References

1. IOC. (2023). Sustainability Progress Report 2021–2024 international olympic committee. https://stillmed.olympics.com/media/Documents/International-Olympic-Committee/Documents/Reports/IOC-Sustainability-Progress-Report-2021-2024.pdf
2. DiMaggio, P.J. and Powell, W.W. (1983) The Iron Cage Revisited: Institutional Isomorphism and Collective Rationality in Organizational Fields. American Sociological Review, 48, 147-160.https://doi.org/10.2307/2095101
3. Suchman, M. C. (1995). Managing legitimacy: Strategic and institutional approaches. Academy of management review, 20(3), 571-610.
4. Boykoff, J. (2025). Greenwash Gold?: The Paris 2024 olympics. Capitalism Nature Socialism, 36(2), 171–178. https://doi.org/10.1080/10455752.2025.2464556
5. Shaffer, D. W. (2018a). Quantitative ethnography. Cathcart Press.

Cultivating Self-Efficacy in Collaborative Repair: A Participatory Quantitative Ethnography Approach to Designing Spaces for Tinkering

Kohei Aida[1][0009-0000-5009-4437]

[1] Advanced Institute of Industrial Technology, Tokyo, JAPAN
b2401ka@aiit.ac.jp

Abstract. This Research Agenda Development (RAD) proposal outlines a plan to investigate interaction dynamics within community-based repair activities in Kokubunji, Tokyo. Moving beyond traditional repair services, these activities serve as collaborative spaces prioritizing the requester's subjective engagement and emotional recovery over functional restoration. By analyzing 1-to-n dialogues, this study aims to uncover how active involvement in the repair process fosters self-efficacy. Crucially, we aim to articulate how this evolving self-efficacy triggers dynamic role fluidity where requesters transition into active repairers within these spaces for tinkering. To evaluate this subjective process and practice meta-design, we adopt a Participatory Quantitative Ethnography (PQE) approach. We propose using Epistemic Network Analysis (ENA) not merely as an analytical tool, but as a boundary object for co-interpretation, empowering the community to self-reflect and continuously reshape their environment.

Keywords: Collaborative Repair, Participatory Quantitative Ethnography (PQE), Self-Efficacy, Meta-Design, Tinkering.

1 Introduction and Research Background

Community-based repair initiatives, such as "Repair Café", offer unique environments for collaborative problem-solving and design. This research focuses on micro-interactions within the "Repair Town" project in Kokubunji, situated in suburban City of Tokyo . Crucially, this initiative is not a commercial service; it positions the requester as the primary actor in a collaborative exploration to realize their hope of reviving a "broken" object.

We hypothesize that the requester's engagement in the repair process triggers a dynamic change in their self-efficacy. Crucially, this heightened self-efficacy acts as a catalyst, transforming the participants themselves from requesters into active repairers. Ultimately, this work aims to articulate how this dynamic role fluidity contributes to designing and sustaining resilient "spaces for tinkering"—a community environment that embraces mistakes and playful experimentation.

The core research questions are: (1) How do collaborative interactions at repair sites shape the trajectory of self-efficacy for both requesters and repairers? (2) How can the

participants themselves utilize Participatory QE (PQE) to evaluate their self-efficacy and meta-design these spaces for tinkering?

2. Proposed QE Approach & Data

The study will collect multi-modal data from local repair events, focusing on 1-to-n interactions. Data sources include audio transcriptions and video observations, coded for elements such as Triggers (invitations to participate), Proactivity (handling tools and objects, sharing ideas), and Self-Efficacy (expressions of confidence, mutual affirmation).

Because self-efficacy is inherently subjective, interpreting its development requires the participants' own voices. To achieve this and practice meta-design, this study adopts a Participatory Quantitative Ethnography (PQE) approach. ENA will be applied to model the co-occurrence of codes. Crucially, the resulting ENA discourse networks will be presented back to the community participants. By establishing a symmetrical premise where representations are open to interpretation, participants will actively co-construct meaning by modifying codes and adding connections based on their emic perspectives.

3. Expected Contribution & Discussion Points

By framing repair practices as mutual learning opportunities, this research provides an empirical foundation for evaluating informal educational spaces. Furthermore, integrating PQE addresses the lack of systematic evaluation in participatory design, demonstrating how communities can continuously measure their own self-efficacy. At ICQE26, I aim to engage with veteran QE scholars on:

- Methodological challenges in applying PQE and using ENA networks as reflective tools for community members in unstructured, non-academic settings.
- Strategies for capturing the catalytic moments where heightened self-efficacy triggers role fluidity within ENA discourse networks.
- Techniques for integrating multi-modal data (e.g., verbal dialogue and embodied physical actions like handling tools) into ENA to accurately capture the non-verbal essence of tinkering.

References

1. Vega, H., & Arastoopour Irgens, G. (2021). Constructing Interpretations with Participants through Epistemic Network Analysis. ICQE.

Detecting Traces of Scaffolding in a Multi-Agent Simulation Environment

Okpokiri Williams Chidera[1][0009-0001-0133-1835], Chengjiu Yin[2][0000-0003-1492-5250] and Juan Zhou[3][0000-0002-2995-4559]

[1] Department of Environmental & Information Science, Tokyo City University, Japan
[2] Research Institute for Information Technology, Kyushu University, Japan
[3] Faculty of Informatics, Department of Sociology & Media Studies, Tokyo City University, Japan

g2683001@tcu.ac.jp, yin.chengjiu.247@m.kyushu-u.ac.jp, juan.z.kt@gmail.com

Abstract. We propose a simulation-based study investigating traces of peer scaffolding in free-form multi-agent discourse. Four similarly configured agents with varying epistemic profiles discuss ill-structured topics across 18 sessions. Transcripts are coded and compared to a similar-in-structure human dataset using a scaffolding rubric adapted from a previous study. Preliminary findings show comparable rates of epistemic moves, while social signaling moves (questioning & confirming) occurred frequently in the simulated data. Repair moves were absent across all simulation conditions, and building moves surfaced only when explicitly prompted, suggesting that multi-agent simulation can reproduce some surface-level features of peer discourse but diverge in other interactional features.

Keywords: Discourse Analysis, Multi-Agent Simulation, Collaborative Learning, Scaffolding

1 Area and Domain of Research

Computer-Supported Collaborative Learning (CSCL), Quantitative Ethnographic (QE) Methods, Multi-Agent Simulation

2 Background

Collaborative learning is central to the learning sciences, as metacognitive processes are often distributed across group members rather than residing solely within individuals [1]. Investigating effective collaboration, particularly peer scaffolding, is both costly and ethically challenging at scale. Multi-agent simulations provide a promising alternative by creating agents that approximate the characteristics of human learners, thus enabling the generation of substantive collaborative discourse data at scale [2,3]. However, existing simulations often impose roles and explicit behavioral instructions [4]. This raises a key question: do scaffolding traces emerge from agent

interaction, or from the simulation design? We address this by instantiating four GPT-4o agents sharing the same configurations but different epistemic starting points, motivated by Tudge's [5] discussion of divergent epistemic starting positions in peer collaboration. No agent was assigned a privileged role, and discussion moves were not explicitly instructed in the prompt. Simulated discussions are then compared against human peer discussion using the same topic and session structure. The scaffolding rubric was adapted from [6] and operationalized by [7].

3 Research Objectives

This study uses a theoretically grounded utterance coding framework to detect discourse moves and the means by which scaffolding occurs [6]. Because identifying scaffolding intentions would require hypotheses about the cognitive states of agents that cannot be inferred from our simulation data, we seek to uncover discourse patterns that can be essential in the design of agents that support collaborative learning. The immediate population is the set of agents instantiated within the AutoGen GroupChat environment [2]. A broader inferential population is collaborative learner groups in text-based asynchronous discussion environments. Our questions are:

1. What traces of scaffolding can we infer from agentic discourse?
2. How do prompt characteristics, including word-count constraints and epistemic profile design, affect the occurrence of discourse moves?
3. How does the distribution of discourse move in simulation diverge from human peer discussions?

Across eighteen sessions of up to 20 turns, we collected agent name, utterance, turn index, reply-to, word count, and timestamps. The main study will begin with a constrained simulation and epistemic starting point setup and systematically remove individual constraints to examine the effects on discourse move detections (RQ2), giving us a broad description of the simulated agent discourse.

4 Prior Experience and Expectations

Building on prior research on the analysis of student educational log data for identifying cognitive and learning-style indices [8], this study seeks structured feedback from QE scholars regarding two methodological decisions: (1) the appropriate moving stanza window size for 20-turn sessions [9] and (2) the validity of applying Epistemic Network Analysis (ENA) to computationally generated discourse as a future analytical step for our study. Can relational structures identified in generated discourse be interpreted through a similar theoretical lens to those observed in human collaborative learning discourse?. More broadly, insights are sought into how QE scholars develop and apply other utterance coding schemes for computationally generated text [10,11].

Acknowledgments: This work was supported by JSPS KAKENHI Grant Number JP25K17091.

References

1. Baker, R., Siemens, G.: Educational data mining and learning analytics. In: Sawyer, R.K. (ed.) Cambridge Handbook of the Learning Sciences, 2nd edn. Cambridge University Press (2014)
2. Wu, Q., et al.: AutoGen: Enabling next-generation LLM applications via multi-agent conversation. arXiv:2308.08155 (2023)
3. Yan, M., et al.: Agentic AI as undercover teammates: Argumentative knowledge construction in hybrid human-AI collaborative learning. arXiv:2512.08933 (2025)
4. Zhang, Z., Zhang-Li, D., Yu, J., Gong, L., Zhou, J., Hao, Z., Jiang, J., Cao, J., Liu, H., Liu, Z., Hou, L., Li, J.: Simulating classroom education with LLM-empowered agents. In: Proceedings of the 2025 Conference of the Nations of the Americas Chapter of the Association for Computational Linguistics: Human Language Technologies, pp. 10364–10379. Association for Computational Linguistics (2025)
5. Tudge, J.R.H.: Processes and consequences of peer collaboration: A Vygotskian analysis. Child Development 63(6), 1364–1379 (1992)
6. Van de Pol, J., Volman, M., Beishuizen, J.: Scaffolding in teacher–student interaction: A decade of research. Educational Psychology Review 22(3), 271–296 (2010)
7. Wong, K.Y.C., Feng, S., Bulathwela, S., Cukurova, M.: Scaffolding reshapes dialogic engagement in collaborative problem solving: Comparative analysis of two approaches. In: Proceedings of the 16th International Conference on Learning Analytics and Knowledge (LAK 2026), pp. 227–238. Society for Learning Analytics Research (2026)
8. Zhou, J., Takada, H., Yin, C.: Analysis of reading behavior based on students' cognitive styles and learning styles. Journal of Internet Technology 25(3), 425–433 (2024)
9. Siebert-Evenstone, A.L., et al.: In search of conversational grain size: Modeling semantic structure using moving stanza windows. Journal of Learning Analytics 4(3), 123–139 (2017)
10. Zhang, S., et al.: Understanding student teachers' collaborative problem solving via epistemic network analysis. Computers & Education 183, 104485 (2022)
11. Zhao, L., et al.: Epistemic network analysis for end-users. Proceedings of LAK '24. ACM (2024)

Evaluation of Machine Translation with Epistemic Network Analysis

Jianxing Yu[1][0009-0002-5073-2472] Tin Yan KOO[1]

Santiago ALONSO SOUSA[1][0009-0005-1180-5087]

[1] City University of Hong Kong
yujianxing@ena3d.org

Abstract. This paper explores how Epistemic Network Analysis (ENA) can be applied to evaluate the quality of Machine Translation (MT). The study is currently in the data collection phase, focusing on domain-specific texts translated by experts. We investigate the combined use of Machine Translation Quality Estimation (MTQE) and ENA to analyze domain-specific MT data, integrating existing MT quality evaluation frameworks, AI-assisted coding, and Large Language Models (LLMs). The proposed framework aims to enhance MT evaluation methods by revealing the relationship between different translation quality dimensions.

Keywords: Machine Translation, Natural Language Processing, Quality Estimation

1 Area and Domain of Research

This proposed research study is in the domain of Machine Translation (MT).

2 Background

The recent years' breakthrough in Machine Translation (MT) is one of the triumphs that Large Language Models (LLM) have achieved. Despite these advances, evaluating translation quality remains a significant challenge due to the inherently uncertainty and subjective nature of translation [1].

Traditional evaluation method such as BLEU [2] and TER [3] evaluates the quality of MT by comparing the translated text with a gold-standard reference using string-matching techniques like n-gram and edit distance. There are also deep learning-based method, such as COMET (Rei et al., 2009), which utilize pretrained multilingual machine-learning model for translation evaluation tasks.

MTQE methods range from feature-based approaches, such as QuEst [4], to modern LLM-based approaches that employ zero-shot or few-shot prompting strategies for translation evaluation [5]. While these methods have demonstrated promising perfor-

mance, many suffer from limited interpretability. Furthermore, human annotation remains necessary for developing and validating evaluation systems, introducing potential subjectivity into the assessment process. Frameworks such as Multidimensional Quality Metrics (MQM) have been developed to standardize error annotation and improve evaluation consistency.

3 Research Objectives

The purpose of this paper is to develop a framework for evaluating the quality machine translation using ENA and MTQE with LLMs. Researchers in the field of MT will be beneficial from a new MT evaluation framework which combines qualitative and quantitative methods. Potential research questions include:

1) How can AI-assisted coding and ENA be utilized to evaluate machine translation quality as a complement to traditional evaluation metrics such as BLEU and COMET?
2) How do different LLM-based translation systems differ in terms of translation quality when analyzed through MQM-based qualitative coding and ENA?

The study is currently in the stage of data collection. We are working with City University of Hong Kong for an AI translation project, where veterinary educational materials are translated by experts and train our own Large Language Models for multimodal translation. For this research, various domain-specific machine translation data (such as the veterinary domain) will be collected. Upon collection of domain specific translation data, data analysis will involve manual and AI-assist coding using framework such as MQM for the MT data and performing ENA analysis on the coded data.

4 Prior Experience and Expectations

I have experienced with multiple Quantitative Ethnographic (QE) tools such as ENA and ONA. My expectation for this study includes exploring how QE can be used as a supplement to the traditional evaluation method for MT content and also gain more experience with using QE methods.

References

1. Zhao, H., et al.: From Handcrafted Features to LLMs: A Brief Survey for Machine Translation Quality Estimation. In: Proceedings of the 2024 International Joint Conference on Neural Networks (IJCNN), pp. 1–10. IEEE, Piscataway (2024).
2. Papineni, K., Roukos, S., Ward, T., Zhu, W.-J.: BLEU: a method for automatic evaluation of machine translation. In: Proceedings of the 40th Annual Meeting of the Association for Computational Linguistics (ACL ’02), pp. 311–318. Association for Computational Linguistics, USA (2002).

3. Snover, M., Dorr, B., Schwartz, R., Micciulla, L., Makhoul, J.: A Study of Translation Edit Rate with Targeted Human Annotation. In: Proceedings of a machine translation evaluation study/workshop, USA (2006)
4. Specia, L., Shah, K.: QuEst – A Translation Quality Estimation Framework. In: Proceedings of a machine translation quality estimation framework presentation/workshop (2013)
5. Kocmi, T., Federmann, C.: Large Language Models Are State-of-the-Art Evaluators of Translation Quality. In: Proceedings of a machine translation evaluation study/workshop, USA (2023)

Exploring Whether Labeling an LLM Interviewer as AI or Human Changes the Structure of Self-Disclosure

Satoshi Kadomatsu[1][0009-0001-5688-1076] and Toru Nagahama[1][0009-0008-5175-057X]

[1] Tohoku University, 6-3-09 Aramaki-aza-Aoba, Aoba-ku, Sendai 980-8579, Japan
kadomatsu.satoshi@gmail.com

Abstract. This proposal builds on a preliminary epistemic network analysis (ENA) study examining whether labeling the same large language model (LLM) interviewer as AI or human changes the structure of self-disclosure. In 30-minute career-focused text-chat interviews, participants were randomly assigned to an AI-label or human-label condition ($N = 6$). The exploratory ENA suggested patterns centered on weakness disclosure under the AI label and positive self-presentation under the human label. We propose a larger confirmatory experiment and seek guidance on the coding scheme and on our ENA and ordered network analysis models.

Keywords: Self-Disclosure, Human–AI Interaction, Social Presence.

1 Area and Domain of Research

This research spans Learning Sciences and Educational Technology as well as the Social and Behavioral Sciences, focusing on artificial intelligence (AI)-mediated support.

2 Background

Self-disclosure conveys internal states [1], which may inform AI-mediated support. Yet evaluative career interviews may elicit defensive impression management [2], potentially discouraging weakness disclosure. Perceiving an interviewer as a computer can reduce evaluation apprehension and increase willingness to disclose [3], but these findings do not reveal how disclosure types co-occur. We therefore model this co-occurrence structure using epistemic network analysis (ENA) [4].

In a preliminary study, participants completed a 30-minute text-chat career interview with the same large language model (LLM) interviewer labeled as either "AI" or "human" ($n = 3$ per condition). A single career advisor coded the utterances into eight categories developed for this study (factual, emotional, weakness, and secret disclosure; lesson drawing; positive self-presentation; hedging; and relationship maintenance), informed by prior research on self-disclosure and impression management [1, 2]. Exploratory ENA with a four-utterance moving stanza window and means rotation suggested that the AI-label network linked weakness disclosure to factual and emotional disclosure, whereas the human-label network linked positive self-presentation to factual disclosure, relationship maintenance, emotional disclosure, and lesson drawing.

3 Research Objectives

We extend this study to a confirmatory experiment addressing two research questions: (RQ1) Does labeling the same interviewer as AI rather than as human shift the co-occurrence structure toward weakness disclosure? (RQ2) Do the AI-label and human-label conditions differ in the ordered relations among disclosure types? The target population consists of working adults reflecting on their careers, with future extensions to learning dialogues. The data will consist of transcripts from 30-minute text-chat interviews, segmented into utterances for coding.

The full-scale experiment is currently being designed. The sample size will be determined through an a priori power analysis, with a provisional target of approximately 30 participants per condition. The design will also include coding-scheme evaluation using Cohen's κ and Shaffer's ρ [5] to test whether agreement generalizes above a prespecified threshold; a manipulation check of perceived interviewer identity plus social-presence and impression-management measures; and ordered network analysis (ONA) [6] for RQ2. We will use means rotation to descriptively visualize group separation, while confirmatory tests will be conducted on projections derived without using group labels, to avoid circular inference.

4 Prior Experience and Expectations

The first author is new to quantitative ethnography (QE) and ENA. We will focus on two aspects: (1) the coding scheme (granularity of the eight categories, whether to merge low-frequency categories, and methods for training coders and assessing inter-rater reliability, including whether nCoder can help scale the coding) and (2) the analytic design (how to justify the size of the moving stanza window used in the preliminary study, compare ENA scores between groups, interpret difference networks, and gauge the appropriateness of ONA for examining ordered relations among disclosure types). More broadly, we hope to establish a sound QE workflow for studying human–AI interactions to illustrate ENA's contribution to this domain.

References

1. Jourard, S.M.: Self-Disclosure: An Experimental Analysis of the Transparent Self. Wiley-Interscience, New York (1971).
2. Bourdage, J.S., Roulin, N., Tarraf, R.: "I (might be) just that good": Honest and deceptive impression management in employment interviews. Pers. Psychol. 71(4), 597–632 (2018).
3. Lucas, G.M., Gratch, J., King, A., Morency, L.-P.: It's only a computer: Virtual humans increase willingness to disclose. Comput. Hum. Behav. 37, 94–100 (2014).
4. Shaffer, D.W.: Quantitative Ethnography. Cathcart Press, Madison (2017).
5. Eagan, B., Brohinsky, J., Wang, J., Shaffer, D.W.: Testing the reliability of inter-rater reliability. In: Proceedings of the 10th International Conference on Learning Analytics & Knowledge (LAK '20), pp. 454–461. ACM, New York (2020).
6. Tan, Y., Ruis, A.R., Marquart, C., Cai, Z., Knowles, M.A., Shaffer, D.W.: Ordered network analysis. In: Damşa, C., Barany, A. (eds.) ICQE 2022. CCIS, vol. 1785, pp. 101–116. Springer, Cham (2023).

Facilitators and Barriers to STEM Persistence for African Diaspora Women in General Chemistry

Jade Pratt[1][0009-0004-2560-4756]

[1] Pepperdine University, Malibu, CA 90263, USA

jadepratt@pepperdine.edu

Abstract. This RAD proposal presents a study that examines factors that act as facilitators and barriers to STEM persistence for African Diaspora women enrolled in general chemistry. The study will focus on the experiential narratives of African Diaspora women enrolled in general chemistry at an R1 research institution in California, USA. Through a quantitative ethnographic lens, Epistemic Network Analysis (ENA) will be employed to explore the connections between these factors that act as facilitators and barriers to STEM persistence.

Keywords: African Diaspora women, STEM, Quantitative Ethnography

1 Area and Domain of Research

The proposed research is in the domain of Social and Behavioral Sciences and Edu cation.

2 Background

Historically, the experiential narratives of Black women in Science, Technology, Engineering and Mathematics (STEM) have often been overlooked or silenced. For this study, the concept of Blackness has been operationalized to a more expansive term, African Diaspora, to provide space for nuanced racial and ethnic categorical associations and to highlight the ways that these identities impact experiences of this demographic of women in STEM disciplines. Particularly, African Diaspora women face unique challenges that are influenced by their racial and gender identities in university-level STEM courses. Their intersectional identities have a significant impact on their desire to continue in the course itself and ultimately, their desire to persist in STEM majors [1-3].

The impact of these foundational courses on confidence and interest for African Diaspora women in STEM cannot be overstated, as they often shape their overall perception of their capabilities and potential success in the field [4,5]. Specifically, general chemistry is one of the first gateway STEM courses that undergraduates encounter, which has a greater impact than any other course in contributing to students' decisions

to change their STEM majors [6]. Successful navigation of undergraduate general chemistry can set a positive tone for future STEM courses, encouraging African Diaspora women to persist and thrive in their chosen disciplines. On the other hand, negative experiences in this foundational course can discourage them from continuing in STEM, highlighting the critical role that this gateway course plays in shaping STEM education trajectories [7].

3 Research Objectives

The primary research questions for this study are: (1) What factors are facilitators or barriers to persistence in STEM majors for African Diaspora women in undergraduate general chemistry courses? (2) What is the relationship between these facilitators and barriers for African Diaspora women in undergraduate general chemistry courses? (3) How do these factors shift or evolve over time during their enrollment in an undergraduate general chemistry course?

The theoretical framework of this study is comprised of Tinto's theory of student integration, intersectionality and misogynoir. By examining overlapping academic, social and peer interactions, this study will explore how these factors contribute to the desire to persist in challenging courses and majors for African Diaspora women [8]. To provide greater clarity on the complexity of this ethnographic group's experience, intersectionality will also be applied. This theory highlights the interconnectedness of identities that cannot be simply understood through a singular recognition of race or gender but instead the additive experiences that these identities combine to create [9]. Misogynoir is another layer of theory that identifies the dual oppression that Black women face due to the combined impacts of sexism and racism. It identifies how negative societal stereotypes translate into feelings of isolation, bias, microaggressions and exclusion in undergraduate general chemistry courses [10].

4 Prior Experience and Expectations

The primary researcher for this study has been intentional in applying QE methodologies to her research work over the past four years. She has previously participated in three data challenges and has presented at ICQE conferences from 2023 to 2025. Currently, she is utilizing QE and ENA methods in her dissertation work and has a focus on multi-faceted engagement with other QE scholars to improve her expertise and build research collaborations through multi-disciplinary engagement. Furthermore, the researcher plans to continue utilizing QE methodology in her post-doctoral work.

References

1. Crenshaw, K. (1989). Demarginalizing the Intersection of Race and Sex: A Black Feminist Critique of Antidiscrimination Doctrine, Feminist Theory and Antiracist Politics. University

of Chicago Legal Forum, 1989(1). https://chicagounbound.uchicago.edu/uclf/vol1989/iss1/8
2. Ong, M., Wright, C., Espinosa, L., & Orfield, G. (2011). Inside the Double Bind: A Synthesis of Empirical Research on Undergraduate and Graduate Women of Color in Science, Technology, Engineering, and Mathematics. Harvard Educational Review, 81(2), 172–209. https://doi.org/10.17763/haer.81.2.t022245n7x4752v2
3. McGee, E. O., & Bentley, L. (2017). The Troubled Success of Black Women in STEM. Cognition and Instruction, 35(4), 265–289. https://doi.org/10.1080/07370008.2017.1355211
4. Ireland, D. T., Freeman, K. E., Winston-Proctor, C. E., DeLaine, K. D., McDonald Lowe, S., & Woodson, K. M. (2018). (Un) hidden figures: A synthesis of research examining the intersectional experiences of Black women and girls in STEM education. *Review of Research in Education*, *42*(1), 226-254.Fink, F.: Article title. Journal **2**(5), 99–110 (2016)
5. Carlone, H. B., & Johnson, A. (2007). Understanding the science experiences of successful women of color: Science identity as an analytic lens. Journal of Research in Science Teaching, 44(8), 1187–1218. https://doi.org/10.1002/tea.20237
6. Fink, A., Frey, R. F., & Solomon, E. D. (2020). Belonging in general chemistry predicts first-year undergraduates' performance and attrition. *Chemistry Education Research and Practice*, *21*(4), 1042-1062.
7. Weston, T. J., Seymour, E., Koch, A. K., & Drake, B. M. (2019). Weed-out classes and their consequences. In *Talking about leaving revisited: Persistence, relocation, and loss in undergraduate STEM education* (pp. 197-243). Cham: Springer International Publishing.
8. Tinto, V. (2012). Leaving college: Rethinking the causes and cures of student attrition (2. ed., paperback ed). Univ. of Chicago Press.
9. Crenshaw, K. (1991). Mapping the Margins: Intersectionality, Identity Politics, and Violence against Women of Color. Stanford Law Review, 43(6), 1241. https://doi.org/10.2307/1229039
10. Bailey, M. & Trudy. (2018). On misogynoir: Citation, erasure, and plagiarism. Feminist Media Studies, 18(4), 762–768.

From Pain Meaning to Clinical Action: Using Quantitative Ethnography to Identify Cultural-Linguistic Pain Phenotypes

Shinye Kim[1], Ji In Woo[1]

[1] University of Wisconsin-Madison, Madison, WI 53706, USA
shinye.kim@wisc.edu

Abstract. Chronic pain assessment depends heavily on how patients describe, contextualize, and negotiate pain with others. Yet dominant pain assessment models often rely on standardized scales that may miss culturally and linguistically patterned meanings of pain. This Research Agenda Development proposal explores how Quantitative Ethnography (QE) can be used to model Korean immigrants' chronic pain narratives as patterned systems of meaning with potential implications for clinical practice. Drawing on Korean-language interviews with adults living with chronic pain, the proposed work examines how concepts such as endurance, weakness, family obligation, causal attribution, medication avoidance, emotional restraint, and healthcare communication become connected in patients' accounts. The central agenda is not simply to identify themes, but to test whether clinically meaningful claims about pain communication are reflected in persistent patterns among coded concepts. This project asks how QE can extend pain research by modeling relational structures of meaning, while also translating those structures into actionable practices for culturally responsive pain assessment. The goal is to develop cultural-linguistic pain phenotypes to inform clinical interviewing, patient-provider communication, and precision pain care.

Keywords: Quantitative Ethnography, Chronic Pain, Korean Immigrants, Cultural-Linguistic Phenotypes, Precision Pain Medicine

1 Background and Rationale

Chronic pain is both a sensory and communicative experience. In clinical settings, pain is rarely directly observable; it must be translated through words, metaphors, gestures, silence, emotional expression, and interactional context. For culturally and linguistically minoritized patients, this translation process can shape whether pain is believed, assessed accurately, and treated appropriately. This proposal explores how QE can be used to study Korean immigrants' chronic pain narratives as relational systems of meaning. Moving beyond whether culture matters for pain, this project examines how cultural-linguistic meanings organize pain experience, communication, and care-seeking to inform culturally responsive and clinically actionable pain assessment.

2 Dataset and Focus

The project draws on Korean-language interviews with adults living with chronic pain in the United States. Interviews addressed cultural and familial pain norms, Korean pain expressions, bodily and emotional descriptions, pain-cause beliefs, endurance, healthcare communication, coping, and changes after migration. Analysis will focus on how "참음"/endurance, culturally specific pain metaphors, emotional restraint, family responsibility, medication avoidance, and communication difficulties with U.S. providers are connected within participants' narratives.

3 Theoretical and Clinical Contribution

Clinically actionable findings require more than broad cultural description. Rather than concluding that Korean immigrants "express pain differently," this project asks how specific meanings, relationships, and communication contexts shape pain disclosure, treatment expectations, and healthcare interaction. These relational patterns could inform more precise clinical interviewing, culturally responsive assessment, and patient-provider communication.

4 Proposed QE Approach

The proposed QE work will begin with qualitative immersion in the transcripts to identify theoretically meaningful and clinically actionable claims. These claims will guide the development of codes, which may include endurance, emotional restraint, family responsibility, medication concern, communication difficulty, treatment seeking, and self-management. The data would be reorganized into sentence-level rows where possible, with each code represented as a binary column. This structure would allow ENA to model how concepts connect within participants' narratives.

5 Long-Term Research Agenda

The long-term goal is to develop a theory of cultural-linguistic pain phenotypes. In precision pain medicine, phenotyping often emphasizes biological, psychological, or behavioral profiles; this project argues that narrative and communicative phenotypes are also clinically meaningful. Patients with similar pain severity may organize pain meaning differently—for example, linking pain with endurance and medication avoidance, family burden, or medical mistrust. These patterns may help explain differences in treatment engagement, delayed care, medication concerns, and communication barriers. Ultimately, this work extends QE into pain psychology, immigrant health, and clinical communication by modeling pain as a culturally situated meaning system with implications for more precise, equitable, and actionable pain assessment.

Honoring Richness in Blackness at Work: A Quantitative-Ethnographic Exploration of Psychological Safety, Joy, and Thriving Among Black Women in Tech

Shelbi Chandler, Ph.D.
Mercy University
schandler5@mercy.edu

Abstract. Black women in technology workplaces often navigate organizational cultures that demand assimilation while simultaneously marginalizing their cultural identity, leadership styles, and lived experiences [1]. While psychological safety has emerged as a critical construct for employee wellbeing and performance, dominant frameworks frequently fail to account for how cultural recognition and identity representation shape safety for historically excluded groups [3].

Keywords: Black women; psychological safety; identity representation; Black joy; resilience; work-place culture

1.1 Background of my research

Black women's workplace experiences are shaped by intersecting systems of race, gender, and power that often result in epistemic exclusion, professional image regulation, and chronic exposure to racialized microaggressions [4; 5; 6]. These conditions undermine psychological safety by signaling that authenticity carries professional risk. Du Bois' [3] concept of double consciousness remains salient, as Black women in corporate environments must continually navigate how they are seen through dominant lenses while attempting to preserve their full selves. Existing organizational research on psychological safety often treats identity as peripheral rather than central, limiting its applicability for Black women in predominantly white, male-dominated industries such as technology.

This study builds on Chandler's [1] dissertation, which argues that psychological safety for Black women is inseparable from identity representation and leadership accountability [8]. Figure 1 conceptualizes a culturally grounded pathway in which honoring the richness of Blackness, rather than merely tolerating difference, creates the conditions for safety, joy, and thriving. Drawing from Black feminist thought and Afrocentric frameworks, including Cooper's [2] articulation of Black joy and Smith's [7] Ubuntic Inclusion Framework, this research positions cultural affirmation as a protective and generative force. By examining how Black women interpret and experience identity recognition at work, this project responds to calls for scholarship that moves beyond deficit-based narratives toward models that center wholeness, resilience, and cultural wealth.

2 Methodology

This study employs a quantitative-ethnographic methodology, leveraging coded qualitative interview data from Black women working in global technology environments. Ethnographic codes related to identity affirmation, cultural recognition, psychological safety, joy, resilience, leadership behavior, and organizational climate are transformed into quantitative representations to examine patterns of meaning-making across participants. Data analysis will examine co-occurrence patterns and relational pathways aligned with Figure 1, with particular attention to how psychological safety functions as a mediating mechanism between honoring Blackness and experiences of joy and thriving. This methodological design supports culturally responsive inquiry while maintaining analytic rigor appropriate for interdisciplinary conference audiences.

This study will address the following RQ's: (1) How does honoring the richness of Blackness contribute to Black women's experiences of psychological safety in the workplace? (2) To what extent does psychological safety mediate the relationship between cultural recognition and Black joy or thriving? (3) How do leadership behaviors, peer interactions, and organizational culture influence the degree to which Black women feel their cultural identity is honored?

3 Expected Findings and Contributions

Expected findings include: (1) Honoring the richness of Blackness will be strongly associated with higher levels of psychological safety among Black women in tech; (2) Psychological safety will function as a key mediator between cultural recognition and experiences of Black joy and thriving; (3) Leadership accountability and peer affirmation will emerge as critical contextual factors shaping identity validation. (4) Quantitative ethnographic patterns will reveal how culturally affirming environments support resilience rather than mere endurance.

Expected contributions include: (1) Advance organizational scholarship by offering a culturally grounded model of psychological safety centered on Black women's experiences; (2) Demonstrate the value of quantitative ethnography for studying identity, culture, and meaning-making in workplace research; (3) Provide empirical support for leadership and organizational practices that honor identity as a pathway to wellbeing and performance; (4) Contribute to interdisciplinary conversations on Black joy, resilience, and thriving in professional environments.

Fig. 1. ***From Chandler's Dissertation***



References

1. Chandler, Shelbi, "More than a seat at the table: examining how psychological safety, identity representation, and leadership accountability impact African American women in global workplace environments" (2026). *Theses and Dissertations*. 1684. https://digitalcommons.pepperdine.edu/etd/1684
2. Cooper, B. (2021). *Eloquent rage: A Black feminist discovers her superpower.* St.Martin's Press.
3. Du Bois, W. E. B. (1903). *The souls of Black folk.* A. C. McClurg.
4. Holder, A. M. B., Jackson, M. A., & Ponterotto, J. G. (2015). Racial microaggression experiences and coping strategies of Black women in corporate leadership. *Qualitative Psychology, 2*(2), 164–180.
5. Roberts, L. M. (2005). Changing faces: Professional image construction in diverse organizational settings. *Academy of Management Review, 30*(4), 685–711.
6. Settles, I. H., Jones, M. K., Buchanan, N. T., & Dotson, K. (2019). Epistemic exclusion: Scholar(ly) devaluation that marginalizes faculty of color. *Journal of Diversity in Higher Education, 12*(1), 22–35.
7. Smith, J. G. (2021). *The Ubuntic Inclusion Framework.* Pepperdine University Press.
8. Wingfield, A. H. (2019). *Flatlining: Race, work, and health care in the new economy.* University of California Press.

Mapping “Responsibility” and “Deservingness” in State Truancy Laws: A Quantitative Ethnography Research Agenda

Ericka S. Weathers, PhD[10000-1111-2222-3333]

[1] University of Pennsylvania, Philadelphia PA 19104, USA
erickasw@upenn.edu

Abstract. State truancy and compulsory attendance laws shape how schools and districts respond to truancy (i.e., the accumulation of unexcused absences). Despite their central role in defining legal responses to truancy, relatively little research has examined how these policies construct responsibility and position students, families, and institutions as targets of intervention. This Research Agenda Development proposal explores how quantitative ethnography (QE) may be used to examine how concepts such as responsibility, support, punishment, discretion, and “deservingness” are relationally organized within state truancy policy discourse. Building on an ongoing national study, the proposed agenda seeks to identify productive QE approaches for studying policy meaning structures.

Keywords: Education Policy, Truancy and Compulsory Attendance Laws, Policy Text Analysis.

1 Area and Domain of Research

This proposed research study is in the domains of Social and Behavioral Sciences, Policy Studies, and Education Policy.

2 Background

Compulsory attendance and truancy laws represent one of the primary mechanisms through which states govern responses to legally defined truancy (and unexcused absences). These policies establish legal expectations for school attendance while defining the responsibilities of students, families, schools, courts, and other entities in responding to truancy [1]. Despite the central role of truancy laws in defining legal responses to unexcused absences, relatively little research has examined how these policies construct the causes of truancy, assign responsibility for addressing it, and position students, families, and institutions as targets of intervention. As a result, we know relatively little about how policy discourse organizes relationships among

responsibility, authority, support, punishment, and compliance. Such knowledge may prove useful in addressing racial disparities in truancy, and educational inequality more broadly.

This proposed research agenda is informed by my ongoing national study of state truancy laws grounded in attribution theory [2] and the social construction of target populations framework [3]. My work examines how policy language assigns responsibility for truancy and positions students, families, and institutions as targets of intervention. While this line of research provides insight into the presence and distribution of policy categories, it raises new questions about how those categories relate to one another within broader systems of meaning. For example, supportive interventions may be positioned as alternatives to punishment in some policies but embedded within escalating compliance structures in others. Similarly, responsibility may be attributed differently depending on how students, families, schools, and courts are connected to concepts such as support, risk, discretion, or punishment.

QE offers a promising framework for addressing these questions because attribution theory and social construction theory conceptualize meaning as fundamentally relational. Social constructions emerge not from isolated policy concepts but from relationships among actors, truancy causes, and policy responses. Likewise, attributions of responsibility are shaped by how policies connect students, families, schools, and institutions to explanations for truancy and to interventions or punishments. Rather than focusing solely on the presence of policy categories, QE provides tools for examining how concepts co-occur and form broader meaning structures. As such, a QE approach may offer new insights into how responsibility, deservingness, intervention, and authority are organized within state truancy laws. Understanding these meaning structures may provide new insights into how states conceptualize and govern truancy, informing future research on policy design, implementation, and equity.

3 Objectives

The aim of this proposed work is to explore how QE approaches can be used to examine the construction of responsibility, deservingness, and intervention within state truancy laws. Specifically, this project seeks to investigate how relationships among actors, truancy causes, and policy responses create broader systems of meaning that shape legal responses to truancy.

The proposed study focuses on state compulsory attendance and truancy laws in all fifty U.S. states. The primary unit of analysis is policy text, including statutory provisions governing unexcused absences, truancy identification, intervention processes, and legal enforcement mechanisms.

I propose the following research questions:

1. How are responsibility, support, punishment, discretion, and institutional authority relationally organized within state truancy laws?
2. How do state truancy laws construct students, families, schools, and courts through interconnected systems of meaning?

3. What patterns of responsibility, intervention, and authority emerge across state truancy laws?

This project is in the research agenda development stage. My ongoing national study has assembled a corpus of state compulsory attendance and truancy laws. The proposed work seeks to identify appropriate QE approaches for extending this work to explore policy meaning structures through a relational lens.

The proposed study would utilize the aforementioned corpus of state compulsory attendance and truancy laws from all fifty states. Existing coding informed by attribution theory and social construction theory may serve as a foundation. Potential units of analysis include policy provisions, statutory subsections, or other legally meaningful segments of text.

4 Prior Experiences and Expectations

My prior research has primarily employed quantitative methods to study education policy and educational inequality, although I am increasingly interested in qualitative and mixed methods approaches into my work. While attending ICQE 2025, I was introduced to a range of quantitative ethnographic approaches and became interested in their potential application to policy research. I am currently supported by a competitive early-career grant that includes explicit goals related to expanding my mixed methods expertise, and I see quantitative ethnography as a promising framework for advancing these objectives.

Although I have experience conducting policy text analysis and applying attribution theory and social construction theory to the study of educational governance, I have not previously conducted a QE study. Through participation in the Research Agenda Development session, I hope to receive guidance on adapting QE approaches to legal and policy texts, including questions related to units of analysis, coding structures, and relational modeling (especially). Beyond the proposed study, I am particularly interested in understanding how QE methods can support deeper investigation of meaning structures within state truancy laws and contribute to a broader mixed methods research agenda focused on educational policy and governance.

Acknowledgments. This study is funded by a Scholars Award from the William T. Grant Foundation (grant number SCH 205069).

Disclosure of Interests. The author has no competing interests to declare that are relevant to the content of this article.

References

1. Conry, J.M., Richards, M.P.: The severity of state truancy policies and chronic absenteeism. *Journal of Education for Students Placed at Risk (JESPAR)* **23**(1–2), 187–203 (2018)
2. Kelley, H.H.: Attribution theory in social psychology. In: *Nebraska Symposium on Motivation*. University of Nebraska Press (1967)

3. Schneider, A., Ingram, H.: Social construction of target populations: Implications for politics and policy. *American Political Science Review* **87**(2), 334–347 (1993)

Modeling Educators' Meaning-Making About Responsible AI Implementation: A Quantitative Ethnographic Comparison Across Educational Contexts and Professional Roles

Juhee Kim[1][0000-0002-8155-4376], Eunsung Park[2][0000-0002-2010-9957]

[1] University of Idaho, Moscow ID, USA juheekim@uidaho.edu
[2] Tennessee Tech University, Cookeville. TN, USA epark@tntech.edu

Abstract. Artificial intelligence (AI) is rapidly reshaping educational leadership and classroom practice, requiring educators to balance instructional opportunities with ethical, organizational, and professional considerations. Although prior research has examined educators' AI attitudes and readiness using surveys and qualitative analyses, less is known about how educators organize these ideas into coherent systems of meaning. Guided by Quantitative Ethnography (QE), this study applies Epistemic Network Analysis (ENA) to model relationships among five constructs of responsible AI implementation: AI readiness, instructional value, professional development, ethics and risk, and system support. Preliminary Model 0 analyses compare discourse across rural and urban educators, educational leaders and teachers, and novice and experienced educators. Rather than identifying isolated themes, this study examines how educator groups construct meaning around AI implementation through interconnected discourse patterns. Findings contribute to AI implementation research and the growing application of Quantitative Ethnography in educational leadership.

Keywords: Quantitative Ethnography; Epistemic Network Analysis; Artificial Intelligence; Educational Leadership; Meaning-Making

1 Area and Domain of Research

Learning Sciences • Quantitative Ethnography • Educational Leadership • Artificial Intelligence in Education

2 Background

Artificial intelligence (AI) is increasingly transforming teaching, learning, and educational leadership. Educational leaders and teachers must simultaneously consider instructional opportunities, educator readiness, ethical responsibility, professional learning, and organizational capacity when implementing AI in schools (Holmes et al., 2022; UNESCO, 2023). While existing studies have documented educators' perceptions, attitudes, and readiness toward AI adoption (Kim 2026; Kim & Wargo 2025), these approaches typically examine factors independently through surveys or thematic qualitative analyses. Consequently, relatively little is known about how educators connect these ideas as integrated systems of meaning during discussions of AI implementation.

Quantitative Ethnography (QE) offers a complementary methodological approach by integrating ethnographic interpretation with quantitative modeling to examine how people construct meaning through discourse (Shaffer, 2017). Rather than treating concepts as isolated variables, QE models relationships among ideas as they co-occur in conversations or written responses, allowing researchers to examine the epistemic structures underlying participants' reasoning (Shaffer et al., 2016). Epistemic Network Analysis (ENA), one of the primary analytical methods within QE, models and statistically compares these conceptual connections across groups.

This study extends previous work examining AI literacy among educators (Kim et al., 2026; Wargo et al., 2026) by moving beyond identifying important AI competencies toward understanding how educators organize knowledge about responsible AI implementation. Specifically, the study investigates relationships among five theoretically derived constructions: AI Readiness, AI Instructional Value, AI Professional Development Need, AI Ethics and Risk, AI System Support.

The project builds upon an exploratory Model 0 developed during the Quantitative Ethnography Fellows Summer Institute. Preliminary ENA analyses suggest that educators construct AI implementation differently depending on educational context and professional role. Rural educators more frequently connect AI readiness with instructional value, whereas urban educators demonstrate stronger relationships among professional development, ethics, and system support. Similarly, educational leaders emphasize organizational capacity and professional learning, while classroom teachers more strongly connect AI readiness with instructional practice. These exploratory findings motivate a larger QE investigation into educators' meaning making surrounding responsible AI implementation.

3 Research Objectives

The purpose of this study is to examine how educators construct meaning about responsible AI implementation by modeling relationships among AI implementation concepts across educational contexts and professional roles.

The study addresses the following research questions:

RQ1. How do rural and urban educators differ in the structure of epistemic connections among AI readiness, instructional value, professional development, ethics and risk, and system support?

RQ2. How do educational leaders and classroom teachers differ in their epistemic networks regarding responsible AI implementation?

RQ3. How do novice and experienced educators organize knowledge differently when discussing AI implementation?

Future analyses will examine whether these discourse patterns also differ across demographic characteristics, including gender and race/ethnicity, as additional data become available.

Qualitative responses collected from K–12 educators participating in AI-related professional learning will be analyzed using Epistemic Network Analysis. Responses have been coded using a five-construct codebook developed from AI implementation

literature and refined through iterative qualitative analysis. ENA will model co-occurrences among coded constructs and compare network structures across educator groups to identify similarities and differences in educators' meaning making (Shaffer et al., 2016). The current Model 0 serves as an exploratory analysis demonstrating the feasibility of applying Quantitative Ethnography to AI implementation research. Future work will expand the dataset, establish inter-rater reliability, refine segmentation decisions, and develop confirmatory ENA models suitable for publication (Shaffer, 2017).

4 Prior Experience and Expectations

This project represents my initial application of Quantitative Ethnography following participation in the Quantitative Ethnography Fellows Summer Institute. Developing the Model 0 demonstrated how QE extends traditional qualitative analysis by modeling relationships among ideas rather than identifying themes independently (Shaffer, 2017). Through continued collaboration with the QE community, I hope to strengthen methodological decisions related to data segmentation, coding reliability, model construction, and interpretation of epistemic networks.

More broadly, I seek to integrate Quantitative Ethnography into my mixed-methods research agenda examining educational leadership, responsible AI implementation, and rural education. By combining qualitative interpretation with quantitative modeling, I hope to contribute methodological advances that help researchers better understand how educators make sense of emerging technologies and how these meaning-making processes inform educational policy, leadership preparation, and professional learning.

References

1. Holmes, W., Bialik, M., & Fadel, C. (2022). *Artificial intelligence in education: Promises and implications for teaching and learning*. Center for Curriculum Redesign.
2. Kim, J. (2026). School leaders and AI-driven education: A comparative study of readiness, perceptions, and implementation strategies. *Artificial Intelligence in Education*. *2*(3), 1-17. https://doi.org/10.1108/AIIE-03-2025-0048
3. Kim, J., Cheah, Y. H., & Wargo, E. (2026). Transforming K-12 Education With AI: Educational Leaders' Perspectives and Policy Implications. *Journal of School Leadership*, *36*(3), 396-415. https://doi.org/10.1177/10526846251413036
4. Kim, J. & Wargo, E. (2025). Empowering Educational Leaders for AI Integration in Rural STEM Education: Challenges and Strategies. *Frontiers in Education*. 10. https://doi.org/10.3389/feduc.2025.1567698
5. Shaffer, D. W. (2017). *Quantitative ethnography*. Cathcart Press.
6. Shaffer, D. W., Collier, W., & Ruis, A. R. (2016). A tutorial on epistemic network analysis: Analyzing the structure of connections in cognitive, social, and interaction data. *Journal of Learning Analytics, 3*(3), 9–45.
7. UNESCO. (2023). *Guidance for generative AI in education and research*. https://unesdoc.unesco.org/ark:/48223/pf0000386693

8. Wargo, E., Kim, J., & Simmons, J. (2026). Considering AI Adoption and Use: Promises and Paradox for Rural Education. *The Rural Educator, 47*(2), 13-24. https://doi.org/10.55533/2643-9662.1470

Re-Examining Globalization: Definition and Biases

Letitia Lee[1][0009-0002-0436-9449] and Andrea S. Matos Devesa[1][0000-0003-0344-833X]

[1] The Chinese University of Hong Kong, Hong Kong SAR China
letitialee@cuhk.edu.hk

Abstract. *Globalization* is used as a catch-all term to introduce rising phenomena, but is rarely formally defined, with critics pointing out that it disproportionately favors a Western-centric lens. Our proposed research aims to examine how scholars and laypeople conceptualize the term to uncover potential biases in knowledge production and perpetuation. We would like to utilize QE techniques to analyze, compare, and visualize the different conceptualizations and their commonalities, ultimately deriving a common understanding of the term and its correlates.

Keywords: globalization, scoping review, cross-cultural research.

1 Area and Domain of Research

Social and Political Sciences, Globalization / Decolonization, Cross-cultural Research.

2 Background

In the social sciences, *globalization* is used as a catch-all term and phenomenon to explain the ongoing societal, cultural, and political impacts of decades-long economic policies. As a result, both left and right political and media entities have attempted to address the alleged consequences of globalization by pushing policies focusing on isolationism, such as anti-immigration laws [1]. Despite its ubiquitousness, there is a lack of consensus on the definition and operationalization of the term [2, 3]. Decolonial scholars have pointed out the disproportionate distribution of knowledge production coming from "more privileged societal contexts" [2], and criticized the usage of globalization as interchangeable with "westernization," "liberalization," and "universalization," noting that definitions reflect and reproduce powers and norms represented by the term [3]. Nevertheless, "globalization" continues to inform scholarship and policy, perpetuating the assumptions linked with the term. Given that definition of terms is relative to their context, "globalization" may be understood differently across sociocultural faultlines [3]. Specifically, we argue that scholars and laypeople from the global North and South may utilize the term differently. For instance, since globalization is often used as a process in discussions of economic policies exacerbating unequal and exploitative labor production between the North and South [4], scholars and laypeople from the global South may perceive and respond differently to the term. Yet, as far as

we know, there have been few attempts to systematically and empirically study how globalization is defined within scholarly work, and to examine any cross-cultural differences or biases in how such a pervasive term is used.

3 Research Objectives

The goal of our proposed research is to understand what different groups mean when referring to *globalization*. We aim to untangle the ways in which the term is defined and used across the social sciences: What is the consensus as to what globalization is, and how/why has the conceptualization of globalization changed over time? We are also interested in whether/what biases exist across conceptualizations. In turn, how do these potential theoretical differences impact the way scholars and laypeople perceive and discuss globalization, its facets, antecedents, and consequences? We are proposing two studies to address these questions. Study 1, a scoping review, will examine papers that focus on conceptualizing globalization. We plan to code the main text of key papers, and conduct network analyses using both the qualitative coding and the quantitative metadata from the journal sources. We have cleaned the data, and are preparing to begin coding. Study 2 will involve cross-national survey data collection with a mixed design. Laypeople will be asked to discuss their understanding of globalization, and we will gather additional attitudinal information to serve as potential moderators, alongside culture, in network visualizations. Study 2 is still in the conceptualization phase.

4 Prior Experience and Expectations

Although we do have some experience with QE and ENA (through prior conferences and workshops), we are hoping to gain more formal insight on how to code and interpret qualitative data and interpret visual networks. We would like guidance on how to work with cross-cultural data sources, and also how to fairly draw comparisons when working with such sources. We have heard about some techniques like association rules mining and ordered network analysis, and feel we may benefit from using these techniques, but would like advice on how best to implement them, in order to better articulate and visualize the associations or potential causal relationships in people's conceptualization of globalization.

5 References

1. Uygur, M. N., Bakan, S.: Agitation strategies of anti-immigration: The populist and neoliberal lens. Journal of International Migration and Integration, 26, 813-840 (2025).
2. Robertson, R., White, K. E.: What is globalization? In G. Ritzer (ed.), The Blackwell Companion to Globalization (pp. 54-66). Wiley. (2007).
3. Scholte, J. A.: Defining globalization. World Economy, 31, 1471-1502 (2008).
4. Hickel, J., Hanbury Lemos, M., Barbour, F.: Unequal exchange of labour in the world economy. Nature Communications, 15(6298), 1-12 (2024).

The Impact of Rapid Prototyping in the AI Era on Organizational Work Validation,

and the Application of Quantitative Ethnography in Service Design Practice.

First Contributor Yuya Kashiwa, ADDIX Co., Ltd.

Abstract. This study examines the impact of rapid prototyping in the AI era on service design processes and organizations (e.g., increased validation costs).
We will conduct in-depth interviews with users and developers within the company and use epistemological network analysis (ENA) and other methods to code and analyze the resulting data. This will allow us to visualize changes in the epistemological worldviews of community members and translate these findings into organizational and methodological knowledge.

Keywords: Prototyping, Service Design Network Analysis .

1 Area and Domain of Research

Design Research,Business and Organizational Studies.

2 Background

With the widespread adoption of AI, the barriers to creating in-house tools and prototypes have dropped significantly. As a result, it has become possible to routinely conduct a volume of prototyping and validation that would previously have been impossible due to time and cost constraints.
On the other hand, as the range of things that can be created has expanded, organizations are likely to face the following challenges:

1. The number of tools and prototype proposals increases, leading to an overload of items requiring validation.
2. The workload associated with providing feedback and evaluations increases.
3. It becomes more difficult to determine what to adopt and what to learn from.
4. The cost of learning rises, and in service design, the importance of “assessing” grows even more than that of “creating.”

In the past, because prototyping itself was costly, learning took place within a limited number of opportunities. However, now that mass production of prototypes has become possible, while the pace of failure and learning continues to accelerate, the burden on organizations to absorb, organize, and make decisions regarding this information may be increasing.
This paper clarifies the organizational impact of rapid prototyping in the AI era.
2

3 Research Objectives

A description of the aims of the proposed work. Answer the following key questions as they apply:

1. What are the aims of the proposed research?

   We will conduct interviews with service professionals and, through analysis using QE/ENA, examine “the impact of AI on the service design process,” and feed these findings back to the company as organizational and methodological knowledge. In addition, by offering QE/ENA workshops, we will enable participants to acquire useful research and analysis methods and apply them to their daily work.

2. To what population does this research pertain?

   We are currently conducting in-depth interviews. We plan to conduct additional interviews in the future.

3. What research question(s) does this research propose to answer?

   How does AI-powered prototyping affect the epistemological frameworks of those involved in service design?

   Furthermore, how will organizations change, and how should these changes be managed?

4. In what stage is your study?

   We are currently conducting in-depth interviews. We plan to conduct additional interviews in the future.

5. What type(s) of data will be collected?

   Qualitative research data, such as in-depth interviews

4 Prior Experience and Expectations

I would appreciate learning how to develop a research plan and gain insights into “which methods to combine and how to combine them” to address the research question.

References

1. Shaffer, D. W. (2017). Quantitative Ethnography. Morgan & Claypool Publishers.

A Conceptual Design of Researcher-Note-Mediated Coding Support in Quantitative Ethnography

Hideki Kondo[1][0009-0001-8775-3712], Sayaka Tohyama[2][0000-0001-5759-9550] and Masayuki Yamada[3][0000-0002-0308-8795]

[1] Kanda University of International Studies, 1-4-1 Wakaba, Mihama, Chiba, 261-0014, Japan
[2] Shizuoka University, 3-5-1 Johoku, Chuo-ku, Hamamatsu, Shizuoka, 432-8011, Japan
[3] Kyushu Institute of Technology, 680-4 Kawazu, Iizuka, Fukuoka, 820-8502, Japan

kondo-h@kanda.kuis.ac.jp

Abstract. The objective of this study is to design an initial workflow for researcher-note-mediated codebook revision in large language model (LLM)-assisted coding for Quantitative Ethnography. This study focuses on notes written while researchers interpret LLM outputs and Epistemic Network Analysis (ENA)-based results, and explores how these notes can support researcher-led codebook revision and iterative analysis.

Keywords: Qualitative Coding, Large Language Models, Epistemic Network Analysis, Researcher Notes.

1 Area and Domain of Research

This study is situated in the fields of Learning Sciences and Educational Technology. It focuses on written and interactional data, such as electronic bulletin board posts and interview data, collected from informal learning environments including online reading groups and self-organized study seminars.

2 Background

In qualitative data analysis, researchers read the data, define codes and coding criteria, and apply those codes to the data. They then revise the definitions, criteria, and interpretations of the data while visualizing and analyzing the results using methods such as Epistemic Network Analysis (ENA). When the definitions and boundaries of a codebook change, the results of the analysis may also change. Coding is therefore not merely a preprocessing step, but an important process through which researchers determine what they conceptualize in the data [1].

LLM-based support for qualitative coding has been explored in Quantitative Ethnography, including the use of ChatGPT to support automated coding and the refinement of human coding schemes [2]. However, there remains room for further investigation into how LLM-based support changes researchers' interpretations, how such

changes can be documented, and how those records can be used to support further analysis.

3 Research Objectives

The objective of this study is to design an initial workflow in which researcher notes mediate codebook revision in LLM-assisted coding for Quantitative Ethnography. In this workflow, a researcher describes coding criteria in natural language, an LLM assigns codes to the target interactional data, and the coded data are modeled and visualized using ENA. The researcher interprets the LLM outputs and ENA-based results, writes notes on discrepancies, unexpected patterns, and code definitions that need reconsideration, and uses these notes with LLM support to revise the codebook before the next cycle of coding and analysis.

The current research questions are as follows.

RQ1: How do researchers use LLM-coded data and ENA-based results to identify needs for codebook revision?

RQ2: How can an LLM use researcher notes to support researcher-led codebook revision?

This research is currently at the conceptual stage. Using a small-scale interactional data as we mentioned above, we will develop a prototype that connects codebook editing, LLM coding, ENA-based results, researcher notes, and version control.

4 Prior Experience and Expectations

Our previous research has focused on supporting online activities in informal collaborative learning communities. In our research, we have designed and evaluated a system in which an LLM generates scripts that prompt learners' next activities based on principles derived from the Learning Sciences. Although quantitative analyses in our study have suggested promising results, there are limitations to those analyses. We are interested in conducting more in-depth evaluations through Quantitative Ethnography and in exploring mechanisms for supporting learners' activities.

For the first author, this is the first study that explicitly incorporates Quantitative Ethnography as a methodology. In the RAD session, we would like to discuss a method for supporting researcher-led interpretive loops rather than automated coding accuracy.

References

1. Shaffer, D. W., & Ruis, A. R. (2021). How we code. In A. R. Ruis & S. B. Lee (Eds.), Advances in Quantitative Ethnography (pp. 62–77). Springer. https://doi.org/10.1007/978-3-030-67788-6_5
2. Zambrano, A. F., Liu, X., Barany, A., Baker, R. S., Kim, J., & Nasiar, N. (2023). From nCoder to ChatGPT: From automated coding to refining human coding. In G. A. Irgens & S. Knight (Eds.), Advances in Quantitative Ethnography: ICQE 2023 (pp. 470–485). Springer. https://doi.org/10.1007/978-3-031-47014-1_32

An AI-Human Collaborative Content Analysis of University Functions in Chinese National Strategies

Xin Li[1*] and Kaixiang Kang[1,2]

[1] Hiroshima University, Hiroshima, Japan
[2] Prefectural University of Hiroshima, Hiroshima, Japan
[*] xinli@hiroshima-u.ac.jp

Abstract. Policy documents describe universities and assign them institutional roles. This RAD presents an AI-human collaborative content analysis of university functions in Chinese national strategies. It places the China pilot in a broader source inventory of 239 Chinese, 101 Japanese, and 113 Korean policy documents, while the reported analysis is confined to screened Chinese material. Chinese paragraph-level segments were organized into four analytical policy types and given stable identifiers. Researchers defined six function domains. Local BERTopic produced a preliminary topic structure; the two authors checked source context, revised the codebook, and made the final coding decisions. Cross-ministerial strategies had their largest coded shares in technology transfer and entrepreneurship and in global cooperation. Education five-year plans had their largest shares in talent cultivation and in ideology and culture. The results describe differences among Chinese policy categories. The workflow links BERTopic analysis to source passages and human judgment, providing a basis for later comparison once sampling and coding procedures are harmonized.

Keywords: BERTopic · content analysis · policy corpus · quantitative ethnography · university functions

1 Purpose, Corpus, and Method

Policy discourse assigns universities roles in innovation, talent, service, culture, and international cooperation [1, 4]. This paper asks: (1) How do coded university functions vary across Chinese policy types? (2) How can BERTopic-assisted analysis remain traceable to source text and human decisions?

The broader project has compiled 453 policy documents, although the present analysis is restricted to China. China has 239 documents (22 State Council or CCP Central Committee comprehensive plans, 105 sectoral strategies, and 112 specific higher-education policies or programs); Japan has 101 (26 cabinet-level economic or development guidelines, 33 sectoral strategies or basic plans, and 42 MEXT higher-education policies or programs); and Korea has 113 (17 presidential or supra-ministerial policies, 19 single- or inter-ministerial sectoral

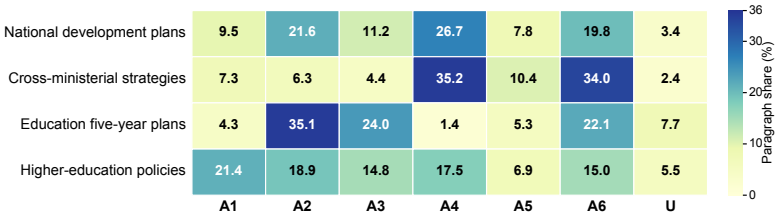


Fig. 1. China pilot: row-normalized paragraph-level shares by policy type. A1: science/technology; A2: talent; A3: ideology/culture; A4: transfer/entrepreneurship; A5: service/think-tank; A6: global cooperation; U: unclassified.

strategies, and 77 ministerial higher-education policies or programs). Sources were PKU Law Archives and Chinese ministries' sites; the Library of Congress and Japanese Cabinet and MEXT sites; and Korea's National Archives and eGov Policy Briefing. These document counts are not heat-map denominators.

The China pilot uses four analytical policy types: national plans, cross-ministerial strategies, education five-year plans, and higher-education policies. After cleaning, segmentation, and screening, each retained segment received a stable identifier and source metadata. BERTopic was run locally for topic modeling [2]. Both authors, who are native Chinese speakers, reviewed its outputs against the source text, revised the codebook, and made the final coding decisions.

The workflow combines content analysis [3, 5] with QE's text–representation–text loop [6, 7]. Figure 1 reports Chinese shares, with each row totaling approximately 100%, including U. The three inventory strata and four analytical policy types are separate classifications. Codes describe assigned roles in sampled texts, not university behavior or policy effects.

2 Preliminary Findings and Contribution

In the China pilot, cross-ministerial strategies have their largest shares in A4 (35.2%) and A6 (34.0%). Education five-year plans have their largest shares in A2 (35.1%) and A3 (24.0%), while A4 accounts for 1.4%. No A1–A6 domain exceeds 21.4% among higher-education policies. These are descriptive differences among Chinese policy categories, not cross-national or causal effects (Fig. 1).

Stable identifiers link aggregate patterns to source passages, and the authors make the final decisions. Further work will refine the validation and semantic-mapping procedures.

References

1. Compagnucci, L., Spigarelli, F.: The third mission of the university: A systematic literature review on potentials and constraints. Technological Forecasting and Social Change **161**, 120284 (2020). https://doi.org/10.1016/j.techfore.2020.120284
2. Grootendorst, M.: BERTopic: Neural topic modeling with a class-based TF-IDF procedure. arXiv preprint arXiv:2203.05794 (2022). https://doi.org/10.48550/arXiv.2203.05794
3. Krippendorff, K.: Content Analysis: An Introduction to Its Methodology. SAGE, Thousand Oaks, 4 edn. (2018)
4. Marginson, S.: Higher education in East Asia and Singapore: Rise of the Confucian model. Higher Education **61**(5), 587–611 (2011). https://doi.org/10.1007/s10734-010-9384-9
5. Schreier, M.: Qualitative Content Analysis in Practice. SAGE, London (2012)
6. Shaffer, D.W.: Quantitative Ethnography. Cathcart Press, Madison (2017)
7. Shaffer, D.W., Ruis, A.R.: How we code. In: Ruis, A.R., Lee, S.B. (eds.) Advances in Quantitative Ethnography. Communications in Computer and Information Science, vol. 1312, pp. 62–77. Springer, Cham (2021). https://doi.org/10.1007/978-3-030-67788-6_5

Connecting Physics and Mathematics Reasoning Across Three Inventories: A Quantitative Ethnographic Approach

Kazi Aatish Imroz[1 [0009−0009−6339−8216]] and Geraldine L. Cochran[1 [0000−0002−8520−6168]]

The Ohio State University, Columbus OH 43210, USA
imroz.1@buckeyemail.osu.edu

Abstract. Students entering transformed introductory physics pathways arrive with varying levels of mathematics preparation, and their conceptual and quantitative reasoning develops in ways that aggregate scores do not capture well. We collected pre- and post-instruction data from students in a transformed introductory physics course using three valid and reliable inventories: the Physics Inventory of Quantitative Literacy [9], the Test of Understanding Graphs in Kinematics [1], and the Gender Force Concept Inventory [4, 3, 5]. Although each instrument is validated to measure a distinct construct, these constructs are deeply connected in the discipline: interpreting kinematics graphs draws on quantitative literacy, and reasoning about force in relation to velocity and acceleration is itself a form of physics quantitative reasoning. We would therefore expect students who are strong on one instrument to be strong on the others, yet anecdotal analysis suggests that total scores do not track one another, and there is little consistency in which items students answer incorrectly. We propose to use Quantitative Ethnography to examine, at the item level, the connections between the kinds of reasoning students demonstrate correctly and incorrectly within and across these instruments, and how those connections change over a semester.

Keywords: Mathematics · Quantitative Reasoning · Introductory Physics.

1 Area and Domain of Research

This research is situated in Physics Education Research, with a focus on quantitative literacy and quantitative reasoning, using pre- and post-instruction assessment data.

2 Background

Our study is part of the national initiative Transforming Introductory Physics Sequences to Support All Students (TIPSSS) [7]. Through this network, institutions across the United States are developing and examining alternative introductory physics pathways designed to support students who enter college with

varying levels of mathematics preparation. Based on prior math preparation, students interested in engineering are often advised to delay enrollment in traditional introductory physics courses until they complete remedial or prerequisite mathematics courses, which can increase their time to degree and create additional barriers to entering engineering programs [8]. The transformed courses provide alternative pathways in which students earn academic credit while developing the mathematical reasoning needed to understand physics, with a central emphasis on quantitative literacy: the ability to use mathematical ideas and representations meaningfully in physical contexts.

The quantitative literacy framework developed by White Brahmia and colleagues identifies three major facets that commonly challenge students in introductory physics: reasoning with signed and unsigned quantities, covariational reasoning, and ratio and proportional reasoning [9]. The Physics Inventory of Quantitative Literacy (PIQL) was developed to assess students' understanding across these areas [9]. Like the PIQL, the Gender Force Concept Inventory (G-FCI) [5, 4, 3] and the Test of Understanding Graphs in Kinematics (TUG-K) [1] are commonly used, valid and reliable inventories for the introductory physics classroom, available to instructors through PhysPort [6]. The G-FCI is a version of the Force Concept Inventory [2], the widely used inventory of conceptual understanding of Newtonian force, with items reformatted to examine gender differences in response patterns.

3 Research Objectives

The proposed research examines the connections between the kinds of reasoning students demonstrate correctly and incorrectly, within and across the three instruments, and how those connections change from pre- to post-instruction. Because these connections exist below the level of the total score, we pursue this question at the item level. Our guiding research questions are:

- RQ1: What connections exist among the kinds of reasoning students demonstrate correctly and incorrectly across the PIQL, G-FCI, and TUG-K at the item level?
- RQ2: How do these connections change from pre- to post-instruction over one semester?

4 Prior Experience and Expectations

Aatish: This will be my first study using Quantitative Ethnography (QE). However, I have some familiarity with the field through reading papers published in the International Conference on Quantitative Ethnography proceedings over the past several years. I hope to gain a clearer understanding of how QE can be applied to pre- and post-assessment analysis. I look forward to thinking through the structure and development of a model 0 for this study. **Geraldine:** I am a physics education researcher and new to QE.

References

1. Beichner, R.J.: Testing student interpretation of kinematics graphs. American journal of Physics **62**(8), 750–762 (1994). https://doi.org/10.1119/1.17449
2. Hestenes, D., Wells, M., Swackhamer, G.: Force concept inventory. The physics teacher **30**(3), 141–158 (1992). https://doi.org/10.1119/1.2343497
3. McCullough, L.: Gender, math, and the FCI. In: Proceedings of the 2002 Physics Education Research Conference. Rochester NY (2002). https://doi.org/10.1119/perc.2002.pr.013
4. McCullough, L.: Gender differences in student responses to physics conceptual questions based on question context. ASQ Advancing the STEM Agenda in Education (2011)
5. McCullough, L., Meltzer, D.E., Semak, M., Willis, C.: Differences in male/female response patterns on alternative-format versions of fci items. In: Proceedings of the Physics Education Research Conference. pp. 103–106 (2001). https://doi.org/10.1119/perc.2001.pr.013
6. McKagan, S.B., Strubbe, L.E., Barbato, L.J., Mason, B.A., Madsen, A.M., Sayre, E.C.: Physport use and growth: Supporting physics teaching with research-based resources since 2011. The Physics Teacher **58**(7), 465–469 (2020). https://doi.org/10.1119/10.0002062
7. Transforming Introductory Physics Sequences to Support all Students: TIPSSS Network Website. https://u.osu.edu/tipsss/
8. White Brahmia, S., Cochran, G.L.: Reframing the narrative on physics readiness. Physics Today **78**(10), 40–47 (2025), https://doi.org/10.1063/pt.qflo.wdfw
9. White Brahmia, S., Olsho, A., Smith, T.I., Boudreaux, A., Eaton, P., Zimmerman, C.: Physics inventory of quantitative literacy: A tool for assessing mathematical reasoning in introductory physics. Physical Review Physics Education Research **17**(2), 020129 (2021). https://doi.org/10.1103/PhysRevPhysEducRes.17.020129

Critical In-Depth Reading for Immigrant and Refugee Background Multilingual Learners

Hazel Vega[1][0000-0001-9077-3522] Jennifer C. Mann[1][0000-0002-8754-5059] and Alison M. Turner[2][0000-0003-4287-6695]

[1] Duke University, North Carolina, USA
[2] North Carolina State University, North Carolina, USA
hazel.vegaquesada@duke.edu

Abstract. This research explores how critical and in-depth reading can impact student academic performance and wellbeing in secondary English as a Second Language classrooms. Using a mixed-methods, quasi-experimental design, this study examines shifts in teacher instructional practices and student outcomes. We expect to learn how to integrate QE at an early stage of the project.

Keywords: Immigrant and refugee students, Multilingual learners, Critical literacy, Teacher learning, Quasi-experiment.

1 Area and Domain of Research

This research is in the domain of teacher education and multicultural education.

2 Background

The current geopolitical circumstances have mobilized large populations, forcing human displacement and relocation. As such, in the U.S., 1 in 4 children lives in an immigrant family, and multilingual learners—students learning English who speak other languages—are the fastest growing K-12 student population [1]. Although immigrant/refugee and multilingual learners are not new in U.S. schools, systemic inequities have restricted their academic achievement. Ubiquitous deficit-oriented views towards them prioritize surface-level literacy skills at the expense of cognitively demanding practices [2]. Immigrant/refugee background students are seen as requiring development of basic skills as a pre-requisite for higher intellectually challenging content.

Critical literacy disrupts these deficit views by broadening access to higher order thinking skills and textual analysis that uncovers power dynamics and sociopolitical issues affecting immigrant/refugee background students [2]. Critical literacy involves exploring texts through critique of values, interests, ideologies to analyze underlying messages regarding social justice [3]. With the extensive misinformation and disinformation and accessibility to online platforms and social media, critical literacy

has become essential for young people. For immigrant/refugee background students, critical literacy can be transformational as they navigating a new country, where they are often the target of stereotypes and inequities [4].

3 Research Objectives

This research explores how Reading for In-Depth Learning –an approach centering text engagement, critical thinking, and real-world language use [3] can impact academic performance and wellbeing in secondary English as a Second Language (ESL) classrooms. Our project aims to equip secondary ESL teachers with evidence-based tools to support students from immigrant/refugee backgrounds through professional development. This five-year project will work with 60 teachers and approximately 600 students over two cohorts in three U.S. school districts. This research addresses these questions: (1) How do secondary ESL teachers use reading for in-depth learning during instruction? (2) To what extent does reading for in-depth learning impact academic outcomes and wellbeing for secondary ESL learners? This research uses a mixed-methods, quasi-experimental design to examine how participation in professional learning influences teacher beliefs, instructional practices, and student outcomes, with attention to identity, belonging, and wellbeing. Our data collection includes teacher and student surveys and artifacts, classroom observations, and teacher interviews.

4 Prior Experience and Expectations

The first author has extensive experience with QE, and for authors 2 and 3, this is the first approach to the method. We want to explore how QE can measure, describe, and visualize growth of student criticality and wellbeing and teacher implementation of critical literacy across data sources. We are developing the data collection and analysis plan, and we expect to learn how to integrate QE at this stage and consider options for data infrastructure that both align with QE principles and serve our research goals.

References

1. National Center for Education Statistics (NCES): Report on the condition of education 2022. (2022)
2. Alford, J.H.: Well, hang on, they’re actually much better than that!’: Disrupting dominant discourses of deficit about English language learners in senior high school English. English Teaching. 13, 71–88 (2014)
3. Lau, S.M.C., Ibrahim, N.C., Destroismaisons, F.: Promoting reading for in-depth learning for critical literacy and interculturality. Children’s Literature in English Language Education. 13, 1–29 (2025)

4. Lee, C.C., Picart, J., Mann, J.C.: Amplifying youth voices through critical literacy and positive youth development: The potential of university-community partnerships. Routledge (2025)

Defining Pharmaceutical Data Science: A Competence Model as an Emerging Artifact

Stine Marie Høgh Jensen[0009-0004-4605-9080], Adrienne L. Traxler[0000-0003-2725-0686], Hendra Agustian[0000-0003-2792-3035], Frederik Voetmann Christiansen[0000-0002-6743-9881], Lykke Pedersen[0000-0002-8935-415X], Arnault-Quentin Vermillet[0009-0003-1227-4199], and Morten Misfeldt[0000-0002-6481-4121]

University of Copenhagen, Copenhagen, Denmark
stine.m.jensen@ind.ku.dk, atraxler@ind.ku.dk

Abstract. In recent years, the pharmaceutical sciences are increasingly dominated by data and digital workflows, and the advent of AI tools in areas such as drug design has only accelerated this change. Future pharma professionals are increasingly called on to work with statistics, complex models, and large datasets. However, the degree programs for these fields have not kept pace, and now data science capabilities must compete with an already full curriculum. The Center for Pharmaceutical Data Science Education represents a coordinated effort to tackle these challenges by weaving the necessary training in data science into all levels of the study programs in Denmark. Here we invite discussion of how to frame quantitative ethnography on the emerging framework of pharmaceutical data science education (PDSE) with a focus on one of the Center's core documents, the prototype PDSE competence model. A collective effort of approximately twenty Center-affiliated teachers and personnel, it represents not only a catalogue of technical skills but also an ongoing conversation about the identity and priorities of an emerging sub-field.

Keywords: data science education, pharmacy education, competence development.

1 Pharmacy education in Denmark

In Denmark, pharmacy education is offered at the bachelor and master level, via 3-year BSc and 2-year MSc programs at two universities. At the University of Copenhagen (UCPH), these programs are housed in the Department of Pharmacy, with additional research projects in the Department of Drug Design and Pharmacology. At the University of Southern Denmark (SDU), pharmacy programs are in the Department of Physics, Chemistry, and Pharmacy or in the Department of Public Health. Students will have careers in research, industry, and clinical practice that range from basic research in drug development through patient-care specialists. Specialization typically occurs at the masters' level, with future paths stemming from a common BSc education in pharmacy.

2 Background and prototype competence model

The Center for Pharmaceutical Data Science Education is a collaboration between UCPH and SDU, across the pharma program departments and including the Departments of Computer Science and Science Education. Course-level efforts to incorporate data science skills were already taking place at both universities, but the Center is intended to drive a larger-scale and more unified revision of the curriculum—not displacing pharmacy content, but updating and rethinking the ways that it is taught. An early key effort toward this goal, taking place in spring through fall 2025, was the development of a common pharmaceutical data science education (PDSE) competence model to guide curriculum adjustments.

Development of the model began as a conversation among early Center personnel about learning goals and data science competency areas (e.g. Statistics, Good Software Practice), with an idea of mapping courses in the program to specific numeric competency levels. These conversations evolved to include a growing number of teachers in the degree programs, and this larger group divided into teams to articulate the provisional categories. Through several cycles of whole-group meetings and team refinement, a final set of topics emerged. This work was synthesized primarily by S. Jensen as the November 2025 version of the competence model.

As PDS education is a new field, there is no existing consensus model for the dimensions [1], but frameworks from data science education provided inspiration [2–4]. The top-level categories in the model are: Ethics, Legislation & Privacy; Mathematics & Statistics; Computing & Programming; Data Acquisition & Management; Exploration, Mining & Analysis; Machine Learning & AI; and Visualization & Presentation.

3 Questions to Discuss at ICQE

Detailed qualitative data was not collected as part of the competence model's development, but its construction illustrated opportunities and challenges inherent in the Center's size, disciplinary range, and blended education/research mission. We would welcome input from the ICQE community on analyzing the existing competence model and possibly collecting new data, with potential focus areas including:

- Cultural negotiation between science education researchers (slower, theory-based) and Center teachers (need for immediate, responsive tools)
- Disciplinary diversity in pharma, from basic-research-oriented drug design projects to issues of production and ethical patient care
- Other culture negotiations: industry and academic Center personnel, data science enthusiasts and relative novices (i.e., many of the students and some of the teachers)

Acknowledgments. This study was funded by the Novo Nordisk Foundation (grant number NNF24SA0092613), the LEO Foundation (grant number LF-ST-24-500013), and the Lundbeck Foundation (grant number R469-2024-629).

Disclosure of Interests. The authors have no competing interests to declare.

References

1. Agustian, H., Jensen, S.M., Traxler, A.L., Hagood, D.: Data science education in pharmacy(-related) programs: Scoping review of student learning and curricular approaches. (under review).
2. Ridsdale, C., Rothwell, J., Smit, M., Bliemel, M., Irvine, D., Kelley, D., Matwin, S., Wuetherick, B., Ali-Hassan, H.: Strategies and Best Practices for Data Literacy Education Knowledge Synthesis Report. (2015). https://doi.org/10.13140/RG.2.1.1922.5044.
3. Echtenbruck, M.M., Fühles-Ubach, S., Naujoks, B., Kaliva, E.: A Data Literacy Competence Model for Higher Education and Research, http://arxiv.org/abs/2504.15690, (2025). https://doi.org/10.48550/arXiv.2504.15690.
4. ACM Data Science Task Force: Computing competencies for undergraduate data science curricula. Association for Computing Machinery, New York, NY, USA (2021).

EFL Students' Agentic Engagement in AI-assisted Writing Revision: A Transmodal Analysis

Yangyang Yu[1][0009-0002-7346-639X]

[1] Shanghai Jiao Tong University, Shanghai, China
florayu0209@sjtu.edu.cn

Abstract.
This study aims to explore how English as a foreign language (EFL) students exhibit agentic engagement in AI-assisted writing revision. It is in the conceptualization and design phase, intending to use a collected dataset of screen recordings and stimulated recalls. I would like to perform QE techniques, particularly Transmodal Ordered Network Analysis (T/ONA), on the dataset to reveal students' agentic engagement patterns over time. The study may deepen the understanding of students' agency in AI-assisted revision and inform the integration of AI into EFL writing instruction.

Keywords: Generative AI, Automated feedback, Second language writing, Agentic engagement, Transmodal analysis

1 Area and Domain of Research

The proposed research is situated within the area of learning sciences and educational technology. It specifically belongs to the domain of computer-assisted language learning and second language (L2) writing.

2 Background

Generative AI technology has brought both chances and challenges to L2 writing instruction. Researchers have pointed to the potential of generative AI in facilitating the entire writing process while showing ethical concerns about overreliance on tools and loss of authorial voice [1]. For the revision stage in particular, generative AI tools such as ChatGPT can provide abundant AI feedback, alleviating teachers' workload and increasing students' L2 learning opportunities. Existing empirical studies on such AI feedback have examined its quality [2], impact on students' writing performance [3], and students' revision behaviors and perceptions [4]. However, it remains underexplored whether students exert their agency during the revision process and whether such agency evolves over time. Tracing students' agentic engagement in AI-assisted revision is crucial for designing effective pedagogical approaches that encourage students' learning from AI rather than reliance on AI. Meanwhile, understanding students' agency in authentic revision contexts may inspire tool

developers to create learning-oriented, pedagogically friendly feedback tools powered by generative AI.

The present study is proposed to explore EFL students' agentic engagement in AI-assisted writing revision. Conceptually, it adapts the machine-in-the-loop writing framework [5] and centers on the construct of agentic engagement [6]. The machine-in-the-loop framework posits that a writer engages in an iterative loop of prompting AI, evaluating AI-generated content, and making use of AI-generated content. Agentic engagement, referring to learners' actively shaping and co-constructing their learning experiences, represents how the writer exercises control within the loop. In AI-assisted revision, agentic engagement can be operationalized through the proactive and intentional level of prompting, evaluation, and integrating AI feedback. Methodologically, the study intends to employ transmodal analysis for QE [7]. Transmodal Ordered Network Analysis (T/ONA) [8], for example, can map asymmetric, directional transitions across distinct modalities (e.g., chatbot interface, draft document, other online sources), which AI-assisted revision often entails.

3 Research Objectives

The proposed study aims to explore how EFL students agentically engage in AI-assisted revision and how their agentic engagement may evolve over time. Research questions may include: 1) How do EFL students exhibit agentic engagement during AI-assisted revision? 2) How do EFL students' agentic engagement networks develop across six writing tasks over time? 3) What factors account for the development of EFL students' agentic engagement? The data for the study have been collected from four Chinese EFL undergraduates who participated in a four-month AI-assisted writing workshop. The primary data sources are 24 screen recordings that captured their AI-assisted revision process (four participants x six writing tasks) and stimulated recall interviews that followed each revision session. The study is currently in the conceptualization, codebook development, and data preparation stage.

4 Prior Experience and Expectations

I have participated in a QE workshop at AERA 2026, which helps me develop a basic understanding of QE. I have also conducted Epistemic Network Analysis on process data in an empirical study. However, I have not tried transmodal analysis before. Recent QE literature inspired me to apply T/ONA to an unanalyzed dataset from my previous work. I wonder whether this analytical approach can work well for the dataset and the proposed topic. During the discussion-based session at ICQE, I hope to learn from QE scholars about methodological challenges such as developing transmodal coding schemes, refining network analysis matrices, presenting longitudinal network trajectories, and connecting networks with qualitative interpretation.

References

1. Barrot, J. S. Using ChatGPT for second language writing: Pitfalls and potentials. Assessing Writing, 57, 100745 (2023).
2. Guo, K. & Wang, D. To resist it or to embrace it? Examining ChatGPT's potential to support teacher feedback in EFL writing. Education and Information Technologies, 29(10), 12653-12683 (2023).
3. Mahapatra, S. Impact of ChatGPT on ESL students' academic writing skills: A mixed methods intervention study. Smart Learning Environments, 11(1), 9 (2024).
4. Cengiz, B. C., Bilki, Z., Ataş, A. H. & Celik, B. Exploring second language writers' engagement with ChatGPT feedback: Revision behaviors and perceptions. System, 134, 103837 (2025).
5. Woo, D. J., Susanto, H., Yeung, C. H., Guo, K. & Fung, K. Y. Exploring AI-generated text in student writing: How does AI help? Language Learning & Technology, 28(2), 183-209 (2024).
6. Reeve, J. How students create motivationally supportive learning environments for themselves: The concept of agentic engagement. Journal of Educational Psychology, 105(3), 579-595 (2013).
7. Shaffer, D. W., Wang, Y. & Ruis, A. Transmodal analysis. Journal of Learning Analytics, 12(1), 271-292 (2025).
8. Wang, Y., Shah, M., Jimenez, F. A., Wilson, C., Ashiq, M., Eagan, B. & Shaffer, D. W. Developing nursing students' practice readiness with Shadow Health® Digital Clinical Experiences™: A transmodal analysis. In: Arastoopour Irgens, G., Knight, S. (eds) Advances in Quantitative Ethnography. ICQE 2023. Communications in Computer and Information Science, vol 1895, pp. 346-361. Springer, Cham (2023).

Examining Students' Experiences in Graduate Physics Programs

Reidyn Wingate[1], and Geraldine Cochran[1]

[1] The Ohio State University, Columbus, OH, 43210, USA
wingate.49@osu.edu

Abstract. This research aims to investigate the experiences of physics graduate students from different backgrounds and at different stages of their graduate education in order to determine the leading factors that help or hinder identity development in graduate physics programs.

Keywords: Physics Identity, Identity Development, Graduate Education

1 Area and Domain Research

The proposed research study is in the domain of Graduate Physics Education Research

2 Background

Many physics graduate education programs host students that hold different backgrounds, identities, and research experiences, which will help shape each student's graduate experience. Just as a student's identity can affect their graduate experience, their graduate experience can likewise affect how their identity as a physicist and as a person develops throughout their physics journey. Prior research indicates that the factors comprising physics identity—physics interests, peer and faculty recognition, and performance in the lab and the classroom—are related to and influenced by other aspects of students identity (Hazari et al., 2020, McDermott et al., 2024, Hyater-Adams, et al., 2018). We are particularly interested in the relationship between these aspects and physics identity that may vary by subdiscipline, research group culture, and student demographics.

3 Research Objectives

The aim of the proposed research is to better examine how factors such as interest, recognition, and performance can affect student identity development and how demographic identity connects to physics identity development. Students from different backgrounds (domestic vs. international), ethnicities, subdisciplines, and institutions

would be interviewed to get a holistic picture of their graduate experiences. Potential research questions include 1) How do students' graduate physics experiences vary across different institutions? 2) How do the graduate experiences of students from different backgrounds and ethnicities vary? 3) How do students' graduate experiences vary across different stages of their education? 4) How does a student's identity develop throughout their graduate education, and what are the primary factors that catalyze this development? 5) How does a domestic student's graduate experience differ from that of an international student? Once the data is collected through interviews, analysis would involve the identification of the primary identities that affect a student's graduate education, and the determination of which factors most affected the experiences and identity development of students of shared and varied identities.

4 Prior Experience and Expectations

This is my first study that would involve QE, but I have reviewed some QE literature as part of my research group. Because this study gets at the lived experiences of physics graduate students, I feel like it would benefit greatly from QE methodology and frameworks. I would love to learn more about QE interview and observation practice from a discussion-based session, as well as how to rigorously code and analyze data collected from interviews. I have not collected data yet or started this study. I would like to design this study with QE in mind from the beginning. My plan is to develop an interview protocol and submit for IRB approval during the Autumn 2026 semester. Then, I plan to collect interviews through Summer 2027. I would like feedback on the kinds of questions I hope to answer, the design of the study, and on a draft interview protocol.

References

1. Hazari, Z., Chari, D., Potvin, G., & Brewe, E. (2020). The context dependence of physics identity: Examining the role of Performance/competence, recognition, interest, and sense of belonging for Lower and upper female physics undergraduates. *Journal of Research in Science Teaching*, 57(10), 1583–1607
2. Hyater-Adams, S., Fracchiolla, C., Finkelstein, N., & Hinko, K. (2018). Critical look at physics identity: An operationalized framework for examining race and physics identity. *Physical Review Physics Education Research*, 14(1).
3. McDermott, L. G., Mosley, N. A., & Cochran, G. L. (2024, February 29). Diverging nonlocal fields: Operationalizing critical disability physics identity with neurodivergent physicists outside academia | Phys. rev. phys. educ. res.

Experiencing Differentiation: A Quantitative Ethnography of Introductory Physics Pathways

Stella F. Nelson[1 [0009−0003−4449−8971]] and Geraldine L. Cochran[1 [0000−0002−8520−6168]]

The Ohio State University, Columbus, OH 43210, USA
nelson.1898@osu.edu

Abstract. Universities increasingly offer differentiated introductory physics pathways tailored to different student populations, yet little is known about how students experience these courses or how those experiences differ across pathways. Framed by a four-dimension differentiation framework [3], we use quantitative ethnography to code semi-structured interviews for students' experiences of curriculum, instruction, support, and assessment, and compare the patterns of connection among those experiences across three introductory physics pathways at a large public university. This extends prior quantitative, curriculum-focused analysis of one sequence of these reforms to the student experience across all four dimensions and multiple pathways.

Keywords: Introductory physics · mechanics · student experiences · differentiation.

1 Area and Domain of Research

This study is situated within Physics Education Research (PER) at the introductory undergraduate level. It investigates students' experiences of differentiation across the multiple pathways The Ohio State University (OSU) offers for completing the introductory physics requirement.

2 Background

Efforts to investigate and improve introductory physics have a long history in PER [1, 2, 4, 5], and in prior work our team proposed a differentiation framework classifying such reforms along four dimensions — differentiated curriculum, instruction, support, and assessment. Using that framework and differentiated curriculum as a lens we analyzed OSU's Physics 1248–1249, an extended introductory mechanics sequence for engineering majors [3, 6]. That analysis was quantitative and curriculum-focused, and it left open questions about how students in that course and others offered at OSU experience differentiation across its other dimensions. This study takes up those questions.

OSU now offers three introductory pathways that differentiate to different degrees and for different populations: the standard one-semester Physics 1250 (all

STEM majors, calculus co-requisite); Physics 1270 (a deeper mechanics course for physics and astronomy majors); and the two-semester Physics 1248–1249 sequence (engineering majors, replacing the calculus co-requisite with college algebra and adding quantitative-literacy support). While each was designed with a distinct population in mind, little is known about how students experience these pathways or how those experiences differ across them. Using quantitative ethnography, we compare how students across the three pathways describe their experiences along the four dimensions of differentiation, and how well they feel their pathway prepares them for future coursework in their major.

3 Research Objectives

Our aim is to understand the experiences of engineering and physics students as they pursue the different pathways now offered at OSU. More specifically our research questions are:

- RQ1: How do students in each introductory physics pathway describe their experiences along the four dimensions of differentiation—curriculum, instruction, support, and assessment?
- RQ2: How do the patterns of connection among these dimensions of experience differ across pathways?
- RQ3: How do students' perceptions of how well their pathway prepares them for future coursework in their major relate to their experiences of differentiation, and how does that differ across pathways?

The data collection for this study will consist of semi-structured interviews to elicit participant storytelling about experiences in the introductory mechanics courses. We will recruit participants who have completed Physics 1248–1249, Physics 1250, or Physics 1270 in the past two years. Interviews will be conducted during the Autumn 2026–Spring 2027. We plan to code transcripts using a scheme that pairs deductive codes drawn from the four dimensions of the differentiation framework with inductive codes capturing experiential qualities the framework does not name. To address RQ2, we will construct and compare network models and compare the structure of connections across the three pathway groups. Following the view that theory should travel through every stage of a QE study, the differentiation framework will guide segmentation and interpretation as well as coding, and we will use member checking with participants to validate the coded experiences.

4 Prior Experience and Expectations

Stella: I am a post-baccalaureate student in physics education research. I am still new to quantitative ethnography. I hope to work with an expert in quantitative ethnography to review the current research plan and revise so that it is best suited for a QE analysis. **Geraldine:** I am a physics education researcher new to quantitative ethnography.

References

1. Etkina, E., Murthy, S., Zou, X.: Using introductory labs to engage students in experimental design. American Journal of Physics **74**(11), 979–986 (2006), https://doi.org/10.1119/1.2238885
2. Hamerski, P.C., Irving, P.W., McPadden, D.: Learning assistants as student partners in introductory physics. Physical Review Physics Education Research **17**(2), 020107 (2021), https://doi.org/10.1103/PhysRevPhysEducRes.17.020107
3. Nelson, S.F., Cochran, G.L., Imroz, K.A., Amalucottage, A.P., Myers, C., Parris, M., Thompson, C.M.: Rethinking introductory physics courses. In: Proceedings of the 2026 Physics Education Research Conference (Accepted)
4. Saul, J.M., Beichner, R.: An activity-based curriculum for large introductory physics classes: The scale-up project. In: Physics Education Research Conference (2001), https://doi.org/10.1119/perc.2001.pr.016
5. Stewart, J., Cochran, G.L., Henderson, R., Zabriskie, C., DeVore, S., Miller, P., Stewart, G., Michaluk, L.: Mediational effect of prior preparation on performance differences of students underrepresented in physics. Physical Review Physics Education Research **17**(1), 010107 (2021), https://doi.org/10.1103/PhysRevPhysEducRes.17.010107
6. White Brahmia, S., Cochran, G.L.: Reframing the narrative on physics readiness. Physics Today **78**(10), 40–47 (2025), https://doi.org/10.1063/pt.qflo.wdfw

Exploring IMPACT: Integrated Modeling Practices Advanced through Computational Thinking

Amanda Peel[1][0000-0002-2704-3911]

[1] New Mexico State University, Las Cruces NM 88003, USA
apeel@nmsu.edu

Abstract. The purpose of this RAD is to further develop and plan research projects related to supporting scientific computational modeling. The work aims to understand how science knowledge and computational thinking co-develop through scientific computational modeling. This RAD discussion will be used to plan data analysis of video data of students constructing computational models of science phenomena using a generative AI tool. It is expected that the RAD session will help hone research questions, inform coding, and plan how to answer research questions with QE tools and approaches.

Keywords: Science Education, Computational Modeling, Computational Thinking, Student Learning.

1 Area and Domain of Research

This work blends science education, learning sciences, and computer science education.

2 Background

Although science content learning has been measured with applications of computational modeling [e.g.,1], most of these approaches focused on *using* computational models, without expanding to *modifying* or *creating* computational models, which would further enhance the engagement in scientific inquiry to support science content learning and further development of computational thinking (CT) practices. Moreover, AI can now be leveraged to lower the barriers to programming [2], which will help students be able to modify and create scientific computational models.

Generative AI has been shown to enhance the interactivity of programming and has been utilized to improve students' code-writing and debugging skills [3]. Novice coders using generative AI is an area that needs to be empirically researched [4]. Early studies indicate that novices who struggle with computational problem-solving also continue to struggle with coding using generative AI, and in some cases, perform worse with these tools [4]. This is a gap in the literature and an essential focus area for research. AI-assisted coding in scientific modeling is so new that there is a significant gap in the science education and science modeling literature. Work needs to be done to identify how AI-assisted scientific computational modeling can be achieved, especially with the goal of enhancing science learning. Empirical work is needed to characterize what

students need to know about CT to be effective modelers. Moreover, we need to understand the impacts of AI-assisted scientific computational modeling on science learning.

3 Research Objectives

The aim of this work is to understand how science content knowledge and CT skills co-develop through the process of creating a computational model of a science topic with a generative AI tool. The interactions between CT, science content, and computational modeling are expected to be very complex. Characterizing these complex interactions and synergies would allow us to better understand how computational modeling can support student learning. The data is expected to be a video of a high school student or a pair of high school students capturing their model creation process. I also hope to have the screen capture to accompany the discussion. The research question guiding this work is: How do students' science content knowledge and CT skills co-develop and interact as they engage in computational model creation? By the conference, I hope to have a pilot completed with data from one attempt at computational model creation.

4 Prior Experience and Expectations

I have worked with ENA and ONA before to do QE work on student artifact data. I have been a QE Fellow and Senior Fellow, so I have experience with the web tool, analyzing data, and publishing QE work. I have not, however, used QE to analyze discourse or video data. It is possible trans-modal analysis (TMA) would be appropriate for this work. I need more support to better understand how to do QE on this type of data. The earlier in the research process that I get support in this, the better, because then I can effectively plan for a scale-up study and/or grant proposal to extend the work.

References

1. Arastoopour Irgens, G., Dabholkar, S., Bain, C., Woods, P., Hall, K., Swanson, H., Wilensky, U. (2020). Modeling and Measuring High School Students' Computational Thinking Practices in Science. *Journal of Science Education and Technology, 29*(1), 137-161.
2. Sarkar, A. (2023, October). Will code remain a relevant user interface for end-user programming with generative ai models?. In Proceedings of the 2023 ACM SIGPLAN International Symposium on New Ideas, New Paradigms, and Reflections on Programming and Software (pp. 153-167).
3. Kazemitabaar, M., Chow, J., Ma, C. K. T., Ericson, B. J., Weintrop, D., & Grossman, T. (2023). Studying the effect of AI code generators on supporting novice learners in introductory programming. *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*, 1–23.
4. Prather, J., Reeves, B. N., Leinonen, J., MacNeil, S., Randrianasolo, A. S., Becker, B. A., ... & Briggs, B. (2024, August). The widening gap: The benefits and harms of generative ai for novice programmers. In *Proceedings of the 2024 ACM Conference on International Computing Education Research-Volume 1* (pp. 469-486).

Exploring Quantitative Ethnography Applications to Bridge Program Evaluation and Research in a Culturally-Sustaining Cascading Mentorship Pilot

Daranee Taychachaiwongse Teng[1][0009-0005-6530-0770]

[1] University of Colorado Denver, Denver CO 80204, USA
Daranee.Teng@ucdenver.edu

Abstract. This proposal solicits support in identifying potential Quantitative Ethnography (QE) applications for a dual-tiered study centering an intergenerational mentoring model that aims to advance postsecondary equity through two integrated components: a developmental program evaluation study of a Cascading Mentorship Pilot (CMP) assessing mentorship demand, capacity, efficacy, and scalability; and (2) a qualitative study exploring student and faculty perspectives on culturally-sustaining mentorship experiences.

Keywords: culturally-sustaining, cascading mentorship, evaluation, student success

1 Area and Domain of Research

This proposal pertains to educational research and evaluation, with a specific focus on culturally-sustaining mentorship in postsecondary contexts.

2 Background

Nationwide, Asian Americans, Native Hawaiians, and Pacific Islanders (AANHPIs) face barriers to postsecondary access and success [1], yet challenges are often overlooked due to model minority stereotypes that obscure the struggles of many members within the broad AANHPI umbrella [2] [3]. This generalization invisibilizes the disparate educational and social outcomes within the AANHPI community, underscoring the need for further research investigating the barriers to postsecondary success that AANHPI encounter during their educational journeys [4].

The proposed research engages cascading mentorship as a nested approach to supporting student success across multiple levels of a scholarly community, where faculty members mentor graduate students, and in turn, graduate students mentor undergraduate students. Cascading mentorship has found success in academic spaces nationally and internationally [5][6]. This study grounds the cascading model in the values of Hānai Pedagogy, which includes hands-on activities, aloha, navigation, authenticity, and interrelations [7][8].

3 Research Objectives

This dual-tiered project integrates a multiple and mixed methods program pilot evaluation study with a qualitative research study. The primary aims of the research

are: (a) to explore AANHPI student and faculty perspectives on postsecondary equity and sense of belonging to inform culturally sustaining research mentorship experiences (The Qualitative Study), and (b) to evaluate demand, capacity, efficacy, and viability of scaling and institutionalizing a cascading mentorship framework to support student success at an urban public research university (The Pilot Evaluation Study). AANHPI faculty, doctoral students, and undergraduate students will hold dual roles acting both as pilot participants and as the research team conducting the qualitative study. A conceptual framework can be found here.

The first tier of the study includes developmental evaluative research of Cascading Mentorship Pilot (CMP) by testing the following hypotheses:

1. **Hypothesis 1: Demand.** CMP participation will foster interest in research for AANHPI undergraduate students through rich, supportive, and culturally sustaining research experiences.
2. **Hypothesis 2: Capacity.** Participating faculty members and doctoral student mentors have the capacity and institutional support to provide culturally affirming guidance to support undergraduate student mentees in navigating pathway options based on their personal and professional goals.
3. **Hypothesis 3: Efficacy.** CMP participation will support AANHPI undergraduate and graduate student belonging and retention at the institution.

Evaluation data will be collected via surveys; journal reflections; and research team meeting observations. Ultimately, this developmental evaluation is designed to yield an evidence-based framework and implementation recommendations to ensure the model's replicability to other institutional contexts beyond the single-site university setting of this Pilot Evaluation Study. Institutional funding has been awarded to conduct this study, and IRB approval has been obtained.

The second tier is a qualitative research study (led by the cascading mentorship team) to understand postsecondary equity, capacity barriers, and factors that would support the long-term success and institutionalization of the CMP model at scale. The overarching research question guiding the qualitative study is “What can we learn from the perspectives of AANHPI students and faculty to improve equity, belonging, and learning through research mentorship?” Interviews and focus groups will be conducted with the broader AANHPI student and faculty population.

4 Prior Experience and Expectations

With less than one year of experience with QE, I would greatly appreciate support in identifying QE applications for this study. I am curious if and how it may be meaningful to map Epistemic Network Analysis (ENA) and/or Ordered Network Analysis (ONA) modeling on to intended outcomes in the pilot’s logic model. For example, with ONA, does the temporal sequence of discourse data in journal reflections demonstrate that participating in certain pilot activities preceded the CMP participant reporting that their mentor helped them develop knowledge, skills, and mindsets that support their long-term academic success? I hope to engage with experienced QE scholars to conceptualize how QE, ENA, and/or ONA might be used to model how learning processes unfold through intergenerational mentorship experiences in higher education and influence postsecondary success outcomes.

References

1. National Commission on Asian American and Pacific Islander Research in Education: The relevance of Asian Americans & Pacific Islanders in the college completion agenda. (2011). http://care.gseis.ucla.edu/wp-content/uploads/2015/08/2011_CARE_Report.pdf
2. Chou, R. S., & Feagin, J. R.: The myth of the model minority: Asian Americans facing racism. Paradigm Publishers (2008)
3. Suzuki, B. H.: Education and the socialization of Asian Americans: A revisionist analysis of the "model minority" thesis. Amerasia Journal, 4(2), 23–51 (1977)
4. Gogue, D. T., Venturanza, R. J., Cuenza-Uvas, A., & Nguyen, M. H.: The role of Asian American and Native American Pacific Islander‐Serving Institutions in reframing leadership education. New Directions for Student Leadership, 2021(171), 101–111 (2021)
5. UC Irvine School of Social Sciences: Cascading Mentorship Groups. (2026) https://www.gradstudies.socsci.uci.edu/archive/archive_cascading_mentorship.php
6. Guo, T., Chowdhury, M., Rasouli, R., & Patel, M.: Exploring the effectiveness of a cascading mentorship model in developing CanMEDS competencies in postgraduate medical education: a qualitative interview study among resident mentors at a medical school in Canada. BMJ open, 13(1), e061338. (2023)
7. Brandehoff, R.: Introducing Hānai pedagogy: A call for equity in education through an AANAPI lens. Pedagogy and Theatre of the Oppressed Journal, *7*(3), 1-15 (2023)
8. Brandehoff, R. Ignite the night to keep talk story: Hānai pedagogy as an act of liberation. Journal of Ethnic and Cultural Studies, 11(5) 140-157 (2024)

Functions of Computational Writing in Pharmaceutical Data Science Teaching

Morten Misfeldt [0000-0002-6481-4121], Adrienne L. Traxler [0000-0003-2725-0686], , Frederik Voetmann Christiansen [0000-0002-6743-9881], Stine Marie Høgh Jensen [0009-0004-4605-9080], Hendra Agustian [0000-0003-2792-3035], Lykke Pedersen [0000-0002-8935-415X], and Morten Misfeldt [0000-0002-6481-4121] and Arnault-Quentin Vermillet [0009-0003-1227-4199]

[1] University of Copenhagen, Copenhagen Denmark
arnault@sund.ku.dk, misfeldt@ind.ku.dk

Abstract: This research discusses a coding framework for studying computational writing in AI-infused computational notebooks in pharmaceutical data science education. We treat notebooks not only as containers for prose, code, data, and visualizations, but as multimodal writing environments in which students document, explore, justify, explain, and position their work. The framework combines four modalities: prose, code, data, and visualization, with five writing functions: transparency, ideation, justification, explanation, and identity work. It is developed to support quantitative ethnographic analysis of student notebooks and to examine when notebook-based teaching fosters inquiry and when it becomes procedural "cookbook" execution. We invite discussion of the framework's analytical usefulness and its implications for reflective AI-infused notebook pedagogy.

Keywords: data science education, pharmacy education, writing, multimodal notebooks.

Computational Writing in University Science Teaching

Computational notebooks (Jupyter, R markdown, Google Colab, etc.) are becoming central for academic work across a range of disciplines. These environments integrate prose, executable code, data, and visualizations, allowing analysis, experimentation, documentation, and communication to occur within a shared computational workspace [1]. As such, they increasingly shape both scientific practice and university teaching in science and health, but their pedagogical role remains ambiguous. While notebooks are often presented as tools for inquiry, exploration, and authentic scientific practice, educational uses may be highly proceduralized. Students may primarily execute predefined workflows, modify isolated parameters, or reproduce teacher-prepared analyses without engaging deeply with the underlying reasoning processes. In this sense, notebook teaching risks becoming a form of "computational laboratory cookbook" pedagogy [2].

This tension motivates the present study. We approach computational notebook work as a writing process and as an educationally situated epistemic practice. We will

discuss an emerging coding framework for analyzing computational writing distinguishing writing modalities and writing functions. The framework will be used to support future empirical investigations of how students coordinate modalities, reasoning processes, and AI-supported forms of delegation while producing computational texts in pharmacy education, where computation increasingly influences reproducibility, regulatory work, empirical investigations and drug design [3].

Worksheets combining text and calculations have existed since the 1980s. But their importance has grown with the emergence of environments that combine prose, code in general purpose programming languages (e.g. python), data, and visualizations. Different research has explored notebook environments. In human-computer interaction research, they have been studied as interfaces for interactive computing, exploratory data analysis, and collaborative work with data and computation [4]. In educational research, recent work on "computational essays" has investigated how notebooks may support epistemic agency in science education [1]. Computational notebook practices also have important continuities with earlier mathematical tools adopted in the classroom where the interplay between tool based techniques and epistemic agency has been a central theme in mathematics education [5]. We also draw on elements from writing and literacy research [6-8], as well as previous QE work on coding code (REF).

To analyze computational writing, we propose a multimodal-functional framework that combines four modalities of computational notebook work with five central writing functions. The modalities are prose, code, data, and visualization. Across these modalities, we distinguish five writing functions: transparency, ideation, justification, explanation, and identity work.

Questions to Discuss at ICQE

We explore how computational writing may be investigated within quantitative ethnography. At the conference, we will present early analyses of pharmacy students' Jupyter notebook reports and discuss how the framework may support quantitative ethnographic studies of computational writing. One question concerns the proposed coding framework itself: whether the distinctions between modalities and writing functions are analytically productive, sufficiently distinct, and useful for studying notebook-based educational practices. Relatedly, we are interested in how similar approaches to writing, multimodality, and epistemic practice have been addressed within existing QE research, including possible connections to coding of discourse, representational work, and epistemic networks. Another question concerns data and methodology. Can computational notebook work be analyzed through the written products alone, or do we need complementary forms of data?

Acknowledgments. This study was funded by the Novo Nordisk Foundation (grant number NNF24SA0092613), the LEO Foundation (grant number LF-ST-24-500013), and the Lundbeck Foundation (grant number R469-2024-629). **Disclosure of Interests.** Nothing to declare apart from above funding acknowledgment.

References

1. Odden, T. O. B., & Zwickl, B. (2025). How physics students build computational literacy by creating computational literature. Journal of the Learning Sciences, 34(5), 700–742. https://doi.org/10.1080/10508406.2025.2494791
2. Wilkerson, M.H., Shin, J.M. & Ban, S.Y. (2026). Open science meets orange skies: K-12 educators' engagement with *computational notebooks to explore scientific events with human impacts. ZDM* Mathematics Education https://doi.org/10.1007/s11858-026-01798-0
3. Peng, K., Huang, Y.N., Sarwal, V., Alachkar, H., Wong-Beringer, A. and Mangul, S. (2021), Integrating big data computational skills in education to facilitate reproducibility and transparency in pharmaceutical sciences. *J Am Coll Clin Pharm*, 4: 1263-1266. https://doi.org/10.1002/jac5.1519
4. Harden, J., Wang, A. Y., Faust, R., Isaacs, K. E., Kirshenbaum, N., Wenskovitch, J. & North, C. (2024). Human-Notebook Interactions: The CHI of Computational Notebooks. In *Extended Abstracts of the CHI Conference on Human Factors in Computing Systems* (pp. 1-6).
5. Anonymous reference
6. Scardamalia, M., & Bereiter, C. (1987). Knowledge telling and knowledge transforming in written composition. Advances in Applied Psycholinguistics, 2, 142–175.
7. Galbraith, D. (2009). Writing as discovery. In V. Connelly, A. L. Barnett, J. E. Dockrell, & A. Tolmie (Eds.), BJEP monograph series II: Part 6 teaching and learning writing (Vol. 2, pp. 5–26). Wiley. https://doi.org/10.1348/978185409X421129
8. Cazden, C., Cope, B., Fairclough, N., Gee, J., Kalantzis, M., Kress, G., Luke, A., Luke, C., Michaels, S., & Nakata, M. (1996). A pedagogy of multiliteracies: Designing social futures. Harvard Educational Review, 66(1), 60–92.
9. Pankiewicz, M., Zambrano, A.F., Barany, A., Baker, R.S. (2024). How We Code Code: Leveraging GPT and Ordered Networks for Understanding Introductory Programming Education. In: Kim, Y.J., Swiecki, Z. (eds) *Advances in Quantitative Ethnography. icqe 2024. Communications in Computer and Information Science*, vol 2279. Springer, Cham.

Mapping Out Who Teaches Computer Science in New Mexico

Adaline De Chenne[1[0000-0001-7574-8505]], Christabel Wayllace[1[0000-0001-8039-2777]], Amanda Peel[1[0000-0002-2704-3911]], Paige Prescott[1[0000-0001-6160-2652]]

[1] New Mexico State University, Las Cruces NM 88003 USA
dechenne@nmsu.edu

Abstract. This research aims to understand how computer science (CS) is taught at the K-12 level in New Mexico (NM), United States by mapping out the backgrounds of current and prospective CS teachers in the state. This study is in the analysis phase, and it relies on data of convenience collected by a professional development program in NM over the span of six years. The goal of our analysis is to use ENA to develop a network for each teacher that includes their background in CS, how they teach CS at their schools, and their experience teaching CS. Then, we will use ENA to compare networks across years. The goal of this project is to inform on the current state of K-12 CS teaching in NM, and provide avenues for improving on this teaching in the future.

Keywords: Computer Science Education, Teacher Experiences, K-12 education, Teacher Learning

Area and Domain of Research

This proposed research study is in the domains of Teacher Education and Computer Science Education.

1 Background

New Mexico (NM) is a located in the southwest United States, and outside three population centers (Albuquerque, Las Cruces, and Santa Fe) the state has a low population density. Many NM public schools are both rural and Title 1 (i.e., serving large numbers of low-income families), which can provide challenges for adopting and adapting to new teaching requirements. New Mexico has recently mandated that all high schools offer a computer science (CS) elective course, but it remains unclear how public schools will do so. One possible elective includes Advanced Placement Computer Science principles (AP CSP), which requires a license to teach. To address the low numbers of teachers with this license, NM has allowed for the license to be granted following the completion of sixty hours of related professional development training. Although, there are other pathways, professional development is the most common.

Computer science is also a discipline that is relatively new at the K-12 level with only 54% of high schools offering a foundational CS course [1]. This means that there is not a well-established path to become a CS teacher. New Mexico State University, the second largest university in the state, only began offering a degree pathway with explicit CS teaching expertise in the last year. Most current CS teachers at the high school level have expertise in other areas [2]. There are questions concerning how teachers develop expertise in CS, how current teachers become responsible for starting and leading CS courses, and what CS teaching in NM currently looks like.

We have collected longitudinal data from a yearly professional development program that aims at helping teachers develop the skills necessary to teach AP CSP. These data were collected over the course of six years, and they are data of convenience in the sense that the original intent of the data was to shape the PD program. However, these data include valuable information about the teachers, including where in the state they are located, their backgrounds with CS, and their courseload.

2 Objectives

The aim of this proposed research is to better understand the current landscape of high school CS teaching in NM, and to provide information that can improve on it. We aim to develop a network for each teacher that informs us on their background, and to develop a mean network per year of data. Then, we will examine how those networks changed from year to year. Possible research questions include: (1) what academic backgrounds are common among current and prospective CS teachers? (2) how does geographic location relate to CS teaching and CS teachers?, and (3) how might teacher development programs adapt to better serve the current need for CS teachers? (4) How do these trends change over time? Data have already been collected for this project.

3 Prior Experience and Expectations

Four of the five authors have no prior experience with QE tools or techniques. Our expectations for QE are still being shaped, although we believe that QE could help us better profile and understand current and prospective CS teachers in NM in a way that is challenging using standard statistical tools. During this session, we are interested in better understanding what QE is, how to apply QE tools to our data, and what insights QE might provide.

References

1. Code.org (2025). *2025 State of AI + CS Education*.
2. Prescott, P., Peel, A., De Chenne, A., Henderson, J., Baeza, S., La Cour, T., Wayllace, C., & Lopez, P. (2026). *A Longitudinal Analysis of Computer Science Professional Development and Implementation in New Mexico.* Proposal submitted to the Society of Information Technology and Teacher Education (SITE) 2026 meeting.

Mapping the Mindset: Reframing Epistemic Network Analysis as a Hermeneutic Tool in Philosophical Doctoral Mentoring

Mercedes Diaz[1]

[1] Rutgers University, New Brunwsick, NJ 08901, USA

[2]Springer Nature Switzerland AG, Gewerbestrasse 11, 6330 Cham, Switzerland
lncs@springer.com

Abstract. This discussion paper evaluates a novel theoretical framework designed to bridge the gap between Epistemic Network Analysis (ENA) and philosophical doctoral mentoring. Traditionally utilized to map cognitive connections in empirical data, ENA faces quantification resistance when applied to disciplines reliant on qualitative nuance. By reframing ENA as a hermeneutic tool grounded in Gadamerian hermeneutics, this framework transforms network visualizations into interpretative texts. This shift allows mentors and mentees to engage in collaborative cartography, tracking the evolution of a student's thesis and diagnosing structural issues like conceptual drifts or logical gaps. Ultimately, this project expands the scope of Quantitative Ethnography (QE) into non-empirical humanities fields.

Keywords: Epistemic Network Analysis (ENA), Quantitative Ethnography, Philosophical Discourse

1. Introduction

The humanities, particularly philosophy, present a unique challenge to modern empirical and analytical tools. Philosophy operates on conceptual analysis, complex argument construction; qualities that resist standard statistical quantification. Within philosophy doctoral programs, mentoring often relies on the traditional "Master-Apprentice" model, which can result in disjointed feedback and raw semantic information that fails to coalesce into reliable philosophical knowledge. Epistemic Network Analysis (ENA) offers a robust methodology for modeling cognitive connections, but its empirical origin creates an ontological and structural mismatch when applied directly to philosophical discourse. There is a critical research gap in determining how to adapt Quantitative Ethnography (QE) tools so that they honor complex human reasoning without reducing philosophy to rigid statistical metrics.

1.1 Theoretical Framework: Gadamerian Hermeneutics

To overcome the quantification resistance of philosophy, this study introduces a theoretical framework grounded in Gadamerian hermeneutics. Rather than viewing ENA as an objective measurement tool, the framework reframes it as an interpretative instrument. The key components of this adaptation include:

- **Network Visuals as Texts:** Visual graphs generated by ENA are treated as qualitative texts requiring active interpretation and deep philosophical reflection.

- **The Web of Mutual Relations:** Following Donald Davidson's assertion that a belief's identity is determined by its location within a network of truth, ENA treats individual philosophical claims as nodes. This positions semantic data into a mutual relation framework, tracking growth over time.

1.2 Research Objectives and Methodology

The overarching goal of this developmental research is to bridge the methodological gap between empirical data visualization and qualitative philosophical introspection. The core objectives focus on:

1. Adapting ENA methodologies specifically for humanities-based doctoral mentoring.
2. Shifting the mentoring paradigm to a collaborative, visual dialogue.
3. Utilizing visual network graphs to diagnose underlying writing issues, including structural gaps, implicit biases, and conceptual drifts in draft dissertations.

1.3 Evaluation Question

- **Conceptual Alignment (Davidson's Network):** When you look at an ENA graph, do you see a node as a static concept or does its meaning change based on the surrounding cluster of nodes? How accurately does spatial distance capture logical dependencies between premises and conclusions?

2. Interview Evaluation Rubric

This qualitative rubric measures how effectively participants internalize Davidson's holistic network concept, engage in collaborative cartography, and use ENA to achieve scholarly independence.

Evaluation Dimension	Low Integration (Level 1)	Moderate Integration (Level 2)	High Integration (Level 3)
Conceptualization of ENA Nodes	Views nodes as isolated, static variables or mere word counts; disconnected from the "web of truth."	Recognizes that nodes derive context from neighboring connections but struggles to translate clusters into arguments.	Synthesizes node proximity with semantic dependency; explains how a shift in one node alters the network's logical truth-value.
Engagement in Cartography	Uses the ENA graph predictably as a top-down, evaluative report card rather than an open interpretive tool.	Discusses the graph jointly but relies on traditional textual margin notes to drive core philosophical breakthroughs.	Uses the graph as an active text; co-constructs new arguments and resolves conceptual mismatches by visually mapping ideas in real-time.

Diagnostic Utility	Fails to notice visual anomalies; dismisses graph outliers or isolated nodes as meaningless technical errors.	Identifies visual anomalies (e.g., conceptual drifts) but requires manual, text-by-text reverse engineering to pinpoint errors.	Uses visual anomalies proactively to unearth unstated premises, blind spots, or structural contradictions in the text immediately.
Epistemic Transition	Cannot see or interpret personal intellectual growth or independence within the shifting geometric density of the graphs.	Acknowledges structural changes between early and late graphs but cannot articulate how those shifts relate to autonomy.	Articulates a clear shift from replicating the mentor's network configuration to carving out an independent semantic cluster.

3. Conclusion and Future Research

Reframing ENA as a hermeneutic tool opens new avenues for Quantitative Ethnography by proving that mathematical consistency and data visualization can be deeply meaningful in disciplines that reject traditional empiricism. By leveraging moving stanza windows and semantic network modeling, graduate coordinators can structurally align and optimize doctoral milestones, validating complex human experiences while maintaining mathematical rigor.

Acknowledgments. This ongoing study is courtesy of the Rutgers Philosophy Department and Rutgers Graduate School of Education.

Disclosure of Interests. The author is not aware of any competing studies on the topic.

References

1. Allington, D., Brouillette, S., & Golumbia, D. (2016). Neoliberal tools (and archives): A political history of the digital humanities. *Los Angeles Review of Books*, 1, 1-24.
2. Bowker, G. C., & Star, S. L. (2000). *Sorting things out: Classification and its consequences*. MIT Press.
3. Davidson, D. (2001). *Inquiries into truth and interpretation* (2nd ed.). Oxford University Press.
4. Floridi, L. (2011). *The philosophy of information*. Oxford University Press.
5. Gadamer, H.-G. (2013). *Truth and method* (J. Weinsheimer & D. G. Marshall, Trans.; rev. ed.). Bloomsbury Academic.
6. Ricoeur, P. (1981). *Hermeneutics and the human sciences: Essays on language, action and interpretation*. Cambridge University Press.
7. Ruis, A. R. (2021). The challenge of the qualitative: Quantitative ethnography and the humanities. *Journal of Learning Analytics*, 8(2), 45-56.
8. Siebert-Evenstone, A. L., Arastoopour Irgens, G., Collier, W., Swiecki, Z., Ruis, A. R., & Shaffer, D. W. (2023). In search of conversational grain size: Modeling semantic structure using moving stanza windows. In A. R. Ruis & S. B. S. Lee (Eds.), *The foundations and fundamentals of quantitative ethnography* (pp. 120-142). Springer.

Neurodiversity and a new science of learning

Hendra Y. Agustian[1][0000−0003−2792−3035]

University of Copenhagen, Denmark
hendra.agustian@ind.ku.dk
lablearning.ku.dk, cpdse.dk

Abstract. This Research Agenda Development outlines a QE framework for studying neurodiverse learning in pharmaceutical data science education. Positioned within a proposed interdisciplinary project, the agenda leverages data science techniques, such as multimodal data integration and structural modeling, to capture learning as an embodied, affective, neurocognitively varied, and socially situated process. QE serves as the analytic scaffold that connects these techniques to ethnographic meaning-making through epistemic network analysis and transmodal analysis. Aligned with priorities in pharmaceutical data science education, the agenda advances transferable methods for modeling complex human–data systems while foregrounding equity, heterogeneity, and interpretive validity in neurodiversity research.

Keywords: Science of learning · Neurodiversity · Data science · Higher science education

1 Research domain

This proposed research falls within the domain of the Science of Learning. Highly interdisciplinary, the project gathers expertise in neurodiversity studies, social psychology, data science, and higher science education.

2 Background

Neurodiversity, i.e. the natural variation in human cognition and neurocognitive functioning, has become an increasingly urgent issue in higher science education [1, 2]. The number of neurodivergent students are increasing in STEMM fields, yet they also experience disproportionately high levels of exclusion, sensory-cognitive overload, and attrition [3–5]. Existing research on learning in advanced science largely draws on cognitive or neuroscientific measures that risk reductionist interpretations [6], which can be problematic in neurodiversity research, or on qualitative approaches that struggle to provide scalable, structural evidence of learning processes in context.

The proposed SOLACE (A New Science of Learning Science) project addresses this challenge by positioning data science as a central epistemic and methodological resource for understanding science learning as a complex and

contextual phenomenon [7, 11]. Within SOLACE, Quantitative Ethnography (QE) serves as the organizing analytic paradigm that allows data science techniques to be tightly coupled with ethnographic sensitivity to meaning, activity, and context.

This RAD articulates a research agenda in which data science and QE are co-constitutive: data science enables scalable pattern detection across multimodal learning data, while QE provides the theoretical and methodological constraints needed to retain interpretive validity, equity, and contextual grounding.

3 Objectives

The RAD aims to advance QE as a rigorous and equity-oriented methodology for studying neurodiverse learning in advanced science education. The objectives are threefold. First, it aims to develop data science pipelines that support QE analyses of multimodal learning data (e.g., discourse, video, observational logs, and phenomenological accounts). Second, it applies QE techniques such as epistemic network analysis (ENA) and transmodal analysis to model structural relationships among epistemic practices, affective states, social interaction patterns, and environmental conditions during authentic science inquiry activities. And last, it leverages data science to represent heterogeneity within and across learners, rather than relying on deficit-based or norm-referenced comparisons. QE network structures are used to model individual and group learning trajectories, aligning with neurodiversity-affirming research principles.

At ICQE, this RAD seeks scholarly dialogue on analytic design, validation strategies, and ethical considerations when combining QE with neurodiversity research.

4 Prior experience and expectations

In higher science education and learning sciences, I have sustained focus on epistemic practices [9], embodiment [10], affect [11], and motivation [12] in naturalistic settings (lablearning.ku.dk). I have led empirical studies using rich qualitative and multimodal data and have published methodological work on evidence synthesis and rigor in educational research [13]. My prior research in inclusive education [14] is also relevant to the critical perspective on neurodiversity.

Parallel to this, I am currently involved in the Center for Pharmaceutical Data Science Education (cpdse.dk), where data science is both the subject matter and the research methodology. This includes experience with designing data-intensive educational studies, and applying machine learning for pattern detection.

SOLACE represents a convergence of these trajectories. Through ICQE, I expect to refine the data science dimensions of QE designs, contribute to methodological conversations on equity-oriented analytics, and build collaborations that advance QE as a core methodology for data-intensive, interdisciplinary research.

References

[1] Doyle, N. (2020). Neurodiversity at work: a biopsychosocial model and the impact on working adults. *Br. Med. Bull., 135*(1), 108-125.

[2] Thorp, H. H. (2024). Science needs neurodiversity. *Science, 384*(6694), 365-365.

[3] Robinson, D. (2022). Neurodiversity in medical education: How can we improve postgraduate learning for neurodiverse doctors? *Med. Teach., 44* (5), 564–566.

[4] Wang, K. D., McCool, J. & Wieman, C. (2024). Exploring the learning experiences of neurodivergent college students in STEM courses. *J. Res. Spec. Educ. Needs, 24* (3), 505–518.

[5] Stone, S. & Arenas, B. E. (2026). Characterising the nature and effect of sensory overload in an undergraduate chemistry teaching laboratory. *Chem. Educ. Res. Pract., 27*(1), 304-316.

[6] Choudhury, S. & Slaby, J. (Eds.). (2011). *Critical Neuroscience: A Handbook of the Social and Cultural Contexts of Neuroscience.* Wiley.

[7] Illeris, K. (2018). A comprehensive understanding of human learning. In *Contemporary Theories of Learning* (pp. 1-14). Routledge.

[9] Agustian, H. Y. (2025). The critical role of understanding epistemic practices in science teaching using wicked problems. *Sci. & Educ., 34*(1), 485-510.

[10] Agustian, H. Y., Gammelgaard, B., Rangkuti, M. A. & Kersting, M. (2026). Embodied, Epistemic Cognition in Higher Science Education: Advances from the Chemistry Laboratory. *Sci. & Educ.*, 1-48.

[11] Agustian, H. Y., Gammelgaard, B., Rangkuti, M. A. & Niemann, J. (2025). "I feel like a real chemist right now": Epistemic affect as a fundamental driver of inquiry in the chemistry laboratory. *Sci. Educ., 109*(3), 722-744.

[12] Agustian, H. Y., Gammelgaard, B., Rangkuti, M. A. & Ryberg, M. L. (2026). Striving to learn to do chemistry in the laboratory: epistemic conation as a fundamental driver of inquiry. *Chem. Educ. Res. Pract, 27*(1), 80-100.

[13] Agustian, H. Y. (2024). Methodological rigor in laboratory education research. *Laboratories, 1*(1), 74-86.

[14] Agustian, H. Y. (2021). Investigating inclusive practice in international schooling: a case study of two schools. *J. Res. Internat. Educ., 20*(3), 226-241.

Reimagining Achievement Goal Theory for Minoritized Engineering Students

Dina Verdín[1][0000-0002-6048-1104]

[1] Arizona State University, Mesa AZ 85212, USA
dina.verdin@asu.edu

Abstract. This project seeks to reimagine achievement goal theory (AGT) so that it aligns with the goal motives of minoritized engineering students. This study is in the pilot data analysis phase; 14 interviews were collected. The pilot data will help gather early insights and refine the interview protocol before the grounded theory work begins. I would like to learn and explore the use of QE in the pilot data so I can be prepared to fully implement its techniques in the large-scale interview data.

Keywords: motivation, minoritized students, grounded theory, theory development, culturally responsive

1 Area and Domain of Research

This proposed research study is in the domain of Engineering Education.

2 Background

Theories are crucial to any research project because they provide a roadmap for data analysis and interpretation and position a research project within the body of existing knowledge. Many theories fall short of providing an inclusive understanding of minoritized students because they were developed from majoritarian viewpoints [1]. The few theories developed specifically for minoritized students (e.g., community cultural wealth, funds of knowledge) shed no light on students' learning and motivational pursuits. Consequently, engineering education scholars seeking to understand the academic motivational pursuits of minoritized students are constrained. Traditional psychological theories often take a generalized approach to human behavior and learning, overlooking the specificities of engineering contexts. Discipline-specific theories can more accurately capture the nuances of engineering ways of knowing, doing, and professional practice. Theories like Achievement Goal Theory (AGT) may not adequately capture the goal motives that drive minoritized engineering students' responses to challenges and setbacks as they navigate an exclusionary culture. AGT uncovers *why* a person behaves as they do when faced with setbacks and how those goals guide their behavior [2]. AGT was developed without considering the influence

of students' culture or unique lived experiences; to date, no study has reconceptualized the theory's core assumptions to consider minoritized students' unique life circumstances. Scholars who contributed to the development of AGT approached its formation from a study of personality [3]. Consequently, the philosophical tradition rooted in studies on personality assumed an individual's motivation was "in participants' heads," not in the learning context or their sociocultural realities [4].

3 Research Objectives

I will demonstrate how personal histories, cultural nuances, and socioeconomic factors can be incorporated to understand Latinx students' achievement motivational orientations. This project is focused on Latinx engineering undergraduate students enrolled in U.S. institutions. My research questions are: RQ1: When using a grounded theory approach to race-reimagine AGT, what context-specific achievement goals emerge among Latinx engineering students? RQ2a: What are the distinct goal profiles of Latinx engineering students, and what are the implications of these profiles for their academic outcomes, engineering identity, and engagement? I currently have a pilot interview dataset of 14 participants intended to help me refine the final interview protocol and capture additional insights to incorporate for the purpose of theory generation. My initial coding cycle generated 520 codes to allow for the development of context-specific goal motivators.

Another wave of interview data will be collected to answer the research questions. RQ1 will be answered through 90-minute interviews with 20-40 Latinx students, analyzed using grounded theory. RQ2 will be answered using survey data after the findings from RQ1 are translated into a validated survey instrument.

4 Prior Experience and Expectations

I have no prior experience with QE, and I wish to gain more insight into how its techniques could be applied to a grounded theory study. During the session, I hope to learn how to properly code the data so that it aligns with grounded theory and QE techniques. I'm also interested in learning how to use nCoder, SCAT, how to validate codes, and how to analyze interview data using epistemic network analysis. I would like to be able to make sense of the ENA visuals to help interpret my data.

References

[1] J. T. DeCuir-Gunby and P. A. Schutz, "The Evolution of Race-Focused and Race-Reimaged Approaches in Educational Psychology: Future Directions for the Field," *Educ. Psychol. Rev.*, vol. 36, no. 1, pp. 1–24, 2024, doi: 10.1007/s10648-024-09873-2.

[2] A. J. Elliot, "Energization and direction are both essential parts of motivation," *Motivation science: Controversies and insights*, pp. 10–14, 2023.

[3] E. M. Anderman, C. C. Austin, and D. M. Johnson, “The Development of Goal Orientation,” *Development of Achievement Motivation*, pp. 197–220, 2002, doi: 10.1016/b978-012750053-9/50010-3.

[4] A. Zusho and K. Clayton, “Culturalizing achievement goal theory and research,” *Educ. Psychol.*, vol. 46, no. 4, pp. 239–260, 2011, doi: 10.1080/00461520.2011.614526.

Revealing Hidden Patterns in Peer Feedback: An Epistemic Network Analysis of Student Beliefs and Practices

Wanju Huang[1][0000-0001-5965-2597]

Ryan Wynkoop[2][0000-0002-4057-5242]

Janet Alsup[3][0000-0002-8247-3674]

[1] Purdue University, West Lafayette IN 47907, USA
wanjuhuang@purdue.edu

[2] Purdue University, West Lafayette IN 47907, USA
rwynkoop@purdue.edu

[3] Purdue University, West Lafayette IN 47907, USA
jalsup@purdue.edu

Abstract. This research aims to understand undergraduate engineering students' perceptions of giving and receiving feedback, as well as examining the relationship between their beliefs and actions during feedback. It involves a medium-sized class (n=30), with data collected from an undergraduate engineering course. Several datasets, including students' peer-assessment written entries and their reflections, have been collected and coded. The authors intend to explore how Quantitative Ethnography (QE) techniques can be applied to analyze the data, providing deeper insights into students' peer-assessment behaviors and perceptions across two time points. This understanding can help instructors who use team-based learning better support teamwork.

Keywords: Team-based Learning, Peer Feedback, Engineering Education, CATME.

1 Area and Domain of Research

This proposed study falls within the domains of Team-Based Learning, Engineering Education, and Learning Analytics.

2 Background

Peer feedback is widely used in higher education to create a student-centered, formative approach that can improve learning outcomes and performance [1]. It enables students to actively participate in the assessment process [2], representing the "learning element"

of peer assessment. Both students and instructors benefit from peer feedback as an assessment method. Students receive multiple, time-limited perspectives compared to relying on a single instructor's comments. Interestingly, while some students believe they benefit from both giving and receiving feedback [3], research frequently suggests that students report learning more from providing feedback than from receiving it [4]. Furthermore, peer feedback fosters reflective practitioners who think cognitively and metacognitively about their performance [2]. It also develops transferable professional skills, such as teamwork and the ability to give and accept criticism, justify positions, and reject suggestions [5]

Despite the pedagogical value of peer feedback, its effectiveness is heavily influenced by student attitudes, perceptions, and beliefs. Researchers argue that a student's emotional readiness to engage, combined with their beliefs about the value of the feedback process, predicts both the accuracy of the feedback they produce and their openness to incorporating it into their own work [6]. In our earlier research on students' peer assessment using the Comprehensive Assessment of Team Member Effectiveness (CATME) system, which is a web-based tool designed to facilitate self- and peer-evaluations, primarily in higher-education teamwork settings [7], we focused exclusively on how closely students' peer assessments matched the five teamwork dimensions defined by CATME, including: Contributing to the Team, Interacting with Teammates, Keeping the Team on Track, Expecting Quality, and Having Relevant Knowledge, Skills, and Abilities. This study expands on our previous work by exploring students' attitudes, perceptions, and beliefs about peer assessment, as well as how these factors influence their giving and receiving of peer feedback. By adopting this approach, our goal is to identify strategies to enhance students' peer-assessment skills and to support instructors in facilitating effective teamwork.

3 Objectives

The purpose of this study is to understand students' attitudes, beliefs, and perceptions of peer feedback through their reflections and feedback entries. Research questions include: (1) What are the students' perceptions of the value and impact of the peer feedback they receive?; (2) How do the students describe their processes and considerations when providing feedback to their peers?; (3) What specific criteria or characteristics do the students identify as contributing to the perceived quality of peer feedback they receive via CATME?; (4) What factors or motivations do the students report as influencing the quality of peer feedback they provide through CATME?; and (5) Which CATME dimension(s) are perceived as most significant for effective teamwork, and why do students in this course believe this to be the case?. Data were collected and coded. However, we have not incorporated QE techniques to visualize and model the connections among the factors related to students' attitudes, beliefs, and perceptions of peer feedback.

4 Prior Experience and Expectations

The first author participated in the QE fellow program in 2025, during which she gained firsthand experience with ENA and developed a basic understanding of QE techniques. The other two authors have minimal knowledge of QE. Both of them are familiar with Qualitative research methods and interested in learning how QE could be integrated into narrative analyses beyond this current study.

5 References

1. Daou, D., Sabra, R., & Zgheib, N. K. Factors that determine the perceived effectiveness of peer feedback in collaborative learning: a mixed methods design. Medical Science Educator, 30(3), 1145–1156. (2020).
2. Ion, G., Sánchez Martí, A., & Agud Morell, I. Giving or receiving feedback: which is more beneficial to students' learning? Assessment and Evaluation in Higher Education, 44(1), 124–138. (2019).
3. Simpson, G., & Clifton, J. Assessing postgraduate student perceptions and measures of learning in a peer review feedback process. Assessment and Evaluation in Higher Education, 41(4), 501–514. (2016).
4. Yu, Q., & Schunn, C. D. Understanding the what and when of peer feedback benefits for performance and transfer. Computers in Human Behavior, 147, 107857. (2023).
5. Planas-Lladó, A., Feliu, L., Arbat, G., Pujol, J., Suñol, J. J., Castro, F., & Martí, C. An analysis of teamwork based on self and peer evaluation in higher education. Assessment & Evaluation in Higher Education, 46(2), 191–207. (2020).
6. Greisel, M., Hornstein, J., & Kollar, I. Do students' beliefs and orientations toward peer feedback predict peer feedback quality and perceptions? Studies in Educational Evaluation, 84, 101438. (2025).
7. Loughry, M. L., Ohland, M. W., & Woehr, D. J. Assessing teamwork skills for assurance of learning using CATME team tools. Journal of Marketing Education, 36(1), 5–19. (2014).

So much data, so many stories: Where should I start with Epistemic Network Analysis?

Wanju Huang[1][0000-0001-5965-2597]

[1] Purdue University, West Lafayette IN 47907, USA
wanjuhuang@purdue.edu

Abstract. This study aims to examine how social network analysis (SNA), used as an intervention, influences students' engagement and perceptions of online learning communities. Quantitative findings derived from SNA indicate a positive impact on students' participation in discussion activities. To strengthen the study, the author plans to incorporate qualitative data analyzed through Epistemic Network Analysis (ENA). This approach will provide a more comprehensive and nuanced understanding of students' learning experiences and the underlying dynamics of interaction within the online community.

Keywords: Online Learning, Community of Inquiry, Discussion, Social Network Analysis.

1 Area and Domain of Research

This proposed research study falls within the domains of Online Learning and Learning Analytics.

2 Background

Online discussions are regarded as impactful, affecting students' satisfaction and learning, and fundamental to developing a community of inquiry [1]. Researchers emphasize that students' interactions in discussions, through posting their initial thoughts and peer responses, reflect their learning [2], underscoring the critical role of online discussion in developing an online learning community. Even though genuine and meaningful dialogue occurs in discussion forums, studies have shown that online learners are more likely to share their own opinions or experiences (explanatory) than to engage (collaborative) with their peers or instructors in co-constructing knowledge or in critically exploring learning materials [3]. In addition, online learners feel confused and frustrated because they often get lost in multithreaded asynchronous discussions and are overwhelmed by information overload, with numerous posts and threads from peers [4]. There is a need to make the invisible learning community visible to the learners. Social Network Analysis (SNA), a "quantitative analytical technique . . . used to analyse and visualize networks" [5], has been used in online learning to illustrate how students participate in online courses, how students interact with each

other in collaborative learning settings (e.g., discussion forums), and how online learning communities evolve through time [6]. The author of this study has experimented with using SNA as an intervention to motivate online Master's students in learning design and technology to engage with the discussion board and to interact with various peers each time they post. The results of these interventions have been positive (Author, 2023) as students were posting more than what was expected, which was one initial post and two peer responses. On average, students created 4-5 posts. Additionally, the data suggested that they engaged with different peers each time they were required to participate in the discussion board. On average, students interacted with 8-9 new peers throughout the semester, which is close to 75% of the class size. In addition to the quantitative data from SNA indicating a positive response to SNA, the preliminary analysis of students' reflections at the end of the class revealed that they were able to broaden their interactions with peers from diverse professional backgrounds, such as a corporate trainer engaging with K-12 educators.

3 Objectives

This proposal explores how additional activities within the same class can improve students' perceptions of SNA, their participation in online discussions, and their view of the online learning community. Data include students' reflections submitted at semester's end, responding to two questions: 1) "How did the social networking maps influence your group interactions?" and 2) "How did these experiences shape your perception of the online learning community?" The analyzed SNA data categorize online learners into four groups: "Guide," indicating full engagement; "Facilitator," actively participating but not leading discussions; "Collaborators," engaging at medium or higher levels, meeting or exceeding expectations; and "Observer," participating minimally, just meeting basic requirements. Corresponding research questions include: 1. Using qualitative data and ENA, what learner types emerge?; and 2. How well can qualitative data be triangulated with quantitative data? In essence, do students emphasizing SNA feel motivated to interact more genuinely with peers? The qualitative data were collected in Fall 2024, Summer 2025, and Fall 2025 from over 100 online students but have not yet been fully analyzed. By the conference, the author hopes to have completed data analysis and to discuss plans for integrating qualitative data to enrich the study.

4 Prior Experience and Expectations

The author participated in the 2025 QE fellow program and has experience using ENA to analyze survey data. Nevertheless, she has not yet explored how to use ENA for data captured in online discussion forums, where conversations among participants take place. During the session, she is interested in learning more about QE analyses and exploring how researchers incorporate different datasets or research methods to tell the stories emerging from the data, visualize the data, and strengthen the study and its narratives.

5 References

1. Martin, F., Wu, T., Wan, L., & Xie, K. A meta-analysis of the Community of Inquiry presences and learning outcomes in online and blended learning environments. Online Learning, 26(1), 325–359. (2022).
2. Poquet, O., Trenholm, S., & Santolini, M. Forum posts, communication patterns, and relational structures: A multi-level view of discussions in online courses. Educational Technology Research and Development, 72(5), 2655-2678. (2024).
3. Wu, X., He, Z., Li, M., Han, Z., & Huang, C. Identifying learners' interaction patterns in an online learning community. International Journal of Environmental Research and Public Health, 19(4), 2245. (2022).
4. Faulconer, E. K., Bolch, C., & Wood, B. Cognitive load in asynchronous discussions of an online undergraduate STEM course. Journal of Research in Innovative Teaching & Learning, 16(2), 268-280. (2023).
5. Jan, S. K., Vlachopoulos, P., & Parsell, M. Social network analysis and learning communities in higher education online learning: A systematic literature review. Online Learning Journal, 23, 249-264. (2019).
6. Saqr, M., Viberg, O., & Vartiainen, H. Capturing the participation and social dimensions of computer-supported collaborative learning through social network analysis: which method and measures matter?. International Journal of Computer-Supported Collaborative Learning, 15. 227–248. (2020).

Students' Opportunities for Meaningful Inclusion as Measured by Classroom Observations and Teacher Interviews

Alexandra Shelton[1[0000-0003-3659-347X]], Rebecca A. Cruz[1[0000-0002-9163-1312]], Lindi R. Shepard[1[0009-0003-5629-575X]], Katy Mullins[1[0009-0001-5849-7967]], Irene Ramirez[1[0009-0004-7522-6079]], Rachel S. McClam[1[0000-0002-4745-6094]], and Isun Malekghassemi[1[0009-0009-7652-965X]]

[1] Johns Hopkins University, Baltimore, MD 21218
ashelt18@jh.edu

Abstract. Our study aims to examine students' opportunities for rigorous learning—a key component of meaningful inclusion—as measured by classroom observations and teacher interviews across two schools. Additionally, we would like to learn how QE can be leveraged to identify patterns in observation and interview data by classroom type. Although data collection is complete, the data have not yet been analyzed.

Keywords: Classroom Observations, Interviews, Inclusion, Special Education

1 Area and Domain Research

This study is in the domain of Inclusive Education and Special Education.

2 Background

Substantial research on inclusion points to positive academic and social impacts for students with and without disabilities (e.g., Hunt et al., 2020). However, some scholars remain unconvinced. For example, Fuchs et al. (2025) argue that placement decisions for students with disabilities whose "parents and teachers have ambitious academic expectations" (p. 11) should be based on where they can receive effective instruction, and this instruction requires "consistency, intensity, and systematicity" (p.11), which the authors argue are likely to occur outside of the general education setting. This viewpoint, however, ignores rigorous studies, such as those previously discussed, that demonstrate the academic (and social) benefits of inclusion, including research demonstrating the positive effects of intensive instruction within the general education setting (Hunt et al., 2020). It also fails to account for decades of research demonstrating that segregated settings do not, in fact, offer access to rigorous, "intensive" instruction (Causton-Theoharis et al., 2011).

Inclusion's success depends on educators' enactments of inclusive practices in everyday classrooms and schools. Yet, much of the research on inclusion investigates students with disabilities' physical presence in general education classes, which is necessary but not sufficient to realize meaningful inclusion—providing fluent and flexible means of achieving rigorous, responsive, and relevant learning experiences to all students within the general education setting. As such, we have adopted Critical Inclusion (Cruz et al., 2024) as our theoretical framework. Critical Inclusion rejects the assumptions behind schools' normative center, which generate false instantiations of inclusion (e.g., inclusion for some but not all students, and only when certain circumstances are met) in favor of meaningful inclusion.

3 Research Objectives

Our study aims to examine students' instructional access to meaningful inclusion. Our specific research questions are:

1. How do educators' practices support or hinder meaningful inclusion?
2. How do enactments of meaningful inclusion vary across classroom types?

This research pertains to general and special education teachers in general and special education settings, as well as students with and without disabilities.

To answer our research questions, we conducted observations and interviews. Though data collection is complete, data analysis has not begun.

4 Prior Experience and Expectations

As part of a multiyear project supporting inclusion in one school district, our team has used QE in two studies. In both studies, we conducted epistemic network analysis (ENA) to identify patterns in interview data. In our second study, we also employed a mixed methods approach, supplementing ENA with reflexive thematic analysis. For the current study, we have not yet begun data analysis and expect that discussions with QE scholars will inform our analytic procedures. Specifically, we hope to learn how we might apply ENA to both observation field notes and interview data. That is, we are investigating how to define and operationalize conversations and units of analysis in ways that meaningfully capture patterns of connection among ideas, practices, and participants, particularly when the data sources are different (e.g., interviews and field notes). We hope to gain more insight into how to identify meaningful discourse codes and determine the appropriate unit of analysis (e.g., utterances, turns of talk, episodes, or activity segments) for each data source within a single larger analysis.

References

1. Causton-Theoharis, J., Theoharis, G., Orsati, F., & Cosier, M. (2011). Does self-contained special education deliver on its promises? A critical inquiry into research and practice. *Journal of Special Education Leadership*, 24(2), 61–78.
2. Cruz, R. A., Firestone, A. R., & Love, M. (2024). Beyond a seat at the table: Imagining educational equity through critical inclusion. *Educational Review*, 76(1), 69-95. https://doi.org/10.1080/00131911.2023.2173726
3. Fuchs, D., Gilmour, A. F., & Wanzek, J. (2025). Reframing the most important special education policy debate in 50 years: How versus where to educate students with disabilities in America's schools. *Journal of Learning Disabilities*, 58(4), 257–273. https://doi.org/10.1177/00222194251315196
4. Hunt, P., Kozleski, E., Lee, J., Mortier, K., Fleming, D., Hicks, T., Balasubramanian, L., Leu, G., Bross, L. A., Munandar, V., Dunlap, K., Stepaniuk, I., Aramburo, C., & Oh, Y. (2020). Implementing comprehensive literacy instruction for students with severe disabilities in general education classrooms. *Exceptional Children*, 86(3), 330–347. https://doi.org/10.1177/0014402919880156

Understanding How the Physics Community Integrates Mathematical Methods in Undergraduate Curricula

Aiswarya P. Amalucottage[1][0000−0002−9673−8186], Lyra D. Moscuzza[2][0009−0009−3039−9650], and Geraldine L. Cochran[1][0000−0002−8520−6168]

[1] The Ohio State University, Columbus OH 43210, USA
[2] University of Washington, Seattle WA 98195, USA
amalucottage.1@osu.edu

Abstract. Mathematics plays a central role in undergraduate physics education, yet institutions differ in how mathematical methods are incorporated into their physics curricula. This study aims to characterize how physics departments integrate mathematical methods into undergraduate programs across institutions. Because a department's approach to mathematical preparation reflects the practices and shared understandings of the physics community, characterizing these curricula reveals how the community approaches and reasons about the role of mathematics in the physics major. Using a database of publicly available course and institutional information, together with faculty interviews, we will examine features such as course structure, prerequisites, sequencing, mathematical content, and connections to other physics courses. We are interested in identifying patterns across institutions and exploring how Quantitative Ethnography methods could support the analysis and visualization of relationships among these curricular features.

Keywords: Physics Education Research · Mathematical Methods · Undergraduate Physics · Curriculum · Quantitative Ethnography.

1 Area and Domain of Research

This proposed research is situated within Physics Education Research, with a particular focus on undergraduate physics curriculum and curricular structures.

2 Background

Mathematics is an essential component of undergraduate physics education, and students are expected to use increasingly sophisticated mathematical tools as they progress through the physics curriculum [2, 4]. However, physics departments may organize and support the development of these mathematical skills in different ways. Some programs offer dedicated mathematical methods courses, others integrate mathematical methods within physics courses, still others rely on

mathematics-department coursework [1]. Understanding these curricular structures is important because the organization and sequencing of mathematical preparation may shape how students encounter and use mathematics throughout the physics major [4]. A systematic comparison across institutions can reveal common and distinct approaches to integrating mathematical methods and provide a broader evidence base for curriculum development and redesign, including efforts to better align mathematical methods with physics content.

3 Research Objectives

The primary aim of this study is to understand how physics departments integrate mathematical methods into their undergraduate physics curriculum. Potential research questions include:

- RQ1: How do institutions differ in the ways mathematical methods are incorporated into undergraduate physics curricula?
- RQ2: What patterns of curricular features characterize mathematical methods courses across institutions and institution types?
- RQ3: How are features such as prerequisites, course sequencing, mathematical content, and connections to other physics courses related across institutions and institution types?

The study focuses on undergraduate physics programs across multiple institutions. Data are being compiled from publicly available institutional and departmental information, including course catalogs, departmental websites, curricula, course descriptions, faculty interviews, and other available course materials. The project is currently in the database development and exploratory analysis stage. At their core, our research questions are relational: they ask how curricular features connect and how those connections differ across institutions and institution types. That is what draws us to QE, an approach designed to model the relationships among coded features rather than the features in isolation [3]. We are interested in exploring whether QE techniques can help identify relationships among curricular features and distinguish different approaches to integrating mathematical methods within undergraduate physics programs.

4 Prior Experience and Expectations

Aiswarya: My prior experience with Quantitative Ethnography has primarily been through collaborative writing on a paper about the role of theory throughout the QE research process. Through that work, I became familiar with some of the methodological considerations involved in QE. I would like suggestions for using AI tools to expand or validate the database. **Lyra:** I am an undergraduate student supporting this project. I have not completed a QE study before, but I look forward to working on this QE project. **Geraldine:** I am a physics education researcher new to QE.

References

1. Caballero, M.D., Doughty, L., Turnbull, A.M., Pepper, R.E., Pollock, S.J.: Assessing learning outcomes in middle-division classical mechanics: The Colorado classical mechanics and math methods instrument. Physical Review Physics Education Research **13**(1), 010118 (2017). https://doi.org/10.1103/PhysRevPhysEducRes.13.010118
2. Redish, E.F., Kuo, E.: Language of physics, language of math: Disciplinary culture and dynamic epistemology. Science & Education **24**(5), 561–590 (2015). https://doi.org/10.1007/s11191-015-9749-7
3. Shaffer, D.W., Ruis, A.R.: Models all the way down: Discourse and other models in quantitative ethnography. In: International Conference on Quantitative Ethnography. pp. 3–20. Springer (2026). https://doi.org/10.1007/978-3-032-12229-2_1
4. Wilcox, B.R., Caballero, M.D., Rehn, D.A., Pollock, S.J.: Analytic framework for students' use of mathematics in upper-division physics. Physical Review Special Topics—Physics Education Research **9**(2), 020119 (2013). https://doi.org/10.1103/PhysRevSTPER.9.020119

A Workshop on Writing in Quantitative Ethnography

Zachari Swiecki[1], Andrew Ruis[2], and David Williamson Shaffer[2]

[1] Monash University, Clayton VIC 3800, Australia
zach.swiecki@monash.edu
[2] University of Wisconsin–Madison, Madison WI 53706, USA

Abstract. Quantitative Ethnography (QE) integrates qualitative and quantitative evidence to support grounded claims about data and its context. Writing up QE work is non-trivial: authors must combine the conventions of qualitative reporting, the conventions of quantitative reporting, and a set of practices specific to QE itself, all in service of a single coherent argument. Existing reporting guidelines address the qualitative and quantitative traditions separately, but say little about the integration challenge that defines a QE write-up — and QE papers are increasingly reviewed by readers who are not steeped in QE conventions. This three-hour workshop walks participants through each section of a QE paper in turn, from methods and results through theory, discussion, and the introduction and conclusion, pairing short talks on conventions with hands-on drafting activities. Facilitators supply example data and results, so participants can take part with or without a draft in hand. It is aimed at graduate students and early-career researchers, though open to anyone curious about QE writing. Participants will leave with a clearer mental model of the QE writing genre, new peer connections, and feedback from facilitators. A post-workshop writing group will extend the conversations and relationships beyond the conference itself.

Keywords: Quantitative Ethnography · Academic Writing · Workshop · Research Methods.

1 Focus and Benefits

Quantitative Ethnography (QE) integrates qualitative and quantitative evidence to support grounded claims about data and its context [3]. Write-ups in academic papers are the primary medium for conveying this evidence. Well-developed guidelines exist for what strong write-ups look like in qualitative [2] and quantitative [1] traditions. However, neither of these guidelines speak to the integration challenge that defines a QE write-up. Weaving qualitative and quantitative strands into a single, coherent argument is non-trivial, and we routinely see strong QE projects whose contributions are undersold because the write-up fails to do this work.

A secondary concern is reception. QE papers are increasingly submitted to venues whose reviewers are not steeped in QE conventions. Authors who can frame their methodological choices clearly, anticipate reviewer questions, and present qualitative and quantitative results in a unified narrative have a substantially better chance of seeing their work published in mainstream outlets. Helping authors write more effectively is therefore both a craft issue and a pipeline issue for the field.

This workshop addresses both. Over three hours, participants work through each section of a QE write-up in turn, learning the conventions, drafting short pieces of text, and receiving feedback. The workshop is designed and delivered by three facilitators with deep ties to the QE community. Dr. Swiecki has worked within QE for over thirteen years and has received several best paper awards at the ICQE conference. Dr. Ruis has been involved in QE since its inception and is a co-author of many of its foundational papers. Prof. Shaffer is the founder of QE. Together, the facilitators bring complementary perspectives on what makes a QE paper work methodologically, rhetorically, and as a contribution to the field.

2 Target Audience

The workshop is open to anyone curious about the conventions of QE writing, with particular focus on graduate students and early-career researchers. It is designed to be useful at any stage of a project: participants with a draft in hand will get the most out of the drafting activities, but those with less well-formed projects are equally welcome. The facilitators will provide example data and results that participants can use for the activities. Two secondary audiences are also welcome: reviewers who want to better evaluate QE submissions for mainstream venues, and established researchers who mentor students through QE write-ups.

Prior experience with QE methods is preferred but not required. The workshop will not teach QE methodology in detail; rather, it focuses on how to write up work that uses QE. We anticipate that the workshop will adapt to the experience mix of registered participants, with more scaffolding for newer researchers and more peer-feedback time for those with drafts.

3 Schedule of Activities

The schedule in Table 1 is a tentative three-hour plan. Specifics will be finalised closer to the conference once we have a clearer picture of participant numbers and prior QE experience; the activities and timing may be adjusted, but the overall arc — moving through each section of a QE paper with paired talk-and-activity blocks — will remain the same.

Table 1. Schedule of activities.

Time	Section	Description
15 min	Overview	*Talk:* Brief facilitator and participant introductions, an overview of what makes QE writing a distinct genre, and a walk-through of the conventional empirical paper structure. We distinguish between drafting order (the order in which we recommend authors compose sections) and presentation order (the order in which sections appear in the final paper).
30 min	Methods	*Talk:* What belongs in a QE methods section, including codebook development, the coding process, and modelling decisions. *Activity:* Participants draft methods text describing a codebook, coding process, and model, using example data provided by the facilitators or their own project if available.
30 min	Results	*Talk:* How to write up qualitative and quantitative findings as a single, coherent story, including the QE conventions of packing (summarising a qualitative claim) and unpacking (walking the reader through a qualitative example that illustrates the claim). *Activity:* Participants draft a results passage that packs and unpacks a qualitative claim and corroborates it with quantitative evidence.
15 min	Break	
30 min	Theory	*Talk:* Theory as an argument: framing prior work to motivate the gap the study addresses and arguing for QE as the appropriate methodology. *Activity:* Participants draft a theory-to-gap outline that moves from existing theory to the gap their study addresses, plus a short argument for using QE in their study.
30 min	Discussion	*Talk:* Recapping results, relating them back to theory, and stating limitations honestly. *Activity:* Participants draft a results-to-theory outline that returns the findings to the theoretical contribution and list several limitations relevant to their study.
20 min	Intro / Conclusion	*Talk:* The problem, promise, and payoff structure for introductions and conclusions in QE papers. *Activity:* Participants draft a three-sentence funnel introduction and a two-sentence conclusion that echoes the introduction's promise and states the payoff.
10 min	Closing	Final questions, comments, and reflections, plus information on how to join the post-workshop writing group.

4 Expected Outcomes

Participants will leave the workshop with three things. First, a clearer mental model of QE writing as a distinct genre — what each section of the write-up does, how qualitative and quantitative evidence fit together, and what reviewers tend to look for. Second, new peer connections with other QE authors at a similar career stage, formed through the small group drafting activities. Third, direct feedback from the facilitators on the short pieces of text drafted during the workshop.

From the facilitators, participants will receive a slide deck covering the conventions discussed in each block, and (where rights permit) a shared drive of example QE papers and write-ups that illustrate the techniques in published work.

Beyond the conference, we plan to establish an informal QE writing group open to workshop participants and other interested members of the QE community. The group will provide a low-stakes space to share drafts, ask questions, and continue the conversations begun at the workshop. The relationships and materials from the workshop are intended to persist well beyond the three-hour session itself.

Longer term, we see this workshop contributing to three goals for the QE community: building writing capacity within the field, helping more QE work reach publication in mainstream venues, and supporting the next generation of QE authors as they develop their voices. We expect to refine and re-run the workshop in future years, incorporating what we learn from this iteration.

Disclosure of Interests. The authors have no competing interests to declare that are relevant to the content of this article.

References

1. Appelbaum, M., Cooper, H., Kline, R.B., Mayo-Wilson, E., Nezu, A.M., Rao, S.M.: Journal article reporting standards for quantitative research in psychology: the APA Publications and Communications Board task force report. American Psychologist **73**(1), 3–25 (2018). https://doi.org/10.1037/amp0000191
2. Levitt, H.M., Bamberg, M., Creswell, J.W., Frost, D.M., Josselson, R., Suárez-Orozco, C.: Journal article reporting standards for qualitative primary, qualitative meta-analytic, and mixed methods research in psychology: the APA Publications and Communications Board task force report. American Psychologist **73**(1), 26–46 (2018). https://doi.org/10.1037/amp0000151
3. Shaffer, D.W.: Quantitative Ethnography. Cathcart Press, Madison, WI (2017)

DIY QE: Demonstrating, Building, and Sustaining Community-Developed Tools for Quantitative Ethnography

Golnaz Arastoopour Irgens[1[0000-1111-2222-3333]], Amanda Barany[2[1111-2222-3333-4444]], Luc Paquette[3[1111-2222-3333-4444]], Andrew Ruis[4[1111-2222-3333-4444]], Yuanru Tan[4[1111-2222-3333-4444]], Zhiqiang Cai[4[1111-2222-3333-4444]] and Cody Marquart[4]
[1] Vanderbilt University, Nashville, TN 37203, USA
[2] Penn Center for Learning Analytics, University of Pennsylvania, Philadelphia, PA 19104, USA
[3] University of Illinois, Champaign, IL 61820, USA
[4] Wisconsin Center for Education Research, University of Wisconsin–Madison, Madison, WI 53706, USA
golnaz.arastoopour.irgens@vanderbilt.edu

Abstract. This interactive workshop responds to a growing challenge within the Quantitative Ethnography (QE) community: while researchers continue to develop innovative tools that extend QE processes, these tools remain scattered across personal repositories, labs, and informal collaborations with little shared infrastructure for dissemination, validation, or stewardship. Developments in AI and large language models have also accelerated tool creation, which offer possibilities for coding, analysis, and data collection while raising important questions about validity, transparency, reproducibility, and alignment with QE's epistemological commitments. Through demonstrations, collaborative design activities, and community discussion, participants will explore emerging QE tools, begin planning or prototyping tools of their own, and contribute to conversations about how the community can collectively organize and sustain open methods infrastructure for QE practice.

Keywords: AI-Supported Methods, Methodological Infrastructure, Community-Developed Tools.

1 Introduction

The Quantitative Ethnography (QE) community has built a core set of methodological tools that includes Epistemic Network Analysis (ENA) through the webtool [4] and rENA, the active-learning classifier nCoder [7], and the related Ordered Network Analysis (ONA) and Transmodal Analysis (TMA) packages [8, 9]. Beyond this core, an expanding set of tools sits scattered across labs and personal repositories, including lightweight Shiny applications that convert ENA, ONA, and TMA outputs for non-R users; large language model (LLM) classifiers that extend nCoder workflows; specialized chatbots; real-time analysis tools that generate interview prompts during data collection; and connectors that route data among platforms. Each tool addresses a

concrete issue that arises when researchers apply QE in different contexts. Activity around nCoder alone illustrates the pattern with successive ICQE proceedings introducing semantic extensions [2], neural network alternatives [3], and LLM-based benchmarks [10], each surfacing design decisions and tools that other QE researchers would benefit from learning about directly. However, these tool innovations remain under-coordinated. News travels by email conversations (if at all), documentation is uneven, and the community has not yet agreed on how to evaluate, share, or steward what gets built.

This DIY (Do It Yourself) workshop convenes the community around three goals. First, it surfaces tools that QE members have built beyond the central pipeline, with explicit attention to the QE problem each tool addresses and the design choices that follow from QE's epistemological commitments. Second, it walks participants through a structured process for scoping and beginning a tool of their own. Third, it opens a discussion about how to organize collective work on QE tools toward open, reproducible, and equitable practice. The workshop responds to a moment in which generative AI is reshaping qualitative research workflows. Recent comparative work [10] shows that LLMs can support coding in ways nCoder cannot, while raising new questions about validity, reproducibility, responsible AI use, and the kinds of meaning QE seeks to recover from discourse [1,6]. The QE community needs venues to work through these questions through direct engagement with the technology, and to do so in ways that hold AI accountable to QE theory rather than the other way around.

2 Target Audience

Workshop participants will include doctoral students, faculty, and practitioner-researchers across QE's contexts of use, including education, healthcare, policy, organizational learning, and human-computer interaction. We expect a mix of experience levels: participants who have an idea for a tool but have not yet started building, participants who have working prototypes and want feedback, participants who collaborate with technical teams and need a shared design vocabulary, and participants curious about the AI integrations entering QE practice. No programming experience is required. Participants should bring a laptop and are strongly encouraged to bring one QE problem from their own work that a new tool or technique might help solve.

3 Schedule of Activities

The workshop runs two hours and has three parts: showcase, build, and organize.

Welcome and Framing. Organizers introduce the workshop, summarize the landscape of QE tools beyond the central pipeline, and frame the design questions and AI considerations that motivate the day's activities. Participants respond to a digital poll describing the QE problem they brought, if they did so.

Tool Showcase. Four short demonstrations of 8 to 10 minutes each feature tools at different stages of maturity and QE analysis pipeline usage. Each presenter will articulate (a) the QE problem the tool addresses, (b) design choices and AI components where relevant, and (c) what it would take for others to use, extend, or contribute to the tool. Confirmed and proposed demonstrations include:

Better Call Print (Yuanru Tan), a lightweight R Shiny application that allows non-R users to extract ENA, ONA, and TMA model output (line weights, node coordinates, and related model details) from RData files exported by the webtool, addressing a practical reporting bottleneck for applied QE researchers.
Code It! (Golnaz Arastoopour Irgens), a re-design of nCoder and Codey that incorporates regular expression-based coding and perfect sampling validation methods offering more flexible workflows than previously designed. It provides an example of "re-mixing" already developed tools for local purposes.
LLMs for QE Coding (Golnaz Arastoopour Irgens & Amanda Barany) will discuss two tools designed using LLMs to code discourse data for QE analysis purposes. These tools explore solutions for leveraging the power of human-LLM partnerships for qualitative coding while centering QE values around fairness, validity, and transparency.
Quick Red Fox (Luc Paquette & Amanda Barany), a data collection tool that connects real-time analysis of interaction log data to identify moments for in-the-moment qualitative interviews, helping with rich QE discourse data collection [5]. The tool is currently deployed in Minecraft-based learning activities and designed to support discourse data collection that can later be analyzed using QE methods.

Tool Planning Sprint. Working in small groups organized by shared problems, participants move through a structured plan: (1) name a QE problem from their own work; (2) sketch or build a prototype of a tool that addresses it, including data inputs, outputs, and intended users; (3) identify dependencies on existing QE tools (the ENA webtool, rENA, nCoder, ROCK, ONA, and TMA); (4) draft a short plan that includes a validation strategy and openness commitments. Organizers rotate among groups as consultants. Each group presents a 2-minute pitch and demo to the larger group.

Community Organization Discussion. Participants discuss governance, hosting, and stewardship: where QE tools should live; how they should be reviewed and validated; what role ISQE can play; and how development can prioritize reproducibility and access. Outcomes are captured on a shared document for follow-up and will include a draft list of community priorities and list of volunteers for particular tasks.

4 Expected Outcomes

Participants will leave with a one-page tool plan they began drafting during the sprint, an initial prototype of their tool, a curated resource list of QE-relevant tools and repositories, and a named connection to at least one collaborator working on a similar

problem. Organizers will compile workshop notes and the planning artifacts into a public document hosted alongside ISQE materials and will seed a collection for QE tools and methodological infrastructure. Follow-up activities extend the workshop beyond the conference. We propose a quarterly DIY QE virtual meet-up where participants share progress, request feedback, and report on tools moving from prototype to use. We also propose a co-authored symposium for a subsequent ICQE that articulates principles for community-stewarded QE tools, including those that integrate AI components. Where appropriate, we will pursue connections to ISQE governance to formalize a *Tools Working Group* with shared documentation, validation, and dissemination practices.

References

1. Arastoopour Irgens, G., Eagan, B.: The foundations and fundamentals of quantitative ethnography. In: Damşa, C., Barany, A. (eds.) Advances in Quantitative Ethnography, ICQE 2022. CCIS, vol. 1785, pp. 3–16. Springer, Cham (2023). https://doi.org/10.1007/978-3-031-31726-2_1
2. Cai, Z., Siebert-Evenstone, A., Eagan, B., Shaffer, D.W., Hu, X., Graesser, A.C.: nCoder+: a semantic tool for improving recall of nCoder coding. In: Eagan, B., Misfeldt, M., Siebert-Evenstone, A. (eds.) Advances in Quantitative Ethnography, ICQE 2019. CCIS, vol. 1112, pp. 41–54. Springer, Cham (2019). https://doi.org/10.1007/978-3-030-33232-7_4
3. Choi, J., Ruis, A.R., Cai, Z., Eagan, B., Shaffer, D.W.: Does active learning reduce human coding? A systematic comparison of neural network with nCoder. In: Damşa, C., Barany, A. (eds.) Advances in Quantitative Ethnography, ICQE 2022. CCIS, vol. 1785, pp. 30–42. Springer, Cham (2023). https://doi.org/10.1007/978-3-031-31726-2_3
4. Marquart, C.L., Hinojosa, C., Swiecki, Z., Shaffer, D.W.: Epistemic Network Analysis (Version 1.7.0) [Software]. http://www.epistemicnetwork.org (2018)
5. Ocumpaugh, J., Paquette, L., Baker, R.S., Barany, A., Ginger, J., Casano, N., Zambrano, A.F., Liu, X., Wei, Z., Zhou, Y., Liu, Q., Hutt, S., Andres, A.M.A., Nasiar, N., Giordano, C., van Velsen, M., Mogessi, M.: The Quick Red Fox gets the best Data Driven Classroom Interviews: A manual for an interview app and its associated methodology. arXiv:2511.13466 (2025).
6. Shaffer, D.W.: Quantitative Ethnography. Cathcart Press, Madison, WI (2017)
7. Shaffer, D.W., Ruis, A.R.: How we code. In: Ruis, A.R., Lee, S.B. (eds.) Advances in Quantitative Ethnography, ICQE 2020. CCIS, vol. 1312, pp. 62–77. Springer, Cham (2021). https://doi.org/10.1007/978-3-030-67788-6_5
8. Shaffer, D.W., Wang, Y., Ruis, A.R.: Transmodal Analysis. Journal of Learning Analytics 12(1), 271–292 (2025).
9. Tan, Y., Swiecki, Z., Ruis, A.R., Shaffer, D.W.: Epistemic network analysis and ordered network analysis in learning analytics. In: Saqr, M., López-Pernas, S. (eds.) Learning Analytics Methods and Tutorials: A Practical Guide Using R, pp. 569–636. Springer, Cham (2024). https://doi.org/10.1007/978-3-031-54464-4_18
10. Zambrano, A.F., Liu, X., Barany, A., Baker, R.S., Kim, J., Nasiar, N.: From nCoder to ChatGPT: from automated coding to refining human coding. In: Arastoopour Irgens, G., Knight, S. (eds.) Advances in Quantitative Ethnography, ICQE 2023. CCIS, vol. 1895, pp. 469–484. Springer, Cham (2023). https://doi.org/10.1007/978-3-031-47014-1_32

Learning SCAT as a Qualitative Data Analysis Method: A Hands-on Workshop from Coding to Theory Writing

Takashi Otani[1][0009-0003-7708-6762] and Daisuke Kaneko[2][0000-0002-4669-4138]

[1] Nagoya University of Economics, 61-1, Uchikubo, Inuyama, Aichi 484-8504, Japan
[2] Hokusei Gakuen University, 2-3-1, Oyachi-Nishi, Atsubetsu, Sapporo, 004-8631, Japan
otani-t@nue.ac.jp

Abstract. In Quantitative Ethnography (QE), coding is methodologically important because codes mediate between qualitative data and quantitative modeling. Since codes are not merely labels but reflect researchers' interpretations of data, QE researchers can benefit from understanding how codes are generated in interpretive qualitative data analysis. SCAT (Steps for Coding and Theorization) is an explicit, formal, and procedural method for qualitative data analysis in which analysts segment qualitative text, conduct four-step coding, write storylines and generate theoretical statements through theory writing. This workshop introduces SCAT as a qualitative data analysis method, rather than as a code-generation tool for QE. In SCAT, codes and storylines are intermediate products, and its primary purpose is theory writing. Through hands-on analysis of a short qualitative dataset, participants will learn the basic principles and procedures of SCAT and critically reconsider the relationship between coding, interpretation, and theorization in relation to QE.

Keywords: SCAT, Qualitative Research, Qualitative Data Analysis, Coding, Theory Writing, Quantitative Ethnography.

1 Introduction

Quantitative Ethnography (QE) has developed powerful approaches for modeling discourse, interaction, and meaning-making through coded qualitative data. In QE research, therefore, the process of generating and applying codes is methodologically crucial. Codes are not merely labels attached to text data; they reflect researchers' awareness and interpretation of the data. For this reason, QE researchers can benefit from a deeper understanding of how qualitative codes are produced within interpretive qualitative data analysis.

SCAT (Steps for Coding and Theorization) offers a useful opportunity for such reflection. SCAT is an explicit, formal, and procedural method for qualitative data analysis, not a qualitative research method [1, 2]. In SCAT, segmented text data are entered into a matrix and coded through four steps; the final codes are then connected into

storylines, which are used for theory writing. SCAT is suitable for the analysis of relatively small-scale datasets. It can also be used to analyze open-ended questionnaire responses [3].

Importantly, this workshop clarifies that SCAT should not be understood merely as a tool for generating codes. In SCAT, codes and storylines are intermediate products in an analytic process whose primary aim is theory writing.

2 Overview of SCAT

In practice, SCAT follows these steps [4]: a) segmenting the text and entering each segment into the text column; b) coding each segment through Steps <1> to <4> using nouns; c) writing a storyline by connecting the terms in Step <4> into a coherent narrative in the past tense; and d) writing theoretical statements based on the storyline in the present tense.

The four coding steps are as follows.

- <1> Extract noteworthy words or phrases from the text. Read the text carefully and extract important words or phrases directly from the original text.
- <2> Paraphrase the extracted words and phrases in Step <1> using terms from outside the original text. One extracted phrase may have multiple paraphrases, or several extracted phrases may be summarized by one paraphrase.
- <3> Put extra-textual concepts that account for Step <2>. It is useful to consider background, conditions, causes, results, effects, comparisons, characteristics, dimensions, and changes. It is also useful to look vertically across multiple rows in the analysis sheet, rather than only horizontally within the same row.
- <4> Generate themes and constructs emerging from Steps <1> to <3> while considering the context. Read Steps <1> to <3> carefully and generate new themes or constructs. These may be newly coined concepts, concise thematic terms, or expressions adapted from the analyst's own words. The terms should be concise and memorable because they may later be applied to other contexts or practices.

After coding, the storyline is a coherent description of the latent meaning and significance of the events described in the data, written mainly by weaving together the themes and constructs from Step <4>.

To write the storyline, connect the Step <4> themes in a meaningful sequence. Consider the relationships among themes: whether they should be linked by "and," "or," "therefore," "however," or another logical connection. The storyline is not a simple summary; it should reveal the deeper context and meaning of the data.

The coding process from Steps <1> to <4> de-contextualizes parts of the data by focusing on segments and concepts. Writing the storyline re-contextualizes those concepts by connecting them into a deeper narrative context.

After writing the storyline, write theoretical statements by breaking the storyline into concise points. In SCAT, "theory" does not mean a universal law. It means what can be said from this data. The statements should remain grounded in the analyzed data while expressing the analytic insights obtained from the storyline.

3 What Makes SCAT Distinct

This method is accessible to beginners in qualitative data analysis and is also used by practitioners and researchers in science and medicine outside the humanities and social sciences, as well as by quantitative researchers. This is because SCAT fully visualizes the analytic process, resulting in high visibility and reflectability of the analysis, which makes validation easier. Furthermore, this transparency not only enables collaborative analysis but also makes SCAT an analytic method with excellent falsifiability. Moreover, SCAT performs de-contextualization of the text through the four coding steps, and re-contextualization through storyline writing. Despite being a compact procedure, SCAT has these capabilities and thereby enables analysts to extract deep, latent meanings from the text.

SCAT is used internationally. As of August 30, 2026, 111 English-language articles in international academic journals across various disciplines had used this method. Including Japanese-language papers, the total reaches 1,187; these include doctoral and master's theses, as well as books written using SCAT. Furthermore, studies examining the usefulness of SCAT in QE have been presented several times at ICQE [5, 6, 7, 8].

4 Target Audience for the Workshop

The target audience includes researchers who are interested in the qualitative dimensions of QE and who would like to understand contemporary qualitative data analysis methods beyond QE. Researchers who have encountered SCAT in ICQE presentations and would like to learn more about it are also part of the target audience.

SCAT is appropriate for interpretivist qualitative data analysis and does not include quantitative procedures, so it can be regarded as contrasting with QE. However, researchers who conduct QE can benefit from experiencing interpretivist qualitative data analysis, because doing so helps them gain a more comprehensive understanding of qualitative analysis in general.

It should be noted that SCAT aims at theory writing; therefore, it should not be used merely as a means of generating codes for QE. In particular, SCAT does not proceed in a single direction, and Step <4> themes and constructs change during writing storyline and theoretical statements. For this reason, it is not appropriate to halt the analysis at Step <4> and use Step <4> terms for other analyses, such as QE, before completing storyline and theory writing. Nevertheless, after completing these processes, the possibility of combining SCAT and QE should continue to be explored.

5 Schedule of Activities

The workshop will begin with an introduction to the basic structure, steps, theoretical orientation, and typical applications of SCAT. Participants will then apply SCAT to a short simulated dataset, working through the process from segmentation and four-step

coding to storyline writing and theory writing. Finally, participants will share and discuss their analytical results, including their storylines and theoretical statements, and receive feedback from the facilitators.

This workshop will be facilitated by a team that includes the developer of SCAT, who is also scheduled to participate in a keynote session at ICQE26.

6 Expected Outcomes of the Workshop

Participants will be able to explain the basic principles and procedures of SCAT as a qualitative data analysis method. They will also be able to apply SCAT to a short qualitative dataset and consider how it may be used in their own research or practice. Through this process, they will gain practical knowledge of interpretive qualitative data analysis.

By gaining a more comprehensive perspective on qualitative data analysis beyond QE, participants will be able to contextualize QE and situate it within a diverse range of analytic approaches.

Acknowledgments. This work was supported by JSPS KAKENHI (JP24K15212 and JP25K13403).

References

1. Otani, T.: Paradigm and Design of Qualitative Study: From Research Methodology to SCAT. The University of Nagoya Press, Nagoya, Japan (2019) (in Japanese)
2. SCAT: Steps for Coding and Theorization, https://www.educa.nagoya-u.ac.jp/~otani/scat/index-e.html, last accessed 2026/8/30. (The official SCAT website is currently being restructured. Interim materials are available at: https://note.com/scatty)
3. Hida, T., Fukushi, M., Nago, N., Otani, T.: How can the free descriptions of questionnaire including quite short ones be analyzed by using SCAT? Bulletin of Ichinomiya Kenshin College **5**, pp. 1–10 (2026) (in Japanese)
4. Goto, A.: Instructions for SCAT. https://note.com/api/v2/attachments/download/044fcefd850f666bb7f83b225d9d5b29, last accessed 2026/8/30
5. Kaneko, D., Ohsaki, A.: An Approach to Quantitative Ethnography with SCAT: The Use of Qualitative Data Analysis Method. In: Fifth International Conference on Quantitative Ethnography: Conference Proceedings Supplement, pp. 61–62 (2023)
6. Kaneko, D., Ohsaki, A.: Improving the Transparency of Code Generation in ENA by Integrating SCAT. In: Conference Proceedings Supplement of the Sixth International Conference on Quantitative Ethnography, pp. 90–94 (2024)
7. Ohsaki, A., Kaneko, D.: Computer-Supported Code Discovery Utilizing Topic Modeling and Stepwise Coding. In: Kim, Y. J., Swiecki, Z. (eds.) Advances in Quantitative Ethnography. Springer Nature Switzerland, pp. 58–73 (2024)
8. Ohsaki, A., Kaneko, D.: Computer-Assisted Code Generation Using Combination of Generative Artificial Intelligence, Stepwise Coding, and Topic Modeling. In: Carmona, G., Lima, C., Santos, M. J., Benítez, H., Montero-Moguel, L., et al. (eds.) Advances in Quantitative Ethnography. Springer Nature Switzerland, pp. 302–317 (2026)

Modeling Complex Learning Interactions with HINA: A Hands-on Workshop on Heterogeneous Interaction Network Analysis

Shihui Feng[1], Baiyue He[1], Alec Kirkley[1], and Dragan Gasevic[1]

[1] Hong Kong University, Pok Fu Lam, Hong Kong
shihuife@hku.hk

Abstract. This workshop introduces HINA (Heterogeneous Interaction Network Analysis), a novel, open-source analytical framework designed to capture the heterogeneous, multi-entity nature of interactions in contemporary learning environments. HINA complements and extends the QE toolkit in several meaningful ways. Where ENA captures the structure of discourse by modeling connections among coded concepts, HINA captures the structure of participation by modeling connections among diverse actors and artifacts. Together, these approaches enable richer descriptions of complex learning processes. Participants will leave with both practical skills and a deeper understanding of how HINA can expand the methodological repertoire of quantitative ethnographers, enabling them to ask new questions about the distributed, multi-faceted nature of learning interactions in contemporary educational environments.

Keywords: Heterogeneous Interaction Network Analysis, Learning Analytics, Learning Processes

1 Workshop Focus and Value to the QE Community

This workshop introduces HINA (Heterogeneous Interaction Network Analysis), a novel, open-source analytical framework designed to capture the heterogeneous, multi-entity nature of interactions in contemporary learning environments (Feng et al., 2025a). Grounded in network science and guided by social constructivism learning theory (Vygotsky, 1978), HINA formalizes multi-entity learning interactions as Heterogeneous Interaction Networks (HINs), where nodes represent distinct entity types—such as students, coded behaviors, subtasks, or AI agents—and edges encode relationships such as engagement or co-occurrence (Feng et al., 2024; Feng et al., 2025c). HINA provides three levels of original analytical methods: (1) node-level measures of interaction quantity and diversity; (2) dyadic-level statistical testing for significant edges using configurable null models; and (3) meso-level, non-parametric clustering of nodes by interaction patterns. These analytical methods are accessible through the HINA web tool (hina-network.com), an openly available platform for conducting analysis and interactive visualization (Feng et al., 2025b).

HINA complements and extends the QE toolkit in several meaningful ways. Where ENA captures the structure of discourse by modeling connections among coded concepts (Shaffer, et al., 2016), HINA captures the structure of participation by modeling connections among diverse actors and artifacts. Together, these approaches enable richer descriptions of complex learning processes. Moreover, HINA's statistical grounding and parameter-free clustering methods align well with QE's commitment to combining interpretive depth with quantitative rigor (Shaffer, et al., 2016). With growing interest in human-AI collaboration in learning contexts, HINA is timely, as it provides principled methods for analyzing how learners interact with AI agents, quantifying their contributions, and detecting emergent engagement patterns to support new theoretical developments.

Participants will leave with both practical skills and a deeper understanding of how HINA can expand the methodological repertoire of quantitative ethnographers, enabling them to ask new questions about the distributed, multi-faceted nature of learning interactions in contemporary educational environments.

2 Target Audience

This workshop is designed for a broad audience within the QE community, including:

- Researchers who work with learning process data (e.g., log files, behavioral codes, multimodal traces, AI interaction records) and seek new analytical methods to model complex, multi-entity interactions.
- Doctoral students and early-career researchers looking to develop expertise in network-analytic approaches applied to learning data.

No prior experience with network analysis or programming is required. The HINA web tool (hina-network.com) is accessible through any standard web browser and requires no installation. Participants are encouraged to bring their own datasets, though curated sample datasets will be provided.

3 Schedule of Activities

Time	Activity & Description
9:00–10:00	**Welcome & Foundations of HINA** Introduce HINA's methodological foundations. This session situates HINA within the QE community's tradition of connecting rich interaction data to meaningful qualitative interpretation. Participants will understand how Heterogeneous Interaction Networks (HINs) model interactions across diverse entity types (learners, behaviors, artifacts, AI agents).
10:00–10:30	**Hands-on Tutorial: HINA Web Tool** A guided, step-by-step tutorial using the HINA web tool (hina-network.com). Participants will prepare a sample dataset, build a heterogeneous interaction network, and explore HINA's three analytical levels.
10:30–10:45	***Break***
10:45–11:30	**Analyze Your Data & Group Discussion** Participants work with their own datasets (or provided sample data) using the HINA web tool, supported by facilitators. This is followed by a structured group discussion on: (1) new questions that can be addressed using HINA; (2) how HINA can be integrated with existing QE methods such as ENA and ONA. The session closes with a collective reflection on methodological synergies between HINA and quantitative ethnography
11:30—12:00	**Collective Sharing & Summary** Sharing across groups. Each group briefly presents key insights from their discussion. Facilitators synthesize emerging themes and highlight opportunities for continued collaboration.

Logistics: *Participants should bring a laptop with a stable internet connection. No software installation is required for the web-tool track.*

4 Expected Outcomes and Follow-up Activities

By the end of the workshop, participants will:

- Understand the methods and applications of HINA.
- Be able to use the HINA web tool and interpret multi-level analytical outputs.
- Have applied HINA to their own or sample datasets, generating preliminary insights and analytical outputs they can take back to their own research.
- Have collective ideas on integrating HINA with existing QE methods.

To ensure impact that extends well beyond the conference, we will pursue several post-workshop activities:

- A dedicated mailing list to support ongoing peer collaboration, trouble-shooting, and the sharing of HINA applications across research groups.

- Potential research collaborations inviting workshop participants to explore the integration of HINA with QE methods. This initiative aims to foster sustained scholarly dialogue and method development.

References

1. Feng, S., He, B., Gasevic, D., & Kirkley, A. (2025a). Heterogeneous Interaction Network Analysis (HINA): A new learning analytics approach for modelling, analyzing, and visualizing complex interactions in learning processes. *arXiv preprint arXiv:2601.06771.*
2. Feng, S., He, B., & Kirkley, A. (2025b). HINA: A learning analytics tool for heterogenous interaction network analysis in Python. *Journal of Open Source Software*, 10(111), 8299.
3. Feng, S., Gibson, D., & Gašević, D. (2025c). Analyzing students' emerging roles based on quantity and heterogeneity of individual contributions in small group online collaborative learning using bipartite network analysis. *Journal of Learning Analytics*, 12(1), 253–270.
4. Feng, S., Yan, L., Zhao, L., Martinez-Maldonado, R., & Gašević, D. (2024). Heterogenous network analytics of small group teamwork: Using multimodal data to uncover individual behavioral engagement strategies. In Proceedings of the 14th *Learning Analytics and Knowledge Conference* (pp. 587–597).
5. Shaffer, D. W., Collier, W., & Ruis, A. R. (2016). A tutorial on epistemic network analysis: Analyzing the structure of connections in cognitive, social, and interaction data. *Journal of Learning Analytics*, 3(3), 9–45.
6. Vygotsky, L. S. (1978). Mind in Society: *The Development of Higher Psychological Processes*. Harvard University Press.

Open-Source Tools for Transcribing and Coding Multimodal Qualitative Data: An Interactive Tutorial

Kirk Vanacore[1,2], Bakhtawar Ahtisham[1,2], Zhuqian Zhou[1,2], Daryl Hedley[1,3], Doug Pietrzak[1,3], Joshua Marland[1,2], Rachel Slama[1,2], and René Kizilcec[1,2]

[1] National Tutoring Observatory
[2] Cornell University
[3] FreshCognate

Abstract. Researchers in Quantitative Ethnography (QE) often work with multimodal discourse data, including interviews, conversations, observations, online discussions, screen recordings, audio, video, and interaction logs. These data create opportunities to study meaning-making, participation, collaboration, identity, expertise, and institutional practice at scale, but they also introduce methodological challenges: multimodal interactions must be transformed into analyzable representations, and qualitative coding workflows must remain valid, transparent, and grounded in human interpretation. This tutorial introduces open-source tools for transcribing and coding multimodal qualitative data. Participants will learn how to use multimodal transcription workflows and *Sandpiper*, an AI-assisted coding system that orchestrates multiple models to support scalable qualitative analysis while preserving human review and researcher control. The session will combine short methodological overviews, live demonstrations, and hands-on coding activities using sample data. By the end of the tutorial, participants will understand how to move from raw video, audio, screen, and dialogue data to structured transcripts, coding tasks, AI-assisted coding outputs, and validation artifacts suitable for downstream quantitative ethnographic analysis.

Keywords: Quantitative ethnography · Qualitative coding · AI-assisted annotation · Multimodal data · Open-source tools

1 Introduction and Motivation

Quantitative ethnography (QE) depends on moving between rich qualitative interpretation and structured representations that support quantitative analysis [1]. This movement is not merely technical: it depends on consequential decisions about data selection, segmentation, coding, and interpretation [2]. Researchers increasingly have access to large collections of interviews, conversations, observations, online discussions, collaborative work sessions, screen recordings, audio,

video, and interaction logs. These data create opportunities to study meaning-making, participation, collaboration, identity, expertise, and institutional practice at scale, including through epistemic network analysis and related QE methods that model patterns in coded discourse and interaction data [3]. However, raw qualitative and multimodal data also present methodological challenges because context-specific meanings can be lost or distorted when complex interactional records are reduced to text or treated as simple classification inputs.

Recent advances in artificial intelligence make it increasingly feasible to support multimodal transcription, qualitative coding, and discourse analysis with computational tools. These tools can reduce the burden of manual transcription and annotation while expanding the scale of qualitative inquiry, but they must be embedded in transparent, auditable, and human-centered workflows that allow researchers to evaluate construct validity, reliability, and alignment with human judgment. Automated coding should therefore support, rather than replace, qualitative interpretation. This tutorial explores open-source workflows for processing and annotating qualitative and multimodal interaction data, focusing on two linked components: multimodal transcription and qualitative coding with *Sandpiper*, an AI-assisted system that uses orchestrated AI models to support scalable and auditable coding [4, 6]. Participants will gain practical experience moving from raw multimodal data to structured transcripts, coding tasks, AI-assisted labels, and validation artifacts suitable for downstream quantitative ethnographic analysis.

2 Tutorial Focus

The goal of this tutorial is to support practical, transparent uses of AI in qualitative data processing and coding, for QE researchers who are interested in using AI to assist them in data processing and qualitative coding. The tutorial introduces two open-source components developed by the National Tutoring Observatory: a multimodal transcription workflow for transforming video, audio, screen, and dialogue data into structured qualitative records, and *Sandpiper*[4], an AI-assisted coding system for configuring coding tasks, applying codebooks, inspecting model outputs, and validating annotations. Rather than treating AI as a black-box shortcut, the tutorial foregrounds researcher control, methodological transparency, and the opportunities and limits of AI-assisted workflows for QE.

3 Open-Source Workflow

The tutorial introduces a two-part workflow for transcribing and coding qualitative and multimodal interaction data. First, participants will examine a multimodal transcription pipeline that aligns audio, video, screen recordings, dialogue, and interaction logs into structured records that preserve timing, modality, and

[4] `www.sandpiperresearch.org`

interactional context (Figure 1). This component emphasizes how transcription decisions shape downstream coding and interpretation.

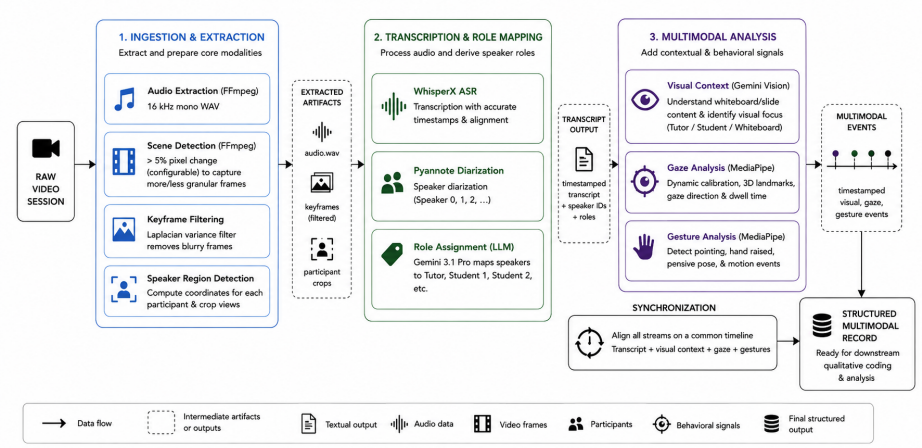


Fig. 1. Multimodal transcription and analysis pipeline.

Second, participants will use *Sandpiper* [4], an AI-assisted annotation application for qualitative and multimodal data. Sandpiper supports codebook-based coding, model-output inspection, and annotation validation. The tutorial emphasizes orchestrated AI models rather than a single one-shot classifier: multiple model calls can generate initial codes, verify labels, surface uncertain cases, and support disagreement-centered review [5, 6]. Participants will examine how code definitions, prompts, model outputs, and human review shape the final analytic dataset.

4 Tutorial Format and Schedule

The tutorial will be organized as a three-hour interactive session combining brief methodological overviews, live demonstrations, and hands-on work with sample data. Participants should bring a laptop. No advanced programming experience is required, though participants should be comfortable working with qualitative data, transcripts, or codebooks.

5 Hands-On Activity and Outcomes

The central activity asks participants to move through a condensed multimodal transcription and AI-assisted qualitative coding workflow. Participants will inspect a short multimodal transcript, review how transcript structure preserves

Table 1. Proposed Three-Hour Tutorial Schedule

Time	Activity	Purpose
0:00–0:15	Framing	Introduce goals and motivate multimodal transcription and AI-assisted coding for QE.
0:15–1:15	Transcription workflow	Demonstrate transformation of video, audio, screen, dialogue, and interaction data into structured transcripts; review outputs and discuss implications for coding.
1:15–1:25	Setup	Access sample data, tutorial materials, and Sandpiper.
1:25–2:20	Coding with Sandpiper	Configure a codebook, inspect or run orchestrated AI-assisted coding, and review model outputs.
2:20–2:45	Validation and auditing	Compare AI-assisted codes to human judgment, inspect disagreements, revise code definitions, and discuss reliability and validity.
2:45–3:00	Wrap-up	Discuss adaptation to participant projects and share open-source materials.

timing and interactional context, configure a coding task in *Sandpiper*, examine AI-assisted codes and rationales, compare outputs to human-coded examples, and revise coding instructions or code definitions. Participants will leave with a practical understanding of how to audit and refine AI-assisted coding, along with a reusable workflow, sample Sandpiper configuration, template codebook, example coding outputs, validation checklist, and links to open-source repositories, documentation, sample data, and tutorial materials.

6 Relevance to Quantitative Ethnography

This tutorial contributes to the QE community by addressing a central methodological challenge: how to transform rich qualitative and multimodal records into structured representations suitable for analysis while preserving interpretive rigor. AI-assisted coding tools are increasingly available, but their use in research requires careful attention to transparency, validity, reliability, and reproducibility. By treating transcription and coding as connected analytic practices, this tutorial shows how open-source tools can support human-centered, auditable workflows. The workshop will support research across multiple domains, including interviews, workplace interaction, collaborative learning, online discussion, tutoring, classroom discourse, and other multimodal interaction settings.

7 Organizer Expertise

Our team has experience developing open-source infrastructure for multimodal transcription and AI-assisted qualitative coding. Our work focuses on building

tools that allow researchers to transform complex interaction data into structured analytic representations while preserving researcher control and interpretive transparency. Members of our team attended ICQE last year and are eager to contribute to and learn from this community.

References

1. Shaffer, D.W.: Quantitative Ethnography. Cathcart Press, Madison (2017)
2. Zörgő, S., Peters, G.-J.Y., Porter, C., Moraes, M., Donegan, S., Eagan, B.: Methodology in the mirror: A living, systematic review of works in quantitative ethnography. In: Wasson, B., Zörgő, S. (eds.) Advances in Quantitative Ethnography. ICQE 2021. Communications in Computer and Information Science, vol. 1445, pp. 144–159. Springer, Cham (2022). https://doi.org/10.1007/978-3-030-93859-8_10
3. Shaffer, D.W., Collier, W., Ruis, A.R.: A tutorial on epistemic network analysis: Analyzing the structure of connections in cognitive, social, and interaction data. Journal of Learning Analytics 3(3), 9–45 (2016). https://doi.org/10.18608/jla.2016.33.3
4. Hedley, D., Pietrzak, D., Dias, J., Burden, I., Ahtisham, B., Zhou, Z., Vanacore, K., Marland, J., Slama, R., Reich, J., Koedinger, K., Kizilcec, R.F.: Sandpiper: Orchestrated AI-Annotation for Educational Discourse at Scale. arXiv preprint arXiv:2603.08406 (2026). https://doi.org/10.48550/arXiv.2603.08406
5. Ahtisham, B., Vanacore, K., Lee, J., Zhou, Z., Kizilcec, R.F., Baker, R.S.: AI annotation orchestration: Evaluating LLM verifiers to improve the quality of LLM annotations in learning analytics. In: Proceedings of the 16th International Conference on Learning Analytics and Knowledge (LAK 2026). ACM, New York (2026). https://doi.org/10.1145/3785022.3785095
6. Ahtisham, B., Vanacore, K., Kizilcec, R.F.: Optimizing LLM annotation of classroom discourse through multi-agent orchestration. arXiv preprint arXiv:2603.13353 (2026). https://doi.org/10.48550/arXiv.2603.13353

Workshop: Introduction to Teaching QE 2026

Ayano Ohsaki[1][0000−0001−5072−0135], Brendan Eagan[2][0000−0002−9189−1747], Zhichun Liu[3][0000−0002−6402−9174], Amalia Daché[4], Yujeong Cho[4], and Amanda Barany[4][0000−0003−2239−2271]

[1] Tohoku University, Sendai, Miyagi 9808576, Japan
aohsaki@ohsaki-lab.net
[2] University of Wisconsin–Madison, Madison, WI 53706, USA
[3] The University of Hong Kong, Hong Kong, Hong Kong China SAR
[4] University of Pennsylvania, Philadelphia, PA 19104, USA

Abstract. This is the second workshop in the series introducing education in Quantitative Ethnography (QE), organized by Teaching (with) QE, a Special Interest Group of the International Society for Quantitative Ethnography (ISQE). This year's focus encompasses a range of educational programs in QE, including practitioner-oriented courses, newly developed undergraduate courses, and QE sessions integrated into graduate-level ethnography courses. During the first half of the workshop, presenting educators will share their approaches, challenges, and best practices for teaching QE. In the second half, participants will engage in small group discussions to explore practical strategies applicable to their own teaching. The workshop is intended for researchers and practitioners currently engaged in QE education, individuals contemplating the integration of QE into their teaching, and learners aiming to deepen their understanding of QE pedagogy.

Keywords: Teaching QE · QE education · Quantitative data analysis · Qualitative data analysis · Quantitative ethnography.

1 Introduction

With the increasing engagement of researchers and practitioners in quantitative ethnography (QE), various educational initiatives have emerged. This expansion has stimulated discussions regarding effective pedagogical methods, suitable curricular content, and assessment strategies. To date, a diverse array of learning opportunities related to QE has been offered, including workshops at international conferences (e.g., [3]), globally delivered QE Master Courses in 2023–2024 (e.g., [2]), and an instructor training program [9]. Educational resources aimed at broader audiences—such as podcasts, webinars, instructional materials, and translated video content [8, 6, 4, 1, 5]—are becoming increasingly accessible. In alignment with these developments, the importance of creating forums for discussion and information sharing to enhance the design and delivery of educational programs is also growing. This year, the Teaching (with) QE, a special interest group (SIG) of the International Society for Quantitative Ethnography (ISQE),

is hosting the second workshop on teaching QE. The theme of this workshop is "the diversity of teaching QE." Last year, we concentrated on formal undergraduate and graduate QE education, with experienced presenters introducing their courses and experiences[7]. Building on last year's success, we are broadening our discussion, as the SIG's interests encompass instruction across contexts, teaching research, community-building, and interdisciplinary approaches. The workshop will provide a collaborative space for designing teaching materials with educators, students, and scholars experienced in QE.

2 Target audience

We extend an invitation to researchers, educators, practitioners, industry experts, and students with an interest in QE education, especially:

- Educators and teaching assistants involved in designing or supporting QE courses
- Instructors incorporating QE into various disciplines or curricula
- Researchers and industry professionals investigating the educational uses of QE
- Students seeking structured methods for learning QE

In this workshop, we will share a registration form for the SIG's online community, through which members can access our educational resources, monthly meeting links, and more.

3 Workshop Structure and Activities

This workshop is structured to promote knowledge sharing and collaborative development across a three-hour session. The first half features a panel discussion with three invited panelists—Brendan Eagan, Zhichun Liu, and Amalia Daché—each of whom will draw on their respective experiences in QE education (Table 1). The panelists will highlight key elements of effective QE instruction, reflecting on the challenges they have navigated and the practices that have proven successful. The panel discussion will be followed by an open Q&A session.

The second half shifts to hands-on collaborative work. Participants will form small groups to discuss their own teaching practices and learning goals, building on insights gained from the panel. These discussions will center on identifying and leveraging best practices in QE education. The workshop will conclude with a synthesis of key takeaways, a conversation about concrete next steps for implementing QE instruction, and an exploration of opportunities for sustained collaboration among participants.

Table 1. Overview of Panelists' Teaching Practices.

	Eagan's case	Liu's case	Daché's case
Target participants	Practitioners	Graduate students	Graduate students
Duration	1 week (in-person), 1 year (online)	4 weeks	2 out of 16 weeks
Format	Hybrid	In-person	In-person
Participants per course	30	23	9
Key characteristics	Over 500 participants since 2016	Newly launched in AY2026	QE instruction integrated into an ethnography course

4 Workshop Panelists

Brendan Eagan is the associate director for Research at the Center for Research on Complex Thinking at UW–Madison. For nearly a decade, he has built and led QE courses, masterclasses, and workshops for diverse research teams worldwide. As the Principal Investigator for the NSF-funded QE Institute, Brendan translates complex methodologies into scalable training programs, bringing a decade of proven pedagogical expertise to this session on teaching QE.

Zhichun Liu is an assistant professor at the University of Hong Kong. His research explores game-based learning and assessment. He is deeply interested in using mixed-method approaches to capture and understand complex learning behaviors outside of formal classroom settings. He recently designed and taught a new QE course at HKU. Additionally, he plays an active role in expanding the methodology's outreach by hosting this year's QE Data Challenge to foster global engagement.

Amalia Daché is an associate professor at the University of Pennsylvania, where she integrates storytelling with data-informed methods to examine educational access and AfroLatin American cultures. In 2024, she gave the ICQE keynote address on "AfroLatin Futurism." She also served as a panelist at last year's workshop and shared her practice, "Humanizing Data: Storytelling, Folklore & Ethnography." In this workshop, she will introduce her updated QE teaching practices in an ethnography course and encourage participants to engage with QE to rethink data through the lenses of ethnography and cultural storytelling.

5 Expected outcomes

This workshop aims to achieve two primary outcomes: identifying core elements of QE education through sharing diverse teaching practices, and discovering

new QE instruction opportunities. These outcomes are particularly relevant to educators who want to integrate QE into existing curricula and those looking to develop new QE courses. Insights from panelists' teaching experiences will serve as practical guidelines for educators in achieving these goals. Moreover, the core elements identified are linked to the knowledge, skills, and dispositions required for conducting QE research. Thus, the workshop is expected to offer meaningful learning opportunities for those studying QE, thereby advancing their understanding of the field.

References

1. Center for Research on Complex Thinking: Resources, http://epistemicnetwork.org/resources/, last accessed 28 August 2026
2. Centre for Academic Development and Leadership Excellence, Universiti Putra Malaysia: Quantitative ethnography masterclass (2023), https://cadelead.upm.edu.my/content/cpd2023_20_feb_quantitative_ethnography_masterclass-71128, last accessed 28 August 2026
3. Choi, J., Siebert-Evenstone, A., Lee, S.: Introduction to automated coding in qe. In: Damşa, C., Barany, A. (eds.) Fourth International Conference on Quantitative Ethnography: Conference Proceedings Supplement. p. 177. International Society for Quantitative Ethnography (2022)
4. International Society for Quantitative Ethnography: Archived webinars, https://www.qesoc.org/webinar-archives/, last accessed 28 August 2026
5. International Society for Quantitative Ethnography: General quantitative ethnography, https://www.qesoc.org/general-quantitative-ethnography/, last accessed 28 August 2026
6. International Society for Quantitative Ethnography: Youtube channel (2020), https://www.youtube.com/@internationalsocietyforqua3967, last accessed 28 August 2026
7. Ohsaki, A., Daché, A., Rivera-Kumar, S., Shaffer, D.W., Williamson, M., Barany, A.: Workshop: Introduction to teaching qe. In: Seventh International Conference on Quantitative Ethnography: Conference Proceedings Supplement. pp. 40–44 (2025)
8. Shaffer, D.W.: EDPSYCH 551: The QE Podcast. Audio podcast (2022), https://podcasts.apple.com/us/podcast/edpsych-551-the-qe-podcast/id1680023896, last accessed 28 August 2026
9. White, L., University of Wisconsin-Madison School of Education: "The Right Place at the Right Time": UW–Madison Unites Researchers at Inaugural Quantitative Ethnography Summer Institute (2023), https://education.wisc.edu/news/the-right-place-at-the-right-time-uw-madison-unites-researchers-at-inaugural-quantitative-ethnography-summer-institute/, last accessed 28 August 2026

Reconstructing Meaning Through Quantitative Ethnography: Critical Perspectives on Race, Immigration, and Identity

Adaurennaya C. Onyewuenyi[1[0000-0002-3818-8091]],
Danielle Espino[2[0000-0002-6885-2125]],
Golnaz Arastoopour Irgens[3[https://orcid.org/0000-0002-5308-3925]],
Nichole M. Garcia[4],
Brendan Eagan[5[0000-0002-9189-1747]]

[1] The College of New Jersey, NJ, USA
[2] Pepperdine University, CA, USA
[3] Vanderbilt University, TN, USA
[4]Rutgers, The State University of New Jersey, NJ, USA
[5]University of Wisconsin-Madison, Madison, WI, USA

Abstract. This symposium brings together scholars examining race, immigration, identity, and learning to explore how Quantitative Ethnography (QE) can reconstruct meaning in ways that challenge dominant narratives and center marginalized experiences. Drawing on critical, feminist, decolonial, and diasporic perspectives, presenters demonstrate how QE can be used not only to model connections through discourse, but also to interrogate power, surface epistemic complexity, and illuminate lived experiences often obscured within traditional analytic frameworks. Across diverse educational and social contexts, the papers examine cross-cultural digital learning among Kenyan and U.S. high school students, intra-Black relational experiences among Black American and Black immigrant college students, sociocritical AI literacies among Black middle school girls, and Latinx undergraduates' reflections on disaster preparedness during COVID-19. Collectively, these studies illustrate how QE can support more expansive and justice-oriented forms of inquiry by tracing relationships among discourse, identity, culture, and systems of inequity. Together, the symposium advances ongoing conversations about the possibilities of critical and interpretive approaches within QE, highlighting how researchers can reconstruct meaning through methodological practices that embrace contextual complexity and marginalized ways of knowing.

Keywords: Critical Methodologies; Quantitative Ethnography; Race & Immigration; Identity development; Epistemic Justice

1 Background

This symposium brings together interdisciplinary scholars of race, immigration, identity, and learning to explore how Quantitative Ethnography (QE) can be reimagined to challenge dominant research paradigms and advance epistemic and social justice in educational research. Collectively, the session examines a broader question of growing

importance within the QE community: *how can QE methodologies deepen our understanding of power, inequity, representation, and marginalized meaning-making?*

Drawing on critical, feminist, decolonial, and diasporic perspectives, presenters demonstrate how QE can surface multiple perspectives, disrupt master narratives, and foreground the lived experiences of racially and ethnically marginalized communities through methodological innovation and interpretive depth.

Presentations span a range of educational and social contexts, including:

- A study of cross-cultural digital learning environments in which QE is used to map affective and epistemic engagement across collaborations between Kenyan and U.S. high school students, challenging Global North dominance in knowledge production and participation hierarchies.
- A study examining identity negotiation among Black American and Black immigrant[1] college students, where QE reveals ethnic variation in how students conceptualize solidarity, belonging, and intra-Black relationships.
- A project investigating how Black middle school girls develop sociocritical AI literacies through scaffolded critique and design activities, using QE to model how youth interrogate, critique, and reimagine artificial intelligence systems.
- A study exploring how Latinx college students from Chicanx[2] and Puerto Rican communities make sense of disaster and systemic disruption during COVID-19, using QE to trace emotional, cultural, and relational discourses surrounding resilience and recovery

Together, these projects demonstrate how QE can be used not only to model discourse and meaning at scale, but also to critique, resist, and transform inherited research logics. The symposium advances ongoing conversations within the QE community regarding interpretive plurality, epistemic justice, and the role of QE in producing more equitable and socially responsive forms of inquiry.

2 Symposium Aims and Contributions

2.1 Symposium Aims

This symposium explores questions such as:

- How can QE be used to examine participation hierarchies, epistemic injustice, and discourse contributions within cross-cultural and Global South–Global North collaborations?

[1] In this study, *Black American* refers to Black individuals with multigenerational roots in the United States, particularly those whose ancestry is linked to the historical enslavement of African people in the United States, whereas *Black immigrant* refers to Black individuals with more recent familial migration histories to the United States, primarily from Africa, the Caribbean, or Latin America.

[2] In this study, *Latinx* refers broadly to individuals with personal or familial roots in Latin America, whereas *Chicanx* refers more specifically to individuals of Mexican descent in the United States. In both terms, the "-x" replaces the gendered Spanish suffixes "-o" (masculine) and "-a" (feminine) with a gender-neutral or non-binary alternative.

- How do patterns of solidarity, conflict, and identity negotiation emerge in QE models of Black American and Black immigrant students, and what do these patterns reveal about racialized meaning-making?
- How can QE illuminate the development of sociocritical AI literacies among youth, including the relationship between technical learning and sociopolitical consciousness?
- How can QE methodologies be adapted to center the epistemologies and lived experiences of Communities of Color while disrupting dominant analytic and statistical norms?
- What possibilities and tensions emerge when critical, feminist, decolonial, and diasporic frameworks are integrated into QE research and interpretation?

2.2 Contributions

This symposium advances a timely and important conversation about how research methodologies can both reflect and transform educational and lived realities. Across varied contexts and populations, the papers demonstrate both the shared and the distinct ways in which QE can uncover epistemic complexity, amplify marginalized voices, and reconstruct research logics grounded in justice, social context, and interpretive plurality. The symposium positions QE as a flexible, critical, and justice-oriented methodological approach capable of engaging complex social realities while maintaining analytic rigor and methodological transparency. In doing so, the session contributes to broader discussions within the QE community regarding the future of Critical Quantitative Ethnography and the ethical responsibilities of researchers working across race, immigration, identity, and educational inequity.

We anticipate that the following benefits and contributions will emerge from the symposium discussion and subsequent research outputs:

- Expanded conversations within the QE community regarding the possibilities and tensions of integrating critical, feminist, decolonial, and diasporic perspectives into QE research and interpretation.
- Greater methodological attention to how QE can be used to surface epistemic complexity, examine power relations, and center the lived experiences and meaning-making practices of racially and ethnically marginalized communities.
- New insights into how QE methodologies can support justice-oriented educational research across varied contexts, including global virtual collaborations, AI learning environments, identity development, and experiences of systemic disruption.
- Opportunities for interdisciplinary collaboration among scholars working across education, psychology, learning sciences, migration studies, critical studies, and computational social science.
- Practical considerations for researchers seeking to design QE studies that balance analytic rigor with contextual sensitivity, ethical reflexivity, and social responsibility.
- Broader recognition of QE as a flexible methodological approach capable of not only modeling discourse and connections at scale, but also reconstructing

research practices in ways that are more equitable, inclusive, and responsive to historically underrepresented communities.

3 Proposed Symposium Structure

The symposium will begin with a brief introduction that frames QE as a humanizing and justice-oriented methodological approach capable of reconstructing meaning by centering marginalized voices, epistemic complexity, and interpretive plurality. The chair will introduce the broader goals of the session, including how critical, feminist, decolonial, diasporic, and participatory perspectives can expand the methodological and ethical possibilities of QE research across contexts related to race, immigration, identity, and learning.

Following the introduction, each presenter will deliver a 12-minute presentation highlighting their research context, methodological approach, and substantive findings and contributions. Presenters will discuss how QE methodologies can be used to surface multiple forms of meaning-making, challenge dominant research paradigms, and illuminate the lived experiences of historically marginalized communities. Across papers, presenters will address the affordances, tensions, and possibilities of applying QE to issues such as cross-cultural collaboration, intra-Black relational dynamics, sociocritical AI literacies, and community responses to systemic disruption.

Following the four paper presentations, the chair will provide a 10-minute synthesis of the symposium's major themes, highlighting points of convergence and divergence across the papers while reflecting on broader implications for methodological innovation, epistemic justice, and the future directions of QE research.

The symposium will conclude with a 20-minute interactive discussion with attendees focused on collective reflection, methodological dialogue, and future collaborations. This discussion will invite participants to consider how QE can continue to evolve as a critical, inclusive, and socially responsive methodological approach capable of reconstructing meaning across diverse educational and social contexts.

The symposium will be structured as follows:

1. **Introduction and Framing: QE as a Humanizing and Justice-Oriented Methodology (5 minutes)**
    a. Overview of the symposium theme and goals
    b. Framing QE as a methodological approach that can center marginalized perspectives and reconstruct meaning
2. **Paper Presentations (48 minutes total; 12 minutes each)** Presenters will discuss:
    a. Their research context and substantive focus
    b. QE methodologies, techniques, and analytic approaches used
    c. How their work engages critical perspectives on race, immigration, identity, and power
    d. Methodological affordances, tensions, and implications for justice-oriented inquiry
3. **Discussant Reflections and Synthesis (10 minutes)**
    a. Synthesis of major themes across presentations

   b. Reflections on methodological innovation and future directions for QE
   c. Discussion of implications for Critical Quantitative Ethnography and epistemic justice
4. **Interactive Audience Discussion (20 minutes)**
   a. Open discussion with attendees
   b. Questions, collaborative dialogue, and reflections on future contributions to QE and related disciplinary communities
   c. Exploration of possibilities for critical, participatory, and humanizing QE research moving forward

4 Symposium Contributors

The symposium brings together a panel of five contributors, each coming from different disciplinary, community, and epistemological backgrounds and practices, and making significant contributions to the conversation of criticality and QE.

Dr. Golnaz Arastoopour Irgens, the author of the paper on sociocritical AI literacies among Black middle school girls, is an Assistant Professor of Human-Centered Technologies, director of the IDEA lab, and core faculty in LIVE, the Learning Innovation Incubator at Vanderbilt. Her research advances equitable and justice-oriented approaches to learning, computing, and STEM education through Quantitative Ethnography (QE). Her work examines how learning, identity, participation, and power emerge within complex sociotechnical systems, with particular attention to inclusive computing and AI learning environments. Drawing on interdisciplinary approaches across the learning sciences, computational methods, and participatory research traditions, her scholarship explores how QE can support more collaborative, humanizing, and socially responsive forms of inquiry. Her previous work on participatory QE examines how researchers and participants can collaboratively engage in meaning-making, interpretation, and visualization throughout the analytic process.

Dr. Brendan Eagan, the symposium chair and discussant, is Associate Director for Partnerships and Community Engagement for the Center for Research on Complex Thinking at the University of Wisconsin-Madison and the principal investigator for the NSF-funded Quantitative Ethnography Institute (NSF Awards #2225240; #2526678). He has contributed to the fields of language learning, cultural exchange, learning sciences, computer-supported collaborative learning, learning analytics, health care, serving people with disabilities, ethics in education, child development, and professional development in a range of domains. He is a bridge-builder within the QE community and has experience connecting research and practice in these different domains.

Dr. Danielle Espino, the author of the paper on cross-cultural digital learning among Kenyan and U.S. high school students, is a co-principal investigator at Pepperdine University for NSF-funded research examining the impact of cross-boundary collaboration in informal STEM learning on adolescent identity development (NSF Award #2215613), utilizing quantitative ethnography (QE) as a primary methodological

approach. Papers she has authored highlight different ways epistemic network analysis has been used to uncover findings on cross-cultural community building, global competencies, connections between affect and discourse, group dynamics, and to advance conversations on justice, equity, diversity, and inclusion (JEDI). Her previous professional experience includes project management roles in the arts nonprofit sector and at various higher education institutions in California, New York, and Virginia.

Dr. Nichole M. Garcia, the author of Latinx undergraduates' reflections on disaster preparedness during COVID-19, is an associate professor of higher education at Rutgers University, New Brunswick. She is a critical race and feminist mixed-methodologist who focuses on the differences and similarities between Latinx sub-ethnic groups, students, and families regarding their college readiness, retention, and completion. Known for her scholarship on QuantCrit, Garcia will discuss her positionality as a QuantCrit (Gillborn et al., 2018; Garcia et al., 2018) and Chicana/Latina feminista entering and deciphering the emerging field of Quantitative Ethnography. Through an epistolary offering, she will groundtruth the use of story, fairness, and community through a Critical Race and Chicana/Latina Feminist lens.

Dr. Adaurennaya "Ada" C. Onyewuenyi, the author of intra-Black relational experiences among Black American and Black immigrant college students, is an Associate Professor of Psychology and Affiliate Faculty of African American Studies at The College of New Jersey. She leads the Identity Development across the African Diaspora (IDAD) Lab. As an applied developmental and educational psychologist with an emphasis in diversity science, her interdisciplinary research examines race, immigration, identity development, and educational inequities, with particular attention to the experiences of Black American and Black immigrant youth and young adults. Drawing on mixed methods and Quantitative Ethnography approaches, her work explores how racial discrimination, belonging, and systems of inequality shape educational and mental health outcomes across diverse learning contexts.

5 Summary

Through this symposium, we aim to create space for imagining new possibilities and futures for QE and critical inquiry. The session invites both emerging and established QE scholars to engage more deeply with the methodological, epistemological, and ethical possibilities of convergence research that integrates qualitative, quantitative, and computational approaches. By foregrounding issues of power, representation, and epistemic justice, the symposium highlights how QE can be used not only to model meaning and discourse, but also to reconstruct research practices in ways that are more critical, inclusive, and socially responsive.

References

1. Arastoopour Irgens, G., Vega, H., Adisa, I., Bailey, C.: Characterizing children's conceptual knowledge and computational practices in a critical machine learning educational program. Int. J. Child-Comput. Interact. 34, 100541 (2022)
2. Arastoopour Irgens, G., Adisa, I., Bailey, C., Quesada, H.V.: Designing with and for youth. Educ. Technol. Soc. 25(4), 126–141 (2022)
3. Battle-Baptiste, W., Rusert, B. (eds.): W. E. B. Du Bois's Data Portraits: Visualizing Black America. Princeton Architectural Press, New York (2018)
4. Bonilla-Silva, E.: Racism Without Racists: Color-Blind Racism and the Persistence of Racial Inequality in America, 2nd edn. Rowman & Littlefield, Lanham (2006)
5. Connell, R.: Southern Theory: The Global Dynamics of Knowledge in Social Science. Routledge, London (2007). https://doi.org/10.4324/9781003117346
6. Covarrubias, A., Vélez, V.N.: Critical race quantitative intersectionality: An anti-racist research paradigm that refuses to "let the numbers speak for themselves." In: Lynn, M., Dixson, A.D. (eds.) Handbook of Critical Race Theory in Education, pp. 290–306. Routledge, New York (2013)
7. Delgado, R., Stefancic, J.: Critical Race Theory: An Introduction, 2nd edn. NYU Press, New York (2012)
8. Fricker, M.: Epistemic Injustice: Power and the Ethics of Knowing. Oxford University Press, Oxford (2007)
9. Gillborn, D., Warmington, P., Demack, S.: QuantCrit: Education, policy, "Big Data" and principles for a critical race theory of statistics. Race Ethn. Educ. 21(2), 158–179 (2018)
10. hooks, b.: Yearning: Race, Gender, and Cultural Politics. South End Press, Boston (1990)
11. Ladson-Billings, G., Tate, W.F.: Toward a critical race theory of education. Teach. Coll. Rec. 97(1), 47–68 (1995)
12. Mignolo, W.D.: Epistemic disobedience, independent thought and decolonial freedom. Theory Cult. Soc. 26(7–8), 159–181 (2009)
13. Morris, A.D.: The Scholar Denied: W. E. B. Du Bois and the Birth of Modern Sociology. University of California Press, Berkeley (2015)
14. Shaffer, D.W.: Quantitative Ethnography. Cathcart Press (2017)
15. Shaffer, D.W., Collier, W., Ruis, A.R.: A tutorial on epistemic network analysis: Analyzing the structure of connections in cognitive and interaction data. J. Learn. Anal. 3(3), 9–45 (2016)
16. Tanksley, T.: Employing an abolitionist, critical race pedagogy in CS: Centering the voices, experiences and technological innovations of Black youth. J. Comput. Sci. Integr. 6(1), 9 (2023)

Analysis of Manzai Characteristics via Clustering Based on Switching Pauses

Akihito Takeuchi[1], Mineaki Ohishi[2][0000−0001−8727−3631], and Takashi Mitsuishi[2]

[1] Graduate School of Information Sciences, Tohoku University, Sendai, Miyagi 980-8577, Japan
akihito.takeuchi.r5@dc.tohoku.ac.jp
[2] Center for Data-driven Science and Artificial Intelligence, Tohoku University, Sendai, Miyagi 980-8576, Japan

Abstract. This paper presents analysis of the characteristics of utterance timing based on the distributions of switching pauses identified for each Manzai performer. A switching pause is defined as the time between the end of one speaker's utterance and the start of another speaker's utterance. We specifically examined 25 Manzai performances by 20 Manzai pairs. All were finalists in the M-1 Grand Prix, a major Manzai competition in Japan. Subsequently, we applied hierarchical clustering to the Manzai performances based on the distributions of the switching pauses. Results revealed the existence of clusters with different characteristics for both tsukkomi and boke.

Keywords: clustering · Manzai · switching pause · turn taking

1 Introduction

This study specifically examines Manzai, a performing art in which two Manzai performers typically present a fast-paced and humorous dialogue to make the audience laugh. In many cases, the two performers assume different roles: boke and tsukkomi. The boke role includes funny or unexpected utterances or actions. In contrast, the tsukkomi role responds to the boke's utterances by correcting or pointing them out. Earlier studies of Manzai have emphasized the study of scripts and utterance content [3, 5]. However, Manzai is a performing art presented on stage. Therefore, not only utterance contents but also switching pauses are important for making the audience laugh. A switching pause is defined as the time between the end of one speaker's utterance and the start of the other speaker's utterance. Earlier studies have also emphasized examinations of switching pauses in Manzai [4, 6]. These studies dealt with representative values of switching pauses for boke and tsukkomi to compare each performer. However, performers might start speaking immediately in some situations, but might start speaking after a short pause in other situations. Therefore, differences in switching pauses appear as distributions not as representative values. This study specifically examines those distributions.

Performers' personas and Manzai pacing styles, which may contribute to humor, are qualitative characteristics and are difficult to evaluate. However, these humor-related characteristics may be reflected in utterance timing observed as switching pauses. Therefore, this study clusters Manzai performances based on their switching pause distributions and returns to the video data to interpret the scenes characterizing each cluster. By combining this quantitative analysis with contextual interpretation, we clarify differences in performers' personas and Manzai pacing styles. This approach demonstrates a new potential application of QE. This paper investigates the following research questions:
RQ1. What types of switching pauses are found in Manzai?
RQ2. Can differences in the distributions of switching pauses reveal differences in performers' personas and Manzai pacing styles?

2 Data and Methods

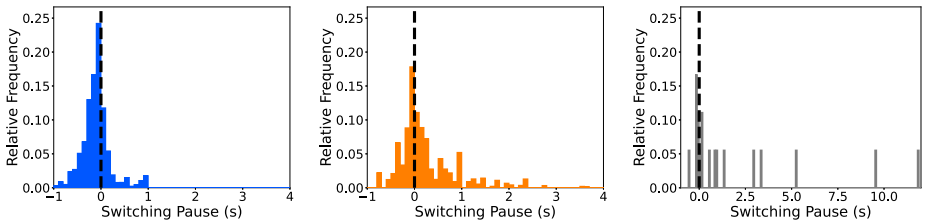
The analysis performed for this study targets switching pauses extracted from 25 Manzai performances by 20 Manzai pairs, consisting of 24 performances [1] presented at the M-1 Grand Prix (a major Manzai competition in Japan) finals and one performance presented at a yose theater [2]. Switching pauses were defined as follows. For a transition from speaker B to speaker A, A's switching pause was the duration between the end time of B's utterance and the start time of A's utterance. B's switching pause was defined in the same way. Positive and negative values indicate gaps and overlaps, respectively. We estimated utterance intervals for the audio of each Manzai performance. For initial estimation of utterance intervals, we used pyannoteAI [7], a speaker diarization service. Then, based on the obtained utterance intervals, we corrected them manually by checking the audio waveforms and listening to the audio. Thereby we ascertained the final start and end times of utterances. We calculated switching pauses for both boke and tsukkomi. Consequently, we obtained switching pauses corresponding to 3,401 speaker transitions. Using Ward's method, we applied hierarchical clustering to the Manzai performance data based on distributions, separately for the boke and tsukkomi roles. To evaluate the closeness between distributions, we adopted the Wasserstein distance (for details, see, e.g., [8], Chaps. 2 and 5) as a dissimilarity measure. In each dendrogram, we selected the cut point based on two criteria: a large linkage distance between clusters and interpretability. Additionally, as an exploratory analysis, we inspected the corresponding video segments when distinctive switching pauses were observed.

3 Results and Discussion

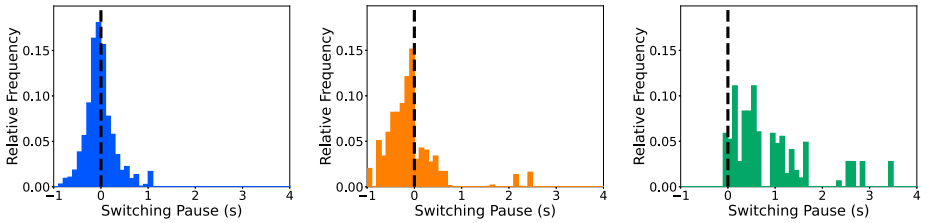
Hierarchical clustering yielded three clusters for each role. Fig. 1 presents histograms of the representative distributions, which are the Wasserstein barycenters of the clusters for tsukkomi and boke, respectively.

For tsukkomi, the cluster T_0 had a sharp peak at approximately 0 s and only slight variation. Therefore, such switching pauses may constitute one element of

the fast-paced dialogue characteristic of Manzai. The cluster T_1 had a peak at approximately 0 s, but compared with T_0, its distribution extended further into the positive range. When the corresponding parts were checked in the video, gestures that appeared to indicate thinking were observed. This observation suggests that the tsukkomi responds intentionally after a short pause to express a state of thinking to the audience. The cluster T_2 consists only of Slim-Club and includes larger positive switching pauses than the other clusters. When the corresponding parts were checked in the video, gestures indicating confusion were observed. This finding suggests that the tsukkomi responds intentionally after a long pause to express confusion to the audience.



(a) Cluster T_0 ($N = 19$) (b) Cluster T_1 ($N = 5$) (c) Cluster T_2 ($N = 1$)



(d) Cluster B_0 ($N = 19$) (e) Cluster B_1 ($N = 4$) (f) Cluster B_2 ($N = 2$)

Fig. 1: Representative distributions defined as the Wasserstein barycenters of (a–c) the tsukkomi clusters and (d–f) the boke clusters.

For boke, the cluster B_0 has a sharp peak at approximately 0 s and only slight variation. Therefore, as with the tsukkomi cluster T_0, this cluster may constitute one element of the fast-paced dialogue characteristic of Manzai. The cluster B_1 showed a distribution including more negative switching pauses than B_0. This timing of speech onset can be interpreted as an expression of selfishness, as if the boke starts speaking without waiting for the tsukkomi's utterance to finish. Therefore, the negative switching pauses observed in B_1 might be used by boke characters intentionally to express selfishness. The cluster B_2 showed a distribution in which many switching pauses took positive values. This timing of speech onset can be inferred as an expression of a self-paced character, as if the boke character starts speaking at their own pace after receiving the other speaker's utterance. Therefore, the positive switching pauses observed in

B_2 might be used by the boke character intentionally to express a self-paced character. Consequently, the distributions of switching pauses obtained by hierarchical clustering can be interpreted as reflecting performers' personas and Manzai pacing style.

4 Conclusion

This study analyzed the switching pause characteristics in Manzai. We obtained three clusters for each role. These clusters might be related to maintaining fast dialogue in Manzai, expressing a state of thinking, expressing confusion, expressing selfishness and expressing a self-paced character. These results suggest that differences in the distributions of switching pauses can be interpreted as reflecting performers' personas and Manzai pacing styles. This study demonstrated that QE can be applied to the analysis of Manzai, a form of Japanese performing art.

Acknowledgements This work was partially supported by JSPS KAKENHI Grant Number 25K21159.

Disclosure of Interests. The authors have no competing interests to declare that are relevant to the content of this article.

References

1. Asahi Broadcasting Corporation, Yoshimoto Kogyo: M-1 Grand Prix DVD Series (2005, 2007, 2008, 2009, 2010, 2015, 2017). Yoshimoto Music Entertainment (2005–2017), (in Japanese)
2. Grape Company Channel: Sandwichman manzai: Police 24 hours, 2018. YouTube (2022), https://youtu.be/mdPUInNdlCc, accessed March 19, 2026. (in Japanese)
3. Kawase, A.: Structural analysis of manzai scripts: Extracting patterns associated with higher ratings in the m-1 grand prix. Mathematical Linguistics **35**(3), 54–69 (2025). https://doi.org/10.24701/mathling.35.3_54, (in Japanese)
4. Kawashima, H., Scoggins, L., Matsuyama, T.: Analysis of the dynamic structure of manzai: Toward a natural utterance-timing control. Human Interface: The Transactions of Human Interface Society **9**(3), 379–390 (2007), (in Japanese)
5. Kinsui, S.: Boke and tsukkomi: An analysis of manzai conversation using pragmatics. Kamigata no Bunka: Kamigata Kotoba no Konjaku **13**, 61–90 (1992), (in Japanese)
6. Komori, M.: Relationship between boke–tsukkomi timing and funniness in manzai: An examination using cross-recurrence quantification analysis. In: Proceedings of the 72nd Annual Convention of the Japanese Psychological Association. p. 3PM112 (2008). https://doi.org/10.4992/pacjpa.72.0_3PM112, (in Japanese)
7. pyannoteAI: pyannoteAI Speaker Diarization Platform. https://www.pyannote.ai/, accessed March 19, 2026
8. Santambrogio, F.: Optimal Transport for Applied Mathematicians: Calculus of Variations, PDEs, and Modeling, Progress in Nonlinear Differential Equations and Their Applications, vol. 87. Birkhäuser Cham (2015). https://doi.org/10.1007/978-3-319-20828-2

Exploring the Effects of Program Years has on Dog Training Programs for Formerly Incarcerated Folks

Jai Oni Dane[000-0003-2572-2379]

[1] Graduate School of Education & Psychology, Pepperdine University

Abstract. This study examines how program duration shapes the experiences of formerly incarcerated individuals who participated in prison dog training programs (PDPs). Epistemic Network Analysis (ENA) was used to visualize connections among eight themes: Emotional Growth, Identity Shift, Dog Connection, Skill Building, Purpose and Pride, Community Bonds, Carceral Culture, and Institutional Barriers. Findings reveal a clear progression across program years. These results suggest that the longer formerly incarcerated individuals participate in prison dog programs, the more their engagement deepens—shifting from personal healing and emotional growth toward an integrated experience that layers vocational skill, community responsibility, and critical structural awareness on top of that ongoing personal transformation.

Keywords: Human-animal interaction, dog training programs, rehabilitation, reintegration, formerly incarcerated, quantitative ethnography

1 Introduction

Prison dog programs (PDPs) represent a growing area of interest within correctional rehabilitation research. These structured, animal-assisted programs place incarcerated individuals in caregiving and training roles with dogs—often service animals or shelter dogs—over extended periods of time. Research suggests that PDPs offer multiple pathways to rehabilitation, including emotional regulation, skill acquisition, and identity transformation [1, 2]. This research analyzes how participants' engagement shifts across two program tenure groups: those who participated for 4–6 years and those who participated for 7–10 years.

The study draws on qualitative interviews with seven formerly incarcerated individuals, analyzed using Epistemic Network Analysis (ENA) to model connections among eight thematic codes. ENA provides a quantitative and visual representation of the co-occurrence of themes within participant responses. This study was guided by two research questions:

RQ1: How do formerly incarcerated individuals reflect on their PDP experience during incarceration?

RQ2 How do individuals who have completed a PDP perceive the relationship between their participation in a prison dog training program and their ongoing rehabilitation and reintegration into society?

While prior work has established that PDPs support rehabilitation broadly [1, 2], this study contributes a methodological demonstration of how ENA can visualize that engagement evolves over program duration— an angle not previously modeled quantitatively in this population.

2 Methods

Data for this study were drawn from semi-structured interviews conducted with seven formerly incarcerated individuals who had participated in a PDP during their incarceration for a dissertation. Participants were recruited through purposive sampling via community reentry organizations. Interviews were audio-recorded, transcribed verbatim, and then eight codes were developed. Epistemic Network Analysis (ENA) was used to model and compare codes across participants, segmented by program tenure. ENA models were constructed using individual coded utterances within transcripts as the unit of analysis. Two tenure groups were examined: 4–6 years (n=3) and 7–10 years (n=3). One participant participated in the PDP for 10+ years (n=1). 10+ years was excluded from the model due to only one participant represented and did not have a comparable cohort for analysis. This participant's data may be incorporated in future work examining long-tenure (10+ year) outcomes as a distinct group Separate ENA models were generated for each research question and each tenure group, producing four network visualizations in total (Figures 1-2).

3 Results and Discussion

The ENA network models for program year participation- by-question combinations are presented below, followed by an interpretation of each. Nodes represent thematic codes and edges represent the strength of co-occurrence between codes; thicker edges indicate stronger connections and thinner edges indicate weaker ones [3].

3.1 Research Question 1

For RQ1, participants in the 4–6 year range (Fig. 1, left, pink) clustered closer to Emotional Growth and Dog Connection. Their strongest connections ran between Emotional Growth, Identity Shift, and Dog Connection, with thinner lines extending to the remaining themes. Carceral Culture and Institutional Barriers appeared less central in their network positioning. For individuals in the program for 4–6 years, reflections focused primarily on personal healing and development. Participants who had been in the program for 4–6 years highlighted significantly their growth as individuals and linked it directly to the relational dimensions of working with dogs. The network pattern reflects a rehabilitation arc centered on personal transformation rather than structural critique, consistent with earlier stages of program engagement.

Participants with 7–10 years of tenure (Fig. 1, right, purple) demonstrated a broader thematic spread across the network. Their connections among Identity Shift, Skill Building, Community Bonds, and structural codes appeared thicker and more interconnected, reflecting greater thematic integration. The 7–10 year mean position shifted lower and slightly rightward compared to the 4–6 year group. Based on the model, participants' responses illustrated that PDPs support leadership development, vocational identity, and increased social skills—participants with more years in the program did not simply repeat the same experience but deepened and broadened it.

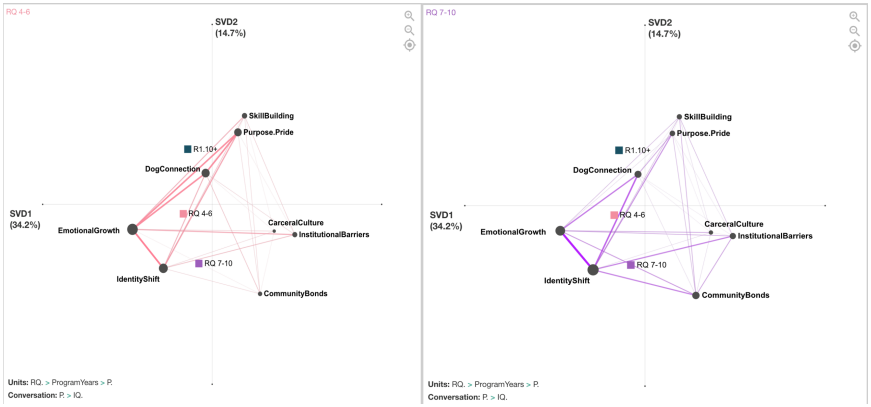


Fig. 1. ENA network models for Research Question 1, comparing participants with 4–6 years (left, pink) and 7–10 years (right, purple) of program tenure.

3.2 Research Question 2

For RQ2, participants in the 4–6 year range (Fig. 2, left, pink) primarily constructed narratives centered on emotional growth and relational trust-building. Their strongest connections ran between Emotional Growth, Identity Shift, Dog Connection, and Skill Building. The network shows moderate connections to Purpose and Pride but thinner links to Community Bonds, Carceral Culture, and Institutional Barriers. This pattern suggests that during the earlier years of program participation, participants engaged primarily with inward-facing themes—those relating to personal change and relational experience—rather than outward-facing structural themes. The network for the 7–10 year group on RQ2 (Fig. 2, right, purple) shows thick edges connecting Purpose and Pride, Skill Building, Community Bonds, The strongest connections run between Purpose and Pride, Community Bonds. Participants perceived reintegration through a wider lens that encompassed not only personal transformation and vocational skills but also collective solidarity and critical awareness of the systems they navigated. Specifically, longer-tenure participants expressed Identity Shift with Skill Building and Purpose while also articulating explicit awareness of Carceral Culture and Institutional Barriers.



Fig. 2. ENA network models for Research Question 2, comparing participants with 4–6 years (left, pink) and 7–10 years (right, purple) of program tenure.

4 Conclusion

The tenure-based comparison presented here reveals that PDP participation is not a static experience but an evolving one. Shorter-tenure participants engaged primarily with themes of personal healing, relational trust, and emotional growth—changes that are meaningful and foundational. Longer-tenure participants demonstrated qualitatively richer networks, integrating vocational identity, community responsibility, and structural awareness alongside the personal themes present in earlier cohorts. These findings suggest that the depth of rehabilitation supported by PDPs is in part a function of sustained engagement. This study offers a continued academic relevance as well as preliminary, methodologically-oriented contribution as the use ENA can show duration-based patterns in PDP participant that qualitative description alone may not capture, and it lays groundwork for larger-sample confirmatory work.

References

1. Cooke, B., & Farrington, D. P. (2016). The effectiveness of dog-training programs in prison. Prisoners and Therapy, 12(3), 1–19.

2. Turner, W. G. (2007). The experiences of offenders in a prison canine program. Federal Probation, 71(1), 38–43.

3. Shaffer, D. W. (2017). Quantitative ethnography. Cathart Press.

Mocking Gender: An Epistemic Network Analysis of *Yinyang Guaiqi* in Online Chinese Gender-Based Offensive Language

Zhenqi Xu[1] [0009-0006-5580-3208] and Ward Peeters[2] [0000-0001-5531-3926]

[1] Monash University, 20 Chancellors Walk, Melbourne VIC 3168, Australia
zhenqi.Xu@monash.edu
[2] Monash University, 20 Chancellors Walk, Melbourne VIC 3168, Australia
ward.peeters@monash.edu

Abstract. This paper examines *yinyang guaiqi* (阴阳怪气), a style of indirect, sarcastic, mocking behaviour in online Chinese gender-based offensive language (GBOL). Drawing on 119 widely circulated Chinese social media GBOL entries, it uses Epistemic Network Analysis (ENA) to examine how this mocking behaviour is taking shape, seeing which rhetorical strategies, lexical features, and evaluative domains are most prominent. The findings show that mockery is linked to irony / humour, and intertextual referencing, formed through direct labels, slogans, metaphors, and established templates. It is also concentrated around topics like gender performance, motherhood or reproduction, politics and ideology, intimate relations, and marriage or family. Combining ENA with a close reading methodology, the paper argues that Chinese GBOL often makes one's gender mockable through short, repeatable, and laughable forms of judgement. These findings reveal how mockery functions as a structured, shareable mechanism for normalising gender-based harm in Chinese digital discourse.

Keywords: Epistemic Network Analysis; Gender-based offensive language; Chinese social media; Mockery; *yinyang guaiqi*

1 Introduction

Studies of gender-based online abuse often focus on visible forms of attack: insults, threats, sexual harassment, hate speech, or derogatory labels [1, 2, 3]. Yet gender-based abuse is not always explicit. This paper asks the following research question: how are subtler attacks, marked by context, memes, labels, and tone of voice, formed in contemporary Chinese online spaces and do they have a discernible linguistic structure?

One example is the Yang Li controversy, involving a Chinese female comedian. Her joke about men who look "so ordinary, yet so confident" helped make *puxin nan* (普信男, "ordinary yet overconfident men") a widely recognisable gendered insult on Chinese social media [4]. *Puxin nan* is not a conventional swear word, but it turns a type of masculinity into something recognisable, repeatable, and mockable. A similar pattern appeared after Quan Xixi's interview with Japanese sociologist Chizuko Ueno,

where comments to the interview contained terms such as *jiaoqi* (娇妻, "pampered wife"). These expressions mocked her intimate life, marital status, feminist identity, and public persona using labels that could be quickly circulated. Here, offence lies not only in words themselves, but in how their use in context can turn gender into a joke.

This paper uses the term *yinyang guaiqi* (阴阳怪气) to describe this Chinese internet style of indirect insults. Existing work describes *yinyang guaiqi* as close to satirical tone, while also having its own phrasing, prosody, and emotional force [5]. In this study, *yinyang guaiqi* is referred to as Mockery, the closest English analytical equivalent used here rather than an exact translation**.** It refers to expressions that do not always insult directly but make a gendered person, relation, or position laughable within a shared context. The analysis examines how Mockery connects ways of speaking, forms of expression, and gendered objects of judgement, showing how Chinese GBOL turns gender into something that can be easily labelled, circulated, and laughed at.

2 Data, Coding, and ENA Setup

The analysis is based on 119 Chinese social media GBOL examples purposively drawn from major gender-related controversies, repeated cross-platform discussions, and recognisable public labels. The dataset captures recurrent and culturally salient forms of GBOL between 2023 and 2026. Each entry is treated as one unit of analysis.

The data was annotated across several dimensions: the direction of the gender attack (e.g., male to female or female to male), lexical features, rhetorical strategies, and evaluative domains. Lexical features describe how each GBOL expression is formally constructed, including direct labelling, sloganisation, metaphor, template formation, proper-name or event reference, genre marking, and homophonic substitution. Rhetorical strategy codes captured how each expression produced its offensive or mocking force, including irony or humour, intertextual reference, metaphorisation, genre parody, and categorisation. Evaluative domain codes captured what social topic was being judged, including gender performance, intimate relations, marriage or family, motherhood or reproduction, and politics or ideology. This paper focuses on Mockery as the analytical code used to capture *yinyang guaiqi* in the dataset. A preliminary co-occurrence check in R showed that this code was highly prevalent. In order to answer the research question, the ENA setup looks at how Mockery relates to three distinct features of language: Mockery × lexical formation, Mockery × rhetorical strategy, and Mockery × evaluative domain.

ENA was used to represent how lexical, rhetorical, and evaluative codes form relational configurations around Mockery, rather than examining their frequencies in isolation [6]. Each GBOL entry was treated as an independent conversation, with the terms used as the conversation variable and Mockery as the unit for all terms. A stanza window of 1 was used, so codes assigned to the same entry were treated as co-occurring. The resulting edge weights represent normalised within-entry co-occurrence rates, allowing prominent relational patterns among codes to be identified and compared [7]. Stronger edges indicate more frequent co-occurrence within the same entries, but do not imply causal or temporal relations.

3 Findings

Figure 1 presents three ENA views of Mockery in Chinese GBOL. Panel A shows clear links between Mockery and Direct Labelling (0.22), Metaphor (0.19), Sloganisation (0.19), and Template Formation (0.14). These links suggest that mockery in Chinese GBOL often travels through compact and repeatable forms. A direct label can turn someone into a recognisable type of person; a slogan or template can be repeated and adapted; a metaphor can make judgement feel immediate. *Jiaoqi* ("pampered wife") is a clear example of this pattern. It does more than name a woman imagined as attached to heterosexual intimacy or marriage. It condenses a gendered persona into a short label: dependent, politically unaware, too invested in intimacy, and, in the mocking logic of the term, not fully legitimate as a "qualified" feminist subject. Its force comes from this compression. A compact label turns gender performance and intimate life into something instantly recognisable and easy to mock.

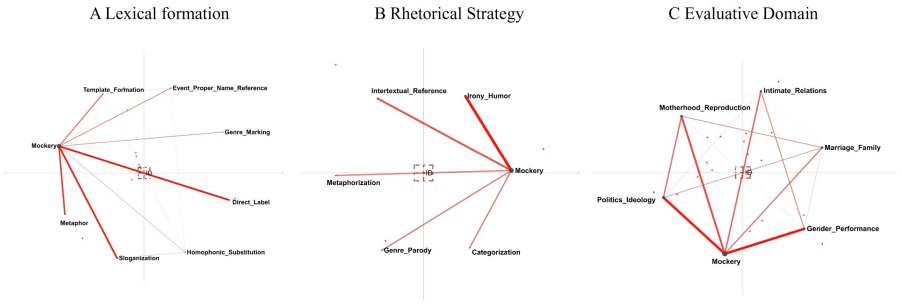


Fig. 1. Three focused ENA views of Mockery in Chinese GBOL. Panel A shows Mockery × lexical formation; Panel B shows Mockery × rhetorical strategy; Panel C shows Mockery × evaluative domain. Nodes show selected codes; edges show co-occurrences within entries.

Panel B shows how mockery becomes indirect. The strongest links are between Mockery and Irony/ Humour (0.31) and Intertextual Reference (0.20). These links support the reading of *yinyang guaiqi* as indirect, context-dependent mockery. A phrase such as "Children will have their own fortune; without children, I enjoy my own fortune" (儿孙自有儿孙福，没有儿孙我享福) makes this pattern visible. It rewrites a familiar Chinese family saying by reversing its moral direction: instead of treating descendants as a source of blessing, it turns the absence of descendants into the speaker's own good fortune. The humour comes from this reversal. Similar phrases such as *ta chao ai* (她超爱, "she loves him so much") also work through context-dependent irony. On the surface, the phrase sounds playful or even affectionate. In GBOL contexts, it can become a sarcastic shorthand for mocking a woman as overly attached, love-brained, or uncritically devoted, reducing her public position to excessive romantic attachment.

Panel C shows what becomes mockable. The strongest links gather around Gender Performance (0.27), Politics / Ideology (0.26), and Motherhood / Reproduction (0.20), with Intimate Relations (0.17) and Marriage / Family (0.15) forming a related relational cluster. These links suggest that Chinese GBOL does not only target isolated individuals. Instead, it mocks gender through the social domains in which gender is performed

and judged: appearing, relating, marrying, reproducing, and taking public positions. The reproductive-political cluster is especially clear in “Fewer births, better births; a lifetime of happiness” (少生优生，幸福一生). The phrase echoes earlier birth-planning discourse and carries a different force in a contemporary context where childbirth is increasingly encouraged. Once reproductive restraint was framed as rational, responsible, and beneficial, but has now returned as an ironic resource for questioning renewed pressure toward marriage and childbirth. The old policy language does not simply remain as memory; it becomes available for mockery.

4 Conclusion

The three ENA views and examples suggest that Chinese GBOL makes gendered roles, relations, and life choices laughable through small, repeatable forms of judgement. *Yinyang guaiqi* is productive because it combines portable lexical forms, indirect rhetorical strategies, and gendered domains of judgement. ENA makes these relational connections visible, while close reading shows how they acquire culturally specific mocking force: light enough to circulate, but sharp enough to wound. These findings characterise platformed online discourse and should not be generalised directly to everyday offline gender discourse.

Disclosure of Interests. The authors have no competing interests to declare that are relevant to the content of this article.

References

1. Jane, E.A.: ‘Back to the kitchen, cunt’: Speaking the unspeakable about online misogyny. Continuum 28(4), 558–570 (2014).
2. Jane, E.A.: Misogyny Online: A Short (and Brutish) History. SAGE, London (2017).
3. Ging, D.: Alphas, betas, and incels: Theorizing the masculinities of the manosphere. Men and Masculinities 22(4), 638–657 (2019).
4. Meng, X., Literat, I.: #AverageYetConfidentMen: Chinese stand-up comedy and feminist discourse on Douyin. Feminist Media Studies 24(4), 661–677 (2024).
5. Zhao, R.: Recognition of weird tone in Chinese communication and improvement of language understanding for AI. In: Stephanidis, C., Antona, M., Ntoa, S. (eds.) HCI International 2022 – Late Breaking Papers: Multimodality in Advanced Interaction Environments, LNCS, vol. 13518, pp. 563–571. Springer, Cham (2022).
6. Shaffer, D.W., Collier, W., Ruis, A.R.: A tutorial on epistemic network analysis: Analyzing the structure of connections in cognitive, social, and interaction data. Journal of Learning Analytics 3(3), 9–45 (2016).
7. Tan, Y., Swiecki, Z., Ruis, A.R., Shaffer, D.W.: Epistemic network analysis and ordered network analysis in learning analytics. In: Learning Analytics Methods and Tutorials: A Practical Guide Using R, pp. 569–636. Springer Nature Switzerland, Cham (2024).

Tracing Interpretive Discourse in Art Appreciation through Discussion: A Quantitative Ethnographic Analysis of Two Sessions on a Japanese Artwork

Tomoki Hirano[0009-0001-8462-9644]

[1] Future University Hakodate, Hakodate, Hokkaido, 041-8655 JAPAN
tomoki@fun.ac.jp

Abstract. This paper presents a quantitative ethnographic reanalysis of two art appreciation through discussion sessions centered on the same Japanese artwork, *Shikami-zu*, a portrait traditionally associated with Tokugawa Ieyasu. It examines how the resources supporting inference — visual evidence, external knowledge, and prior talk — shift across phases within each session. The two sessions showed similar overall rates of inferential activity but differed in their trajectories: Session 1 sustained visually grounded inference from the early phase, whereas Session 2 moved through a knowledge-mediated middle phase before returning to close looking and interpretation in the late phase.

Keywords: Art appreciation through discussion; interpretive discourse; Japanese art; quantitative ethnography.

1 Introduction and Research Questions

Art appreciation through discussion is a setting in which viewers construct meanings from visual evidence while responding to the comments of others. Visual Thinking Strategies exemplifies this approach by encouraging viewers to articulate what they notice, what they infer, and what in the artwork supports their claims [1]. Art-viewing dialogue is therefore a useful site for examining how observation, inference, and interpretation are socially organized. Quantitative ethnography (QE) offers a relevant perspective because it uses segmentation, coding, and quantification to augment ethnographic accounts of discourse without reducing them to a single analytic technique [2, 3].

The present study reanalyzes data originally examined from the perspective of information provision and knowledge building, which found that knowledge-building moves increased in the second of two art appreciation through discussion sessions on the same artwork [4].

The study is guided by two research questions:

RQ1. How do the temporal patterns of participant discourse differ across two art appreciation through discussion sessions on the same artwork?

RQ2. How do the resources supporting inference — visual evidence, external knowledge, and prior talk — shift across phases within each session?

2 Data and Context

The data consist of participant utterances from two approximately 30-minute art-appreciation discussions, referred to as Session 1 (S1) and Session 2 (S2). S1 was held on November 19, 2017 and S2 on November 20, each involving approximately 20 adult viewers and facilitated by the same student facilitator. Both sessions centered on *Shikami-zu*, a Japanese portrait traditionally associated with young Tokugawa Ieyasu, founder of the Edo shogunate, the military government that ruled early modern Japan, after a major military defeat at the Battle of Mikatagahara. The artwork differs from conventional heroic portraits: viewers often notice the figure's tense facial expression, unstable posture, awkward bodily arrangement, and striking details in clothing and objects. The two sessions were selected because they hold the artwork and facilitator constant while involving different groups of viewers.

3 Coding and Analytic Approach

Participant talk was segmented by utterance, with multiple codes allowed per utterance. The three primary codes were informed by Housen's account of aesthetic development [5]. Housen's framework describes a progression from visible features to local inference and integrative meaning-making. Here, these distinctions were adapted as utterance-level discourse categories rather than stages of individual aesthetic development. The two auxiliary flags captured the resources viewers drew on when making inferences and interpretations. All five codes are summarized in Table 1.

Table 1. Participant coding scheme.

Code	Definition
observation	Utterances describing visible features of the artwork (shape, color, posture, expression, clothing, objects)
inference	Local conjectures about the figure's emotion, bodily state, situation, intention, or production context
interpretation	Integrative meaning-making connecting multiple observations, inferences, prior information, or prior talk; a novel conjecture alone was not sufficient
ext.reference	Explicit reference to external knowledge, everyday experience, historical knowledge, or cultural references
uptake	Explicit uptake of prior talk, including another participant's comment or facilitator-provided information

The two sessions were compared in terms of three analyses. First, overall code distribution was examined using utterance-level presence/absence counts for each code. Second, each session was divided into equal thirds to provide a common temporal basis for comparing shifts in code frequencies. Third, because inference was highly prevalent, its combinations with observation, ext.reference, and uptake were examined descriptively to characterize the resources accompanying inference across phases.

Percent agreement and Cohen's kappa were, respectively, 90% and .81 for observation, 77% and .33 for inference, 79% and .39 for interpretation, 83% and .45 for ext.reference, and 89% and .73 for uptake. Because inference was highly prevalent, its low kappa is partly attributable to the prevalence effect; the prevalence-adjusted, bias-adjusted kappa (PABAK) for this code was .54. Disagreements were resolved through discussion.

4 Results

Table 2 summarizes the utterance-level distribution of the participant codes across 58 participant utterances in S1 and 72 in S2. Inference was similarly prevalent in both sessions, appearing in over 90% of participant utterances in each. By contrast, observation and interpretation were both somewhat more frequent in S1 than in S2, and the auxiliary flags showed relatively small overall differences. These distributions indicate that the two sessions should not be contrasted as a less interpretive session and a more interpretive session.

Table 2. Utterance-level distribution of participant codes.

Session	observation	inference	interpretation	ext.reference	uptake
S1	31 (53.4)	53 (91.4)	11 (19.0)	12 (20.7)	17 (29.3)
S2	31 (43.1)	66 (91.7)	9 (12.5)	13 (18.1)	18 (25.0)

Using this exploratory segmentation, S1 comprised 19, 21, and 18 utterances per phase; S2 comprised 23, 26, and 23. In S1, observation and inference were both frequent in the early phase, and interpretation, ext.reference, and uptake became more prominent in the late phase. In S2, inference was already frequent in the early phase, but ext.reference peaked in the middle phase and interpretation was most frequent in the late phase.

Table 3 presents selected code combinations. The overall frequency of inference accompanied by observation was identical across sessions, but its phase-wise distribution differed. In S1, observation-supported inference was concentrated in the early phase; in S2, it was less frequent in the early and middle phases but became most prominent in the late phase. In S1, inference accompanied by ext.reference and uptake was most frequent in the late phase. In S2, inference accompanied by ext.reference peaked in the middle phase, whereas inference accompanied by uptake increased from the middle to the late phase.

This pattern is consistent with a qualitative reading of the two sessions. S1 began with sustained attention to visible features of the artwork and moved from observation to local inference. S2, by contrast, introduced Tokugawa Ieyasu as an interpretive resource near the beginning of the discussion, and the middle phase contained more inferences that drew on external knowledge or experience. The late phase of S2 then returned more strongly to visual features while also showing the highest concentration of interpretation.

Table 3. Phase-wise distribution of resources accompanying inference.

Session	Phase	inference + observation	inference + uptake	inference + ext.reference
S1	Early	14 (73.7)	5 (26.3)	3 (15.8)
	Middle	7 (33.3)	2 (9.5)	1 (4.8)
	Late	8 (44.4)	10 (55.6)	6 (33.3)
S2	Early	11 (47.8)	3 (13.0)	4 (17.4)
	Middle	4 (15.4)	6 (23.1)	7 (26.9)
	Late	14 (60.9)	9 (39.1)	2 (8.7)

5 Discussion

The findings suggest that the two sessions differed less in the overall amount of inferential activity than in the temporal organization and resources of meaning-making. S1 exemplified a visually grounded mode of art appreciation through discussion: in its early phase, viewers frequently paired observation with inference, moving from visible features to conjectures about the figure's emotion, bodily state, or situation. S2 showed a different trajectory. Tokugawa Ieyasu was identified by a participant near the opening, and external knowledge became a salient resource in the middle phase. This was followed by increased observation and interpretation in the late phase, suggesting that historically mediated inference was later reconnected with close looking.

For QE, this case illustrates how utterance-level coding and temporal segmentation can characterize discourse trajectories in small, contextualized datasets, even when corpus size limits tools such as ENA. Because the sessions involved different groups on consecutive days, the observed trajectories should be interpreted as case-specific patterns rather than session effects.

References

1. Yenawine, P.: *Visual Thinking Strategies: Using Art to Deepen Learning Across School Disciplines.* Harvard Education Press, Cambridge, MA (2013).
2. Arastoopour Irgens, G., Eagan, B.: The foundations and fundamentals of quantitative ethnography. In: Damşa, C., Barany, A. (eds.) *Advances in Quantitative Ethnography*. Communications in Computer and Information Science, vol. 1785, pp. 3–16. Springer, Cham (2023).
3. Shaffer, D.W., Ruis, A.R.: Is QE just ENA? In: Damşa, C., Barany, A. (eds.) *Advances in Quantitative Ethnography*. Communications in Computer and Information Science, vol. 1785, pp. 71–86. Springer, Cham (2023).
4. Hirano, T., Anzai, Y., Yamauchi, Y.: Taiwagata kansho no facilitation ni okeru joho teikyo no arikata [Uses of information in facilitation on art appreciation through discussion]. *Japan Journal of Educational Technology* 43(4), 285–298 (2020).
5. Housen, A.: Aesthetic thought, critical thinking and transfer. *Arts and Learning Research Journal* 18(1), 99–132 (2002).

Beyond Discourse Data: Positioning Observation Notes in Computational Quantitative Ethnography

Ayano Ohsaki[1][0000−0001−5072−0135] and Daniel Spikol[2][0000−0001−9454−0793]

[1] Tohoku University, Sendai, Miyagi 9808576, Japan
aohsaki@ohsaki-lab.net
[2] University of Copenhagen, DK-2200 Copenhagen, Denmark

Abstract. Observation notes are a foundational element of ethnography, yet how they are integrated into computational Quantitative Ethnography (QE) workflows remains largely implicit. A systematic search of ICQE papers (2019–2025) found only 15 that explicitly referenced field or observation notes, and only four addressed their role in data analysis. This study examines how explicitly positioning observation notes within computational QE analysis enriches the characterization of practices. Using data from an entrepreneurship hackathon, we apply KeyBERT (keyword extraction) across three corpus configurations (Observation notes, Discourse data, and combined D&O), hierarchical clustering, and line-level cross-referencing. Results show that 28 of 30 key phrases from Observation notes were unique to that dataset, and clustering revealed qualitatively distinct patterns absent from Discourse data alone. Integrating observation notes into computational QE workflows yields richer characterizations of practice and warrants broader methodological attention.

Keywords: Observation note · Coding process · Hackathon · Quantitative ethnography · Human-computer interaction.

1 Introduction

Within the Quantitative Ethnography (QE) community, ongoing discussions have focused on how to extract meaningful cultural insights from large-scale datasets while preserving the richness of ethnographic interpretation. A range of computational approaches have been examined toward this goal: e.g., large language model (LLM)-supported coding [10] and AI-assisted coding table construction designed to ensure the transparency and reflexivity of qualitative data analysis [7]. However, one foundational element of ethnography as systematized in cultural anthropology remains unaddressed in QE: observation notes. Field notes, including observation notes, capture culturally situated knowledge — observations and interpretations accessible only to those who understand the culture under study. Shaffer [8] positioned field notes as 'important parts of the mechanical grip of ethnography,' framing them as a means of capturing what

occurs in the world and rendering it amenable to analysis. This perspective was further consolidated by Arastoopour Irgens and Eagan [1] in their systematic articulation of the QE process, and continues to underpin QE research and practice more broadly.

A keyword-based search of full and short papers published at ICQE (2019–2025) using three terms to account for non-standardized terminology in the QE literature identified only 15 papers that explicitly mentioned at least one of: *field note* (n = 10; 2025: 3, 2022: 5, 2020: 2), *observation note* (n = 4; 2020: 2, 2019: 2), or *observation record* (n = 1; 2025: 1). Of these 15 papers, nine used the terms in general contexts such as citations or discussions of ethnographic approaches, and two mentioned them only in the data collection section. Only four papers explicitly addressed how field or observation notes were handled in relation to data analysis. Kim and Tscholl [6] identified observation notes as the primary analytical dataset. The remaining three employed them in a supporting role: to aid interpretation of multimodal data [3], inform the analytical process [5], or derive analytical content [2].

The overall paucity of explicit references to field notes and observation notes across this body of work suggests that their use has become so naturalized within our community that the specific ways in which they are employed tend to go unstated. This implicit mode of practice, however, poses a significant challenge to the goal of achieving reflexive and transparent analysis of large-scale data. Therefore, this study aims to examine how observation notes can be integrated into QE's computational workflows. Our research question (RQ) is: *How does explicitly positioning observation notes within computational QE analysis enrich the characterization of practices?*

2 Methods

Data were collected at an entrepreneurship hackathon in Estonia where Ukrainian refugees developed business ideas. A four-member team, selected from over 50 participants, gave informed consent. Discourse data (550 lines) were auto-transcribed, translated into English, and diarized; a bilingual researcher recorded 100 lines of timestamped observation notes. For details, see [9]. Our analytical approach combines key phrase extraction, similarity visualization, and line-level cross-referencing, examining the same dataset through both qualitative and quantitative lenses and tracing findings back to the source. To assess how observation notes affect key phrase extraction, we construct three corpus configurations: Observation notes, Discourse data, and combined (D&O). First, we apply KeyBERT [4] to each corpus configuration to extract the top 30 key phrases as candidates for characterizing practices. These key phrases are then visualized using a Ward-method dendrogram to reveal inter-phrase similarities within and across corpus configurations. To allow analysts to trace each key phrase back to its origin in the raw data, we also extract key phrases at the line level—the dataset's minimum unit—and cross-reference them against the corpus-level key phrase list derived in the first step.

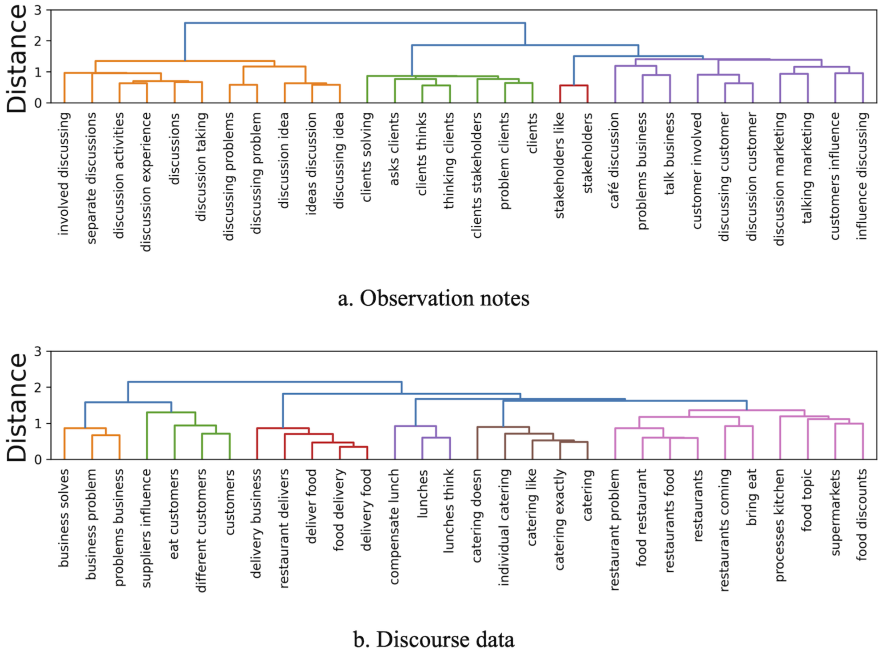


Fig. 1. Dendrograms of key phrases from Observation notes and Discourse data.

3 Results and Discussion

The results reveal that Observation notes provide complementary information not captured by Discourse data alone. Applying KeyBERT to the three corpus configurations, 28 of 30 key phrases extracted from Observation notes were unique to that dataset. Comparing Discourse data and D&O, 19 key phrases (63%) were shared with Discourse data, only 1 (3%) with Observation notes, and 10 appeared exclusively in D&O, likely because the volume imbalance shifted overall document characteristics in D&O toward Discourse data, influencing KeyBERT's output. Hierarchical clustering (Ward's method) revealed qualitative differences across datasets (Fig. 1). Observation notes yielded four clusters reflecting discussion dynamics, client interactions, stakeholders, and business contexts. Discourse data produced six domain-specific clusters centered on business, customers, catering, food delivery, lunch, and restaurants. Consistent with the quantitative findings, D&O yielded five domain-specific clusters, further supporting the interpretation that document characteristics shifted due to the volume imbalance. To close the interpretive loop, we traced key phrases back to the original data. Of 90 extracted instances, 69 (76%) appeared only once, indicating that extraction is not driven solely by term frequency. For example, 'café discussion,' extracted from Observation notes, traced to a member sharing their family's café experience, personal context that later significantly shaped

team activities. These results suggest that integrating observation notes into QE workflows yields findings that differ meaningfully from those of discourse analysis alone. How best to incorporate observation notes into computational QE warrants further investigation and could enrich the field more broadly.

Acknowledgments. This work was supported in part by JSPS KAKENHI Grant Number JP26K03044 and the Creative Impact Research Centre Europe (CIRCE).

Disclosure of Interests. The authors have no competing interests to declare that are relevant to the content of this article.

References

1. Arastoopour Irgens, G., Eagan, B.: The foundations and fundamentals of quantitative ethnography. In: Damşa, C.and Barany, A. (ed.) Advances in Quantitative Ethnography. pp. 3–16. Springer (2023)
2. Barany, A., Shah, M., Foster, A.: Connecting curricular design and student identity change: An epistemic network analysis. In: Ruis, A.R.and Lee, S. (ed.) Advances in Quantitative Ethnography. pp. 155–169. Springer (2021). https://doi.org/10.1007/978-3-030-67788-6_11
3. Buckingham Shum, S., Echeverria, V., Martinez-Maldonado, R.: The multimodal matrix as a quantitative ethnography methodology. In: Eagan, B., Misfeldt, M., Siebert-Evenstone, A. (eds.) Advances in Quantitative Ethnography. pp. 26–40. Springer, Switzerland (2019). https://doi.org/10.1007/978-3-030-33232-7_3
4. Grootendorst, M.: Keybert: Minimal keyword extraction with bert. (2020). https://doi.org/10.5281/zenodo.4461265
5. Jackowska, M.: How is team membership change manifested in collective interactions? – using epistemic network analysis to explore the challenges of contemporary team composition. In: Ruis, A.R.and Lee, S. (ed.) Advances in Quantitative Ethnography. pp. 292–303. Springer (2021). https://doi.org/10.1007/978-3-030-67788-6_20
6. Kim, Y., Tscholl, M.: The dynamic interaction between engagement, friendship, and collaboration in robot children triads. In: Eagan, B., Misfeldt, M., Siebert-Evenstone, A. (eds.) Advances in Quantitative Ethnography. pp. 307–314. Springer (2019). https://doi.org/10.1007/978-3-030-33232-7_27
7. Ohsaki, A., Kaneko, D.: Computer-assisted code generation using combination of generative artificial intelligence, stepwise coding, and topic modeling. In: Carmona, G., Lima, C., Santos, M.J., Benítez, H., Montero-Moguel, L., Galarza-Tohen, B. (eds.) Advances in Quantitative Ethnography. vol. 2677, pp. 302–317. Springer, Cham (2026). https://doi.org/10.1007/978-3-032-12229-2_20
8. Shaffer, D.W.: Quantitative Ethnography. Cathcart, Madison (2017)
9. Spikol, D., Li, Z., Nolte, A., Ohsaki, A., Rapur, K.: Investigating hackathons with collaboration analytics. In: Proceedings of the 8th International Conference on Game Jams, Hackathons and Game Creation Events. p. 1–8. ICGJ '24, Association for Computing Machinery, New York, NY, USA (2024). https://doi.org/10.1145/3697789.3697797
10. Zambrano, A.F., Liu, X., Barany, A., Baker, R.S., Kim, J., Nasiar, N.: From ncoder to chatgpt: From automated coding to refining human coding. In: Arastoopour Irgens, G., Knight, S. (eds.) Advances in Quantitative Ethnography. pp. 470–485. Springer (2023). https://doi.org/10.1007/978-3-031-47014-1_32

Calibrated LLM Coding in Quantitative Ethnography: A Comparison Study

James Mitchell[1][0009-0002-7274-7604], Curtis Isozaki[2][0009-0005-1375-8990], Amalfi Gayosso[3][0009-0009-1393-1027], Alison Lawton[4][0009-0006-1316-3900], & Liane Weintraub [5] [0009-0008-5944-8420]

[1-5] Pepperdine University, Malibu, CA 90263 USA
james.mitchell2@pepperdine.edu

Abstract. This study evaluated whether prompt calibration improves large language models' (LLMs) ability to perform theory-informed discourse coding within Quantitative Ethnography (QE) without distorting downstream Epistemic Network Analysis (ENA) representations. Uncalibrated and prompt-calibrated versions of ChatGPT and Gemini were compared with socially moderated human coding using the Party Line codebook comprising 18 codes across four domains of cross-partisan dialogue. Uncalibrated LLM conditions substantially undercoded relational and affective discourse constructs relative to human coders, particularly in areas involving outgroup warmth, shared humanity, and vulnerability. Calibration improved performance but did not fully eliminate interpretive divergence. Coding differences produced meaningfully different ENA network visualizations, with LLM-coded models underrepresenting relational connections, thus altering epistemic patterns. Findings suggest that current LLMs are not yet suitable replacements for human coders. However, iteratively calibrated human-AI collaboration may hold promise for scalable, reproducible QE discourse analysis.

Keywords: Quantitative Ethnography, Epistemic Network Analysis, Coding, Large Language Models, Artificial Intelligence,

Introduction

Quantitative Ethnography (QE) has emerged as a powerful methodological framework for integrating qualitative depth with quantitative rigor in the study of discourse [1]. QE transforms qualitative constructs into codes and, through Epistemic Network Analysis (ENA), models their co-occurrences to visualize patterns of sensemaking [2]. Traditionally, human coders apply a theory-informed codebook and use ENA to examine the structure of connections among codes, inferring how participants reason, frame issues, and construct relationships [1, 2]. While powerful, this approach is labor and resource-intensive and especially challenging with large datasets [3].

Artificial intelligence (AI) large language models (LLMs) have generated growing interest as coding assistants in qualitative and mixed-methods research [4, 5]. Recent studies have explored LLMs for codebook development, scalable reflexive and hermeneutic coding [4], deductive qualitative coding with structural interventions [6], and applied thematic or content analysis [7]. In these studies, LLMs have shown

promise in approximating human judgment and reducing manual coding time, with some research achieving levels of agreement comparable to benchmarks for human-to-human reliability [5].

However, most of this work has taken place in general qualitative settings, rather than in QE-specific settings that feed into ENA models [5, 8, 9]. This gap is likely because QE workflows impose structural constraints on segmentation, co-occurrence, and ENA model integrity that general qualitative LLM coding has not been designed to meet, test against, or validate. Furthermore, studies vary widely in the LLMs used, prompt design, and task configuration, and often provide only partial procedural information [7, 9]. Without a structured and replicable framework for LLM integration in QE, researchers risk introducing systematic coding errors. Within QE, such errors are not merely inaccuracies; they may distort the epistemic networks themselves. This raises a central methodological question: under what conditions can LLMs participate in coding workflows without compromising the interpretability of the resulting ENA models?

Politically charged discourse presents a particularly demanding context for LLM-assisted QE coding because its meaning often depends on tone, relational context, and subtle shifts in framing that experienced human coders are trained to detect through interpretive sensemaking [10]. In cross-partisan dialogue, the same utterance may signal warmth, irony, or deflection depending on conversational context, distinctions that carry significant interpretive weight in QE analysis and directly shape the structure of ENA models. Even trained human coders require detailed calibration and social moderation to reach a reliable agreement in exchanges [1, 2]. Therefore, cross-partisan political dialogue offers a useful stress test for examining the boundaries of LLM-assisted coding in QE.

This study analyzed transcribed conversations between politically diverse participants using the Party Line codebook, which operationalizes four domains of cross-partisan discord: outgroup orientation, relational humanization, identity framing, and escalation/polarization [10]. The study compared calibrated and uncalibrated versions of Gemini and ChatGPT LLMs against socially moderated human coding to evaluate the role of calibration in theory-informed discourse coding to examine (a) agreement patterns across discourse constructs, (b) the effects of calibration on LLM coding reliability, and (c) the consequences of coding variation on ENA network structure. The following research questions guided the study:

1. How do calibrated and uncalibrated LLM coding conditions converge with or diverge from socially moderated human coding when applying the Party Line codebook [10]?
2. To what extent does calibration/training improve interrater reliability between LLMs and human coders?
3. How do different LLMs differ in their coding tendencies, particularly regarding relational and affective discourse constructs?
4. How do differences in calibrated and uncalibrated LLM coding affect resulting ENA network structures?

1 Literature Review

1.1 Quantitative Ethnography Research Interrater Reliability

Interrater reliability (IRR) plays a central role in QE research, where coding serves as an ongoing interpretative process that involves constructing, organizing, and negotiating meaning across complex discourse data [1]. Because ENA relies on consistent, coherent coding to model relational structures in dialogue, the quality of coding and interpretive alignment is methodologically consequential. As LLM-assisted coding is increasingly integrated into QE research, there is a growing need to examine how AI systems may influence coding decisions, interpretation, and analysis of dialogue within QE methodologies [13].

1.2 AI-Augmented Qualitative Analysis & HITL Interpretive Partnership

In response to growing interest in applying AI to QE and qualitative analysis, recent scholarship has increasingly explored how LLMs can support coding workflows through natural language processing and AI-augmented reflexive analysis [12, 13]. Across this literature, LLMs are positioned as potential complements to human coders, assisting with large-scale pattern recognition, linguistic analysis, and code generation, while still requiring iterative refinement, structured workflows, and human oversight to preserve interpretive trustworthiness and rigor [9, 13].

The literature highlights the importance of prompt design, reliability, and validation in the responsible application of AI to qualitative data analysis [12, 13]. Using LLMs can achieve IRR levels comparable to those of human coders when guided by structured prompts, standardized coding, and clearly defined procedures, supported by human validation and training [5, 7, 8]. However, AI outputs remain dependent on prompt construction, design, and researcher supervision, particularly when balancing deductive consistency with inductive interpretation [5].

Rather than replacing human interpretation altogether, researchers advocate for human-in-the-loop (HITL) approaches that embed AI within reflexive qualitative workflows [4, 13, 16]. HITL approaches emphasize iterative codebook development, concise definitions, structured outputs, and opportunities for researchers to justify coding decisions while maintaining interpretive depth [4]. For example, the six-stage Hybrid Human-in-the-Loop Thematic Analysis framework integrates AI support across data familiarization, coding, theme generation, review, definition, and reporting, ensuring rigor and accountability within a replicable process [12].

Collectively, the literature suggests that AI-assisted qualitative analysis is most effective when positioned as a collaborative and reflexive research partner rather than as a fully autonomous coding system. Although AI tools can substantially reduce analytical time, these efficiencies introduce trade-offs in reliability and interpretation, reinforcing the value of hybrid HITL approaches that balance scalability with methodological rigor [12].

1.3 Reliability, Calibration, and Human Oversight in LLM-Assisted Qualitative Analysis

Through generative content analysis and discourse interpretation, LLMs have demonstrated variable reliability when contextual ambiguity, interpretive nuances, and socially constructed meaning are present, requiring careful validation and methodological transparency to prevent inconsistent outputs [14]. While chain-of-thought prompting, structured deductive coding frameworks, and interpretive reasoning frameworks can improve LLMs' interpretive coherence, the literature emphasizes LLM limitations such as hallucinations, subjective coding assumptions, embedded bias, inconsistent outputs, and a lack of reproducibility, which may undermine confidence in interpretively complex analyses [4, 6, 7, 9, 14]. Therefore, AI-assisted outputs are best aligned when standardized prompts, explicit codebook definitions, and iterative calibration are employed.

However, iterative prompting, deductive codebook alignment, and structured feedback loops can strengthen AI-assisted analysis, reinforcing the importance of human oversight for validity and accuracy [9]. These findings also underscore the importance of HITL oversight as an ethical practice, through transparent reporting, methodological validation, and explicit documentation of AI use [9].

Related work identified ongoing challenges involving code definitions, construct validity, and IRR in qualitative analysis [14]. These studies suggest that GPT-4 can be integrated with tools such as nCoder to support QE workflows by providing explanations that aid coding, social moderation, and interrater alignment [1, 15]. These findings are particularly relevant to QE because they position LLMs not as autonomous coders but as interpretive supports embedded within collaborative, iterative coding.

Building on these methodological directions, Prahl [16] advanced the AI-Reflexivity Checklist (ARC), which operationalizes reflective safeguards through pre-analysis and evidence-informed checkpoints for qualitative coding. ARC emphasizes methodological accountability through mechanisms such as deductive codebook alignment, contextual and interpretive prompting, ethical risk logging, and transparency safeguards, including version histories, audit trails, and raw data traceability [16].

Collectively, these methodological developments highlight the growing role of human-LLM research partnerships in QE through adaptive codebooks, iterative prompting, explanatory feedback loops, and human-led calibration. Comparisons between LLMs and nCoder further foreground reliability, ethical reflexivity, and governance as considerations for integrating LLMs into QE research practice. The present study builds on emerging research on structured prompting and interpretive reliability in AI-assisted analysis [8] by comparing human and LLM coding within a shared QE framework.

2 Methodology

2.1 Research Design and Coding Instrument

In this study, discourse patterns in cross-partisan conversational data were examined through structured coding and ENA modeling. The dataset consisted of six publicly shared recordings of conversations between strangers in Abilene, Texas, and San Francisco, California, selected because of their strong conservative (Abilene) and liberal (San Francisco) political associations [11]. The conversations occurred via an installation that linked two pay phones labeled "call a Republican" and "call a Democrat," inviting dialogue across presumed political differences [17]. Each participant's full utterance served as the unit of analysis. The procedure comprised four sequential phases: human consensus coding, uncalibrated LLM coding, prompt-calibrated LLM coding, and ENA modeling. The identical data, segmentation procedure, and coding framework were used across all conditions. The 18-construct codebook was developed for use in the Party Line study [10], including codes and constructs related to relational discourse, political identity, disagreement, escalation, and humanization in cross-partisan dialogue. Operational definitions and coding guidance for all constructs are shown in Table 1.

Table 1. Party Line Project: Coding Framework for Discourse Analysis

Domain	Construct	Operational Definition
Outgroup Orientation	Outgroup Warmth	Expression of positive affect toward the conversation partner or political outgroup
	Outgroup Hostility	Expression of negative affect, distrust, contempt, or moral suspicion toward the outgroup
	Hopefulness	Optimism about cross-partisan understanding, democracy, future of the country
	Hopelessness	Expression of pessimism about political division, democracy, or cross-partisan relations
Relational Humanization	Shared Humanity	Framing of self and other as fundamentally similar; recognition of commonality
	Gratitude	Expression of appreciation for the interaction or the other participant
	Curiosity	Genuine inquiry about the other person's experiences, views, or background
	Vulnerability	Disclosure of uncertainty, nervousness, sensitive personal info; emotional exposure
	Surprise at Similarity	Expression that similarity or agreement w/ the presumed outgroup was unexpected
	Social Dist. Reduction	Statements indicating reduced perceived distance between self and outgroup
Identity Framing	Political Ident. Disclosure	Self-identification of partisan/ideological affiliation
	Issue Positioning	Clear stance taken on a specific political issue
	Issue Disagreement	Acknowledgment of disagreement about politics

Domain	Construct	Operational Definition
	Difference Normalization	Framing political difference as natural, legitimate, or acceptable
	Political Lament	Expression of dissatisfaction with the political climate without specifying issues
Escalation/ Polarization	Sarcasm	Use of irony or mockery directed toward political actors or groups
	Cancelling	Delegitimizing, shunning, or rejecting the validity of a person, group, or viewpoint
	Moral Distancing	Framing the outgroup as morally inferior or flawed, creating an "us vs them" dichotomy

Human coding was conducted by four human researchers who collaboratively coded the first 50 turns of talk. The remaining xx lines were coded and socially moderated in pairs until consensus was established, a common QE practice in which coders discuss and resolve coding discrepancies to strengthen shared interpretive understanding [1]. The coding process required substantial discussion of theoretical grounding, operational definitions, and contextual application of constructs, particularly for relational or nuanced codes. For example, the codebook operationalized *curiosity* as "genuine inquiry about the other person's experiences, views, or background." Consequently, participants' inquiries unrelated to their conversational counterpart did not satisfy the construct definition.

2.2 LLM Coding Procedure

The LLM coding was conducted using ChatGPT 5.5 and Google's Gemini 3 Flash. Researchers initiated new accounts to ensure the LLM had no retained model memory or potential bias. LLM coding occurred in two iterations. The first pass was a tabula rasa, or uncalibrated, coding approach, where researchers uploaded only the codebook and the dataset without providing examples or further explanations and prompted the LLMs to assign 0 if the construct was not present and 1 if it was present for each of the 18 codes. Identical coding materials and procedural instructions were supplied across models.

In the prompt-calibrated condition, researchers supplemented the codebook with a 50-turn human-coded training set and interpretive guidance, including clarified coding rules and explicit instructions for relational and affective constructs. This was intended to test whether structured examples and rules improve prompt-based coding. The 50-turn training set and moving-stanza choices were selected from existing comparable research studies [4]. Table 2 displays the coding conditions used in the comparison study.

Table 2. Coding Conditions Used in the Human and AI Reliability Comparison Study.

Condition	Model	Calibration	Description
Human	n/a	Social moderation	Consensus
Uncalibrated	ChatGPT 5.5	None	Initial prompt; no calibration
Prompt-calibrated	ChatGPT 5.5	Iterative examples & clarified rules	Prompt-calibrated

Uncalibrated	Gemini 3 Flash	None	Initial prompt; no calibration
Prompt-calibrated	Gemini 3 Flash	Iterative examples & clarified rules	Prompt-calibrated

2.3 Interrater Reliability Assessment

Inter-rater reliability was assessed using Cohen's Kappa and Spearman's Rho [1, 18]. Kappa evaluated chance-corrected agreement between human consensus and LLM coding conditions, while Spearman's Rho assessed relationships among coding frequencies. Interpretation followed established conventions [18, 19].

2.4 Epistemic Modeling Procedure

Each turn of talk was treated as a unit of interaction, preserving the structural relationships among the constructs as they emerged in discourse [1]. The ENA Webtool was configured using a four-moving-stanza window because it is narrow enough to capture discourse focused on a topic and wide enough to incorporate utterances that built on the speakers' interaction. The resulting visualizations depict patterns among the coded constructs, with nodes representing discourse constructs and edges representing the relative strength of co-occurrences between constructs. Larger nodes and thicker edges indicated stronger or more frequent co-occurrence patterns. Comparative ENA models were generated for the human consensus benchmark, uncalibrated ChatGPT, prompt-calibrated ChatGPT, and prompt-calibrated Gemini conditions.

3 Findings

3.1 Agreement Metrics

Agreement analyses comparing the human consensus benchmark with each LLM condition are summarized in Table 3. Representative examples illustrate major patterns of agreement and disagreement. Agreement levels ranged from poor to moderate across most constructs, with stronger agreement in specific discourse constructs such as those involving explicit political or identity-related markers. For example, uncalibrated ChatGPT achieved strong agreement on *political identity disclosure* ($\rho = .73$, $\kappa = .73$) and *political lament* ($\rho = .66$, $\kappa = .65$), while calibrated ChatGPT showed strong agreement for *out-group hostility* ($\rho = .86$, $\kappa = .86$). Across many constructs, ChatGPT conditions demonstrated stronger agreement than Gemini conditions. For example, on *out-group warmth,* uncalibrated ChatGPT showed moderate agreement ($\rho = .51$, $\kappa = .48$), while calibrated Gemini showed significantly lower agreement ($\rho = .17$, $\kappa = .17$).

Table 3. Interrater Reliability Statistical Scores.

Code	Gemini Untrained Rho	Gemini Untrained Kappa	Gemini Trained Rho	Gemini Trained Kappa	GPT Untrained Rho	GPT Untrained Kappa	GPT Trained Rho	GPT Trained Kappa
Outgroup Warmth	0.4	0.34	0.17	0.17	0.51	0.48	0.33	0.32

Outgroup Hostility	-0.01	-0.01	NA	0	0.46	0.44	0.86	0.86
Hopefulness	0.2	0.19	0.12	0.1	0.52	0.52	0.51	0.48
Hopelessness	0.35	0.3	NA	0	0.26	0.25	0.44	0.33
Shared Humanity	0.21	0.14	0.19	0.19	0.4	0.32	0.22	0.2
Curiosity	0.26	0.2	0.26	0.26	0.29	0.25	0.26	0.2
Vulnerability	0.43	0.4	0.43	0.4	0.45	0.42	0.44	0.4
Surprise at Similarity	-0.03	-0.01	0.07	0.06	0.25	0.24	0.44	0.43
Social Dist. Reduction	0.08	0.08	NA	0	0.24	0.23	0.12	0.11
Political Ident Disclosure	0.2	0.14	NA	NA	NA	NA	0.35	0.33
Issue Positioning	0.54	0.54	0.69	0.69	0.73	0.73	0.59	0.56
Issue Disagreement	NA	0	NA	0	0.29	0.23	0.27	0.26
Difference Normalization	0.5	0.4	NA	NA	0.19	0.19	0.16	0.15
Political Lament	0.49	0.44	0.24	0.11	0.66	0.65	0.59	0.62
Sarcasm	NA	NA	NA	NA	0.19	0.19	0.28	0.28
Cancelling	1	NA	1	NA	NA	NA	1	1
Moral Distancing	NA	0	NA	0	0.4	0.28	-0.01	-0.01

Calibration effects were inconsistent across LLMs and constructs. In some cases, calibration increased alignment with human coding. For example, prompt-calibrated ChatGPT demonstrated stronger agreement for *vulnerability* (ρ = .44, κ = .43) than uncalibrated ChatGPT (ρ = .25, κ = .24). However, calibration reduced performance in other areas, such as *out-group warmth*, where uncalibrated ChatGPT (κ = .48) outperformed prompt-calibrated ChatGPT (κ = .32). Constructs involving relational nuance, emotional inference, and contextual interpretation consistently produced weaker agreement across all LLM conditions. Even in the strongest-performing condition, *shared humanity* produced only modest agreement (e.g., uncalibrated ChatGPT: ρ = .40, κ = .32). Kappa scoring was weak in both Gemini conditions (κ values between .14 and .19). Collectively, the findings suggest that LLM performance is strongest for explicitly coded discourse constructs, but more limited for relationally nuanced or emotionally subtle features.

3.2 Code Distribution Analysis

The human consensus coding identified a broader range of nuanced relational and emotional constructs than the LLM conditions. Frequency distributions provided a complementary view of coding behavior, beyond agreement metrics alone (see Table 4). The largest divergences from human coding appeared in socially contextualized constructs such as *curiosity*, *outgroup warmth*, and *shared humanity*.

Table 4. Frequency Comparison: Human vs LLM Coders.

Construct	Human	Calibrated ChatGPT	Uncalibrated ChatGPT	Calibrated Gemini	Uncalibrated Gemini
Outgroup Warmth	86	113	48	71	33
Outgroup Hostility	3	4	6	0	1
Hopefulness	13	7	13	4	6
Hopelessness	10	2	5	0	3
Shared Humanity	47	26	13	40	8
Gratitude	22	5	7	5	5
Curiosity	35	68	65	69	69
Vulnerability	26	20	15	10	2
Surprise at Similarity	10	5	6	0	8
Social Dist. Reduction	11	6	11	0	2
Political Ident Disclosure	9	5	10	0	9
Issue Positioning	11	16	39	0	0
Issue Disagreement	0	2	3	0	1
Difference Normalization	4	8	6	0	1
Political Lament	16	9	11	1	6
Sarcasm	4	3	6	0	0
Cancelling	0	0	2	0	0
Moral Distancing	1	3	6	0	0

Across many constructs, both prompt-calibrated and uncalibrated ChatGPT conditions generated higher coding frequencies than Gemini conditions. Calibration effects varied across LLM models and construct types. Calibration improved the detection of some relational constructs, although prompt-calibrated ChatGPT occasionally exceeded the human benchmark. The *outgroup warmth* construct illustrates these interpretive differences. Human coders often interpreted conversational openness as *outgroup warmth*. For example, the utterance, "What's up, Shane? How are you doing?" was coded by the research team as *outgroup warmth* because it reflected openness toward a presumed political out-group member. In contrast, uncalibrated Gemini often required more explicit affective cues before assigning the construct. Both LLMs appeared to assign the *curiosity* construct more broadly than the human coders did, using it as a catch-all for questions, conversational prompts, and friendly engagement that human coders recognized as not reflecting genuine curiosity about the conversational partner. Overall, the frequency patterns shown in Table 4 suggest that calibration improved sensitivity to some nuanced interpersonal constructs.

3.3 ENA Model Comparison

The analysis examines how coding variations across human and LLM conditions affect downstream ENA representations. Using identical procedures across conditions, Figs. 1-5 compare the networks for human, uncalibrated ChatGPT, prompt-calibrated ChatGPT, uncalibrated Gemini, and prompt-calibrated Gemini conditions. Across all five models, the human-coded networks display a more distributed structure with stronger connections among relational constructs. In contrast, LLM networks are narrower, with stronger clustering around *warmth* and *curiosity*. These patterns align with earlier agreement and frequencies, which show underrepresentation of relational connections relative to the human-coded benchmark.

Fig. 1 demonstrates a strong connection between *shared humanity* and *out-group warmth,* suggesting that human coders perceived participants as avoiding out-group hostility. In contrast, all LLM networks show a dominant association between *curiosity* and *out-group warmth*, consistent with earlier evidence that the LLMs assigned the *curiosity* code more broadly. Taken together, these differences suggest that coding variation carries epistemic consequences extending beyond agreement statistics. Differences in coding decisions impacted the relational architecture of the resulting ENA models and, consequently, the interpretive story implied by the networks. These findings highlight a key methodological issue for AI-assisted ENA workflows: even moderate differences in the application of constructs may meaningfully reshape epistemic implications.

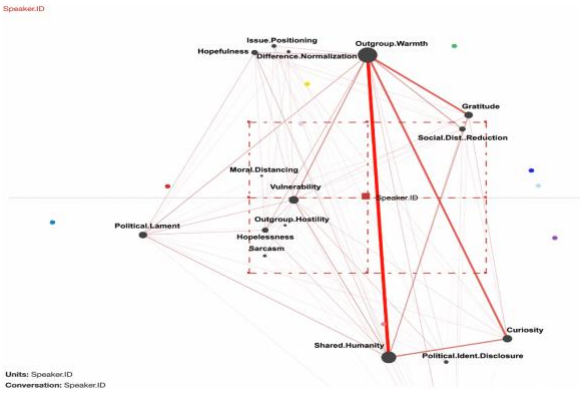


Fig. 1. Human Coded with Social Moderation & Consensus

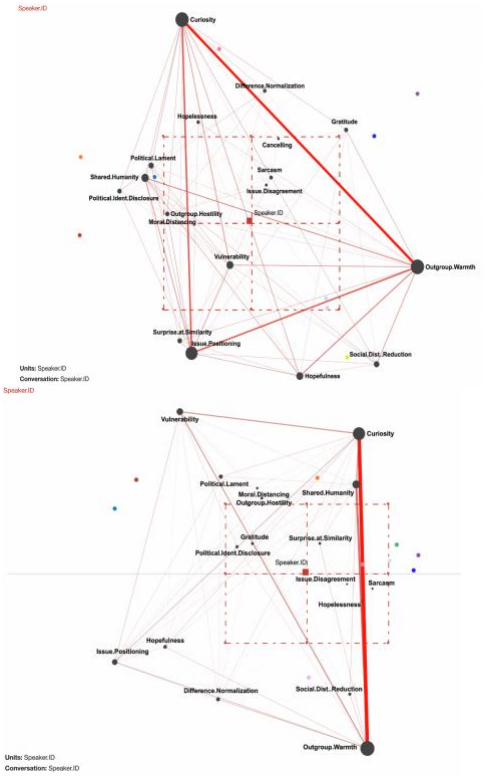


Fig. 2. Prompt-Uncalibrated and Calibrated ChatGPT



Fig. 3. Prompt-Uncalibrated and Calibrated Gemini

4 Discussion

4.1 Interpretation of Findings

Consistent with prior work [7, 8, 6], the findings suggest that LLMs can approximate human coders when the constructs are explicit and deductive; however, LLMs in this study struggled to reproduce the relational, emotional, or nuanced sense-making. Humans tended to interpret discourse relationally and contextually, whereas the LLMs relied more heavily on explicit semantic and textual cues. These differences suggest distinct LLM approaches to operationalizing meaning across coding conditions rather than simply automating human coding patterns. Accordingly, integrating LLMs into QE workflows may influence not only coding efficiency but also interpretive modeling.

Human and LLM coders converged on explicit constructs but diverged sharply on relational constructs. Humans and LLMs converged when coding d explicit or direct issue-related statements, such as *political lament* and *political identity disclosure*. However, they varied greatly in how they coded interpersonal or relational constructs such as *shared humanity, vulnerability, warmth,* and *gratitude,* which require contextual interpretation and the ability to infer emotional dispositions. In relationally dense discourse, QE LLM-assisted coding therefore involves more than simply applying a codebook. Furthermore, ChatGPT and Gemini showed distinct coding

tendencies, accuracy, and computational orientations, suggesting that LLM-assisted coding is not a homogenous category. Different LLMs appeared to apply varying thresholds for relational inference, contextual interpretation, and semantic coding. Consequently, model selection may become a consequential decision in LLM-assisted QE workflows.

Coding variances across the human, uncalibrated, and prompt-calibrated coding conditions substantially altered the resulting ENA models and the relational patterns and co-occurrences within the same dataset. The human-coded model depicts strong connections among *warmth, vulnerability, shared humanity*, and *reduced social distance*, resulting in a comparatively richer relational network. By contrast, the LLM-generated models are comparatively narrower in relational structure, more dependent on explicit in textual cues, and focus on literal inquiries. Consequently, LLM networks provided comparatively weaker representations of humanizing discourse constructs.

This study's key contribution is the suggestion that LLM calibration does not necessarily produce uniform improvements across models. Calibration effects depended on both the LLM used and the construct being coded. For several constructs, calibration improved agreement in one model but reduced it in the other, even with the same code. This indicates that calibration strategies are not model-agnostic and cannot be assumed to transfer between LLMs.

4.2 Implications for QE Research

This study highlights a limitation of current LLM coding in QE: that LLMs integration is not analogous to human social moderation, in which coders negotiate the meaning of codebook constructs and refine interpretations collectively and iteratively. Because social moderation serves as a mechanism for interpretive alignment, future work might explore whether iterative human-AI interaction, staged prompting, or smaller-segment coding approaches could approximate aspects of negotiated coding review. Future QE workflows may increasingly involve human-AI partnerships and HITL coding systems.

QE coding depends not only on codebook application, but also on interpretive judgment, contextual inference, and socially negotiated coding decisions [1, 2]. Currently, LLMs can effectively assist with certain deductive coding workflows; however, in this study, LLMs did not adequately replicate the relational coding patterns reflected in human-coded ENA networks. Consequently, human review remains crucial for relational codes that require high inference.

Although this study highlights discrepancies between LLM and human coding, AI shows considerable potential as a scalable support tool. Manual coding is time-intensive, especially with large datasets, and LLMs offer advantages in speed, repeatability, and scalability. If improved calibration and iterative guidance prove effective in mitigating interpretive discrepancies within HITL workflows, LLMs may function effectively as a first-pass coding layer, applying codebooks across large datasets and generating coding suggestions, while human coders retain responsibility for relational, nuanced, or emotionally inferred meaning-making. In such contexts, calibrated LLM systems may operate less as replacements for human interpretation than as force multipliers supporting discourse analysis.

4.3 Limitations and Future Research

Approximating social moderation under LLM conditions is a limitation of this study. Because social moderation serves as a central mechanism in QE [1], its absence limits direct comparisons between human and LLM coding. The study also relied on only two LLMs, limiting generalizability across AI systems. The reported findings, therefore, reflect specific behaviors of ChatGPT and Gemini rather than universal properties of LLM-assisted coding. Additionally, because of the single-discourse dataset and coding framework, findings may not transfer to other settings or studies.

LLM coding behavior is highly dependent on instructional procedures, as prompt structure can substantially affect coding outputs, and iterative refinement can improve coding quality [19]. Accordingly, outcomes in LLM-assisted workflows may depend not only on codebook definitions but also on calibration strategies and procedural guidance, raising questions regarding reproducibility, reporting transparency, and methodological standardization. Additionally, calibration in this study was prompt-based instead of fine-tuning (i.e., adjusting model weights on the 50-turn training set). This made the approach more accessible but also a less detailed intervention than a durable change to the model itself. An API was not used to reflect common practices as most QE researchers use consumer interfaces. Temperature and decoding parameters were not controlled, adding stochastic variability on top of model differences.

Future work should compare additional LLM systems using controlled datasets, codebooks, and prompts to clarify whether observed coding discrepancies are model-specific or more broadly characteristic of LLM-assisted QE coding. Research should also investigate whether structured human-AI interaction, iterative calibration, multi-LLM review, or critique-and-revision workflows can approximate aspects of social moderation and negotiated refinement.

5 Conclusion

As LLM systems evolve, continued evaluation of AI-assisted coding remains important for QE research. This study examined whether calibrated and uncalibrated conditions for ChatGPT and Gemini could approximate human QE coding without compromising interpretive alignment or downstream ENA modeling. Across conditions, LLMs showed weaker detection of relational, affective, and contextually inferred constructs than the human consensus benchmark, while applying simpler constructs such as *curiosity* more broadly. More importantly, coding variation materially altered ENA structures, suggesting that differences in how constructs are applied can reshape relational architecture and interpretation in QE workflows.

Despite these limitations, LLM-assisted coding retains considerable promise through calibrated workflows, HITL systems, and approaches that approximate human coders. a. Because QE depends on interpretive meaning-making, the most promising role for AI in QE may be as a collaborative interpretive partner to help surface and make sense of complex relational, affective, and contextually inferred dimensions of discourse central to human meaning-making.

Disclosure of Interests. The authors have no competing interests to declare that are relevant to the content of this article.

References

1. Shaffer, D.W.: Quantitative Ethnography. Cathcart Press (2017)
2. Shaffer, D.W., Ruis, A.R.: Epistemic Network Analysis: A worked example of theory-based learning analytics. In: Handbook of Learning Analytics, pp. 175–187. Society for Learning Analytics Research (2017)
3. Herrenkohl, L.R., Cornelius, L.: Investigating elementary students' scientific and historical argumentation. J. Learn. Sci. **22**(3), 413–461 (2013)
4. Dunivin, Z.O.: Scaling hermeneutics: a guide to qualitative coding with LLMs for reflexive content analysis. EPJ Data Sci. **14**(1), 28 (2025)
5. Zhang, H., Wu, C., Xie, J., Rubino, F., Graver, S., Cai, J., Kim, C., Carroll, J.: Exploring inductive and deductive qualitative coding with AI: investigating inter-rater reliability between large language model and human coders. In: Ahram, T.Z., Karwowski, W., Rau, P.L. (eds.) Human-Computer Interaction & Emerging Technologies. AHFE Open Access, vol. 195. AHFE International (2025)
6. Hila, A., Hauser, E.: Assessing the reliability of large language models for deductive qualitative coding: a comparative intervention study with ChatGPT. Proc. Assoc. Inf. Sci. Technol. **62**, 275–285 (2025)
7. Wachinger, J., Bärnighausen, K., Schäfer, L.N., Scott, K., McMahon, S.A.: Prompts, pearls, imperfections: comparing ChatGPT and a human researcher in qualitative data analysis. Qual. Health Res. **35**, 951–966 (2024)
8. Mehta, S., Paul, S., Awiti, E., Young, S., Zulaika, G., Otieno, F., Phillips-Howard, P., Mason, L., Bhaumik, R.: Evaluation of large language models within GenAI in qualitative research. Sci. Rep., **15** (2025)
9. Tai, R.H., Bentley, L.R., Xia, X., Sitt, J.M., Fankhauser, S.C., Chicas-Mosier, A.M., Monteith, B.G.: An examination of the use of large language models to aid analysis of textual data. Int. J. Qual. Methods, **23** (2024)
10. Weintraub, L., Lee, S., & Hamilton, E.: Affective depolarization in cross-partisan dialogue: An epistemic network analysis of the Party Line conversations. In *Advances in quantitative ethnography*. Springer [Forthcoming, 2026]
11. DeSilver, D.: Chart of the week: the most liberal and conservative big cities. Pew Research Center (2014)
12. Saldaña, J.: The Coding Manual for Qualitative Researchers, 4th ed. SAGE (2021)
13. Wiebe, J.P., Khan, R., Burns, S., Slotta, J.D.: Qualitative research in the age of LLMs: a human-in-the-loop approach to hybrid thematic analysis. In: Proceedings of the International Conference of the Learning Sciences (ICLS 2025), pp. 1123–1131 (2025)
14. Rughiniș, C., Matei, Ș., Corcaci, A.: Generative content analysis for policy research: comparing LLM reliability in analyzing institutional AI discourse. In: Proceedings of the 25th Intl. Conf. on Control Systems and Computer Science (CSCS 2025), pp. 596–603 (2025)
15. Zambrano, A.F., Liu, X., Barany, A., Baker, R.S., Kim, J., Nasiar, N.: From nCoder to ChatGPT: From automated coding to refining human coding. In: Arastoopour Irgens, G., Knight, S. (eds.) ICQE23. CCIS, vol. 1895, pp. 470–485. Springer, Cham (2023)
16. Prahl, A.: The AI-reflexivity checklist (ARC): A pre-analysis pause for LLM-assisted coding. Qual. Health Res. **36**(2–3), 181–190 (2026)
17. Matter Neuroscience. (n.d.): The party line project. https://matter.xyz/partyline
18. Landis, J.R., Koch, G.G.: An application of hierarchical kappa-type statistics in the assessment of majority agreement among multiple observers. Biometrics **33**, 363–374 (1977)
19. Ornelas, T., Araújo, A.A., Araújo, J., Araújo, M., Trinkenreich, B., Kalinowski, M.: LLM-assisted thematic analysis: opportunities, limitations, and recommendations. arXiv (2025)

The Codebook Teaches Scope: Coding Student Experience with Retrieval-Augmented LLMs

Kaixiang Kang[1,2][0009−0001−9124−8073], Ran Nakao[3][0000−0002−5772−972X], and Yizhou Fan[1][0009−0004−5690−1227]

[1] Hiroshima University, Hiroshima, Japan
[2] Prefectural University of Hiroshima, Hiroshima, Japan
[3] Ehime University, Matsuyama, Ehime, Japan
kougaisyou@pu-hiroshima.ac.jp

Abstract. Student free-text responses show how students describe campus life and their own role in it. Manual coding is useful for this work, but it does not scale well. We build a retrieval-augmented LLM coding pipeline for quantitative ethnography (QE). The pipeline represents each response as one or more aspect–stance pairs and uses previously coded similar cases as local codebook examples. The benchmark generation model is `deepseek-v4-flash`.
A three-condition within-sample ablation on 300 SemEval-2014 sentences examines what retrieved examples contribute to aspect–polarity coding. Exact-pair F1 was .5987 with anchors only, .6384 with polarity-corrupted retrieval, and .6630 with full retrieval; the full–corrupted contrast was inconclusive, and the design does not isolate scope from other example features. In the Japanese student responses, the institution-focused codebook excluded personal goals and self-directed actions. Leaving these uncoded followed the codebook rather than constituting model errors. A separate responsibility-locus dimension could capture this distinction but requires independent validation. The findings suggest that retrieval-augmented coding can make codebook boundaries inspectable while researchers decide whether they fit the research question.

Keywords: Retrieval-augmented generation · Large language models · Qualitative coding · Student experience

1 Introduction

For QE, coding determines what enters the analysis, unit boundaries, and stance [1]. These decisions determine the coded events and co-occurrences available to epistemic network analysis (ENA) [2], so changing a coding boundary can change the representation even when source text is unchanged.

An a priori codebook defines the analytical perspective [3], which a large language model (LLM) applying it should follow. An uncoded response may therefore conform to the codebook while exposing a mismatch between its scope and the research question. We examine this through a SemEval mechanism audit and a Japanese student-response case.

Table 1. Results for 300 laptop-review sentences. Pair scores require an exact aspect–polarity match.

Condition	Pair P	Pair R	Pair F1	Aspect F1	Exact-set acc.
Anchor-only	.5909	.6067	.5987	.6784	.4133
Polarity-corrupted	.6242	.6533	.6384	.7402	.4600
Full retrieval	.6572	.6689	.6630	.7580	.5133

2 Method and Ablation

The pipeline returns aspect–stance pairs with positive, neutral, or negative stance (SemEval: *polarity*). A sentence encoder retrieves six coded examples by cosine similarity [6]. These examples, fixed anchors, and a coding instruction prompt the LLM; anchors stabilize the task and format, while retrieval supplies local precedents [4].

Using a deterministic label-balancing heuristic (seed 42), we selected 300 usable laptop-review sentences (450 normalized gold pairs) from the public SemEval-2014 validation split [5]. Because the sample and five fixed anchors had already been used in development and evaluation, this is a within-sample mechanism audit rather than a held-out estimate. Filtering and deduplication left 1,038 of 1,060 public training rows; removing two exact overlaps yielded 1,036 candidates, none with normalized-string similarity at or above .95 to a sampled sentence.

Using all-MiniLM-L6-v2, we retrieved six examples once per target and fixed their order across conditions. *Anchor-only* omitted them. *Polarity-corrupted* retained text, aspects, evidence, IDs, counts, and order but changed all 2,693 displayed labels. *Full retrieval* retained gold labels. Thus, the last two conditions differed only in displayed polarity, while both differed from anchor-only in the full retrieved set.

The API model `deepseek-v4-flash` generated JSON with low thinking effort and a 4,096-token cap [7]. All 900 runs completed; seven truncated first attempts were retried, and one completed generation was retained per target–condition pair. Outcomes were exact-pair micro precision, recall, and F1; aspect F1; and exact-set accuracy. We computed 95% paired percentile intervals from 10,000 sentence resamples (seed 20260825), conditional on the selected sample and completed generations.

The differences in the Pair-F1 were +.0398 for the corrupted anchor (95% interval [.0052, .0742]), +.0643 for the full anchor ([.0268, .1028]) and +.0246 for the full corrupted anchor ([−.0125, .0613]; Fig. 1). The last interval included zero, so the full–corrupted difference was inconclusive.

The pre-specified retention ratio was .618 (95% interval [.103, 1.319]); its lower bound did not exceed .50 and its point estimate did not reach .75. The denominator can approach zero, so we make no percentage claim; the design also does not isolate scope from text, evidence, format, or prompt length.

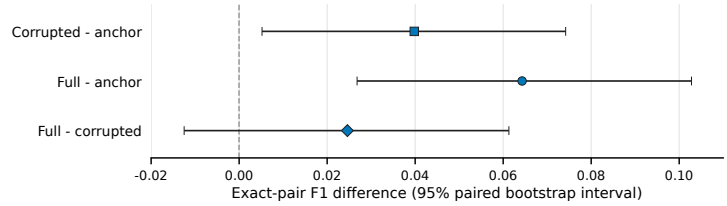


Fig. 1. Contrasts in exact-pair F1 with 95% paired sentence-bootstrap intervals.

3 Japanese Student Responses: Errors Reveal a Missing Category

The extract contains 230 complete response rows but no response or survey-item IDs, so student and item counts and student-level ENA cannot be reconstructed. The codebook covered evaluations of institution-provided services, facilities, policies, courses, and support, while deliberately excluding personal goals and self-directed actions. This scope fits an analysis of university services, but not every account of university life.

The files record the API request alias `deepseek-reasoner`. On the 31 May 2026 run date, provider documentation mapped this alias to the thinking mode of `deepseek-v4-flash`. The files do not preserve a resolved model ID or system fingerprint [7]. The prompt used paraphrase-multilingual-MiniLM-L12-v2, 10 anchors, and six examples retrieved from a 31-example development corpus. Because example selection was not fully documented, we treated 40 exact source matches as development/reference overlap and the remaining 190 responses as prompt-held-out.

The prediction-visible review sheet contained 284 expanded coding rows: one for each predicted aspect and one placeholder when no aspect was returned. One judgment was blank. A researcher judged 243 of the remaining 283 rows strictly correct. Thus, 85.9% describes a prediction-visible row-level review, not response-level accuracy or agreement with an independently coded reference set.

Uncoded examples included, in translation, "participate actively in various activities" and "balance study and club activities." Their omission followed the codebook; both were also development examples. A responsibility-focused study could instead define a response-level *responsibility-locus* dimension (institution, student-self, mixed, no-content, or ambiguous), distinct from aspect–stance. This proposal is not independently validated.

4 Discussion and Conclusion

Retrieval-augmented LLM coding can expose coding boundaries while leaving interpretation to researchers [8]. Retrieval improved the SemEval point estimate, but the design does not identify which part of the example bundle produced the gain.

In the Japanese case, an institution-focused codebook appropriately omits self-directed actions; a broader responsibility question requires a researcher-defined and validated dimension. Such a change could alter frequencies, co-occurrences, and ENA networks without changing source text.

This study has limits: one previously analyzed sample, one model configuration, and one generation per target–condition pair; its intervals omit generation and provider variability. The prediction-aware, single-researcher Japanese review lacked IDs for network comparison and preserved only the request alias. Validating the proposed dimension requires at least two coders applying frozen rules independently to prompt-held-out responses while blind to model outputs. The benchmark does not validate Japanese predictions or cross-language transfer.

Retrieved examples can execute a fixed coding task and expose its boundaries, but researchers must define the lens and validate changes before downstream QE analysis.

References

1. Shaffer, D.W., Ruis, A.R.: How we code. In: Ruis, A.R., Lee, S.B. (eds.) Advances in Quantitative Ethnography, CCIS, vol. 1312, pp. 62–77. Springer, Cham (2021). https://doi.org/10.1007/978-3-030-67788-6_5
2. Shaffer, D.W., Collier, W., Ruis, A.R.: A tutorial on epistemic network analysis: Analyzing the structure of connections in cognitive, social, and interaction data. *Journal of Learning Analytics* **3**(3), 9–45 (2016). https://doi.org/10.18608/jla.2016.33.3
3. Hsieh, H.-F., Shannon, S.E.: Three approaches to qualitative content analysis. *Qualitative Health Research* **15**(9), 1277–1288 (2005). https://doi.org/10.1177/1049732305276687
4. Lewis, P., et al.: Retrieval-augmented generation for knowledge-intensive NLP tasks. In: Advances in Neural Information Processing Systems, vol. 33, pp. 9459–9474 (2020)
5. Pontiki, M., Galanis, D., Pavlopoulos, J., Papageorgiou, H., Androutsopoulos, I., Manandhar, S.: SemEval-2014 Task 4: Aspect based sentiment analysis. In: Proceedings of SemEval 2014, pp. 27–35 (2014). https://aclanthology.org/S14-2004/
6. Reimers, N., Gurevych, I.: Sentence-BERT: Sentence embeddings using Siamese BERT-networks. In: Proceedings of EMNLP-IJCNLP 2019, pp. 3982–3992 (2019). https://aclanthology.org/D19-1410/
7. DeepSeek: DeepSeek V4 preview release. API documentation (24 April 2026). https://api-docs.deepseek.com/news/news260424/
8. Liu, X., Zhang, J., Barany, A., Pankiewicz, M., Baker, R.S.: Assessing the potential and limits of large language models in qualitative coding. In: Advances in Quantitative Ethnography, CCIS, vol. 2278, pp. 89–103. Springer, Cham (2024). https://doi.org/10.1007/978-3-031-76335-9_7

When Construct Complexity Extends Beyond Connections: Quantifying Recurrence in Affect Dynamics

Seiyon Lee[1 [0009−0001−9041−3313]], Anthony Botelho[1 [0000−0002−7373−4959]], and Daryn Dever[1 [0000−0002−9150−1155]]

Educational Technology, School of Teaching and Learning, College of Education, University of Florida, Gainesville, FL, USA

Abstract. Quantitative Ethnography (QE) methods such as Epistemic Network Analysis and Ordered Network Analysis examine relations among constructs by accumulating connections within a specified window of context. A separate temporal question is whether a state returns and whether those returns form repeated sequences. We use auto-Recurrence Quantification Analysis to examine these questions in affect, whose relationship with learning is not fully explained by occurrence alone. Recurrence rate, determinism, and entropy were calculated separately for concentration, boredom, confusion, and frustration. Each measure was first calculated using the first 10 observations and recalculated after each subsequent observation through observation 45. Differences between the learning groups were generally larger when more observations were included. Group means differed at shorter sequence lengths for recurrence rate than for entropy, while determinism showed that recurrent observations formed repeated sequences in both groups. The findings show how recurrence can extend the temporal questions examined in QE and inform future comparisons with established QE methods.

Keywords: Quantitative Ethnography · Affect Dynamics · Recurrence Quantification Analysis · Temporal Dynamics

1 Introduction

Quantitative Ethnography (QE) aligns theory with data collection, analysis, and representation to examine relationships among constructs in context [7]. Epistemic Network Analysis (ENA) models co-occurrences among coded constructs, while Ordered Network Analysis (ONA) incorporates directionality into those connections [8]. Both accumulate connections within a specified window of context. Recurrence addresses another temporal property by examining whether observations return and how repetitions are organized across the analyzed sequence. Two students may spend similar amounts of time in confusion, for example, while one experiences a sustained episode and the other repeatedly enters and exits the state.

Understanding affect requires a broader temporal account because its relationship with learning may not be adequately represented by occurrence or moment-to-moment transitions alone. Confusion may support learning when productively regulated but become detrimental when unresolved [4]; frustration can accompany either continued effort or disengagement [3]. Affective chronometry similarly emphasizes how states emerge, persist, and dissipate over time [1]. Examining whether a state returns and forms repeated sequences may therefore clarify differences that remain ambiguous in local transitions. Building on recurrence research on confusion and frustration [2], we analyze four affective states while increasing the sequence length from 10 to 45 observations.

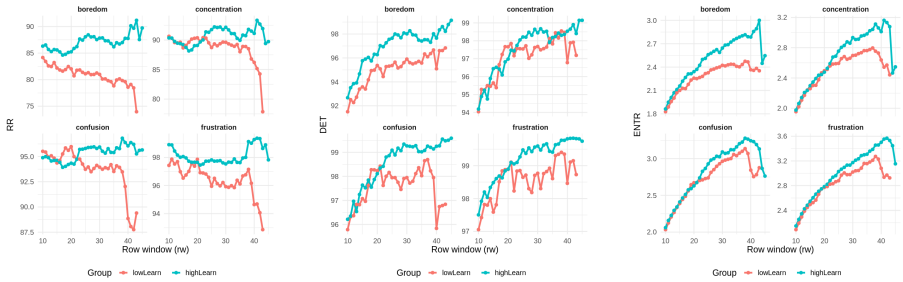
Recurrence Quantification Analysis (RQA) compares observations at every pair of time points and marks sufficiently similar pairs in a recurrence matrix [9]. Auto-RQA (aRQA) compares a series with itself. Recurrence rate (RR) describes the extent of repetition, determinism (DET) indicates how much recurrence forms repeated sequences, and entropy (ENTR) describes variation in the lengths of those sequences [5]. Applying these measures to concentration, boredom, confusion, and frustration shows how QE researchers can examine recurrence as a temporal property distinct from co-occurrence and directionality.

2 Method

Data came from 791 seventh-grade students across 13 Maine schools in an ASSISTments efficacy trial [1]. End-of-year TerraNova mathematics scores defined highLearn (above the full-sample median) and lowLearn (at or below it) groups before filtering. We retained students with at least 10 consecutive affect observations, no internal missing observations, and sessions no longer than 45 minutes. The resulting sample included 190 students (highLearn $n = 96$; lowLearn $n = 94$). Analyses began with the first 10 observations and added one observation at a time through observation 45.

Affect was inferred using published interaction-based detectors trained with observations collected through the *Baker Rodrigo Ocumpaugh Monitoring Protocol* [6]. Detector probabilities were converted into binary indicators using state-specific thresholds calibrated to approximate training-data prevalence [1]. The corresponding rates were 90.2% for engaged concentration, 10.0% for boredom, and 4.0% each for confusion and frustration. Values above each threshold were coded active (1) and all others inactive (0).

We conducted a separate aRQA for each student's state-specific binary series using the `crqa` package. Because the series preserved observed state order, embedding dimension and delay were set to 1, without rescaling or normalization. Euclidean distance and radius 0 defined recurrence as an exact match. A minimum diagonal length of 2 required two consecutive matches to form a repeated sequence. RR, DET, and ENTR were averaged within group for each sequence length from 10 to 45 observations.



(a) Recurrence Rate (RR) (b) Determinism (DET) (c) Entropy (ENTR)

Fig. 1: Mean (a) RR, (b) DET, and (c) ENTR for highLearn (blue) and lowLearn (red) students across sequence lengths and affective states.

3 Results and Discussion

Figure 1 shows that differences between group means generally increased with sequence length, although the point at which they appeared varied by state and metric. For RR, boredom differed earliest. HighLearn remained near 85–90%, while lowLearn declined from approximately 84% toward 79%. Concentration was similar across groups at shorter sequence lengths, but highLearn remained near 89–91% as lowLearn declined. Mean RR for confusion and frustration began to differ mainly after approximately 20–25 observations, with highLearn remaining relatively stable and lowLearn decreasing.

DET increased for both groups as matching points extended into diagonal sequences. Differences were smaller than for RR and appeared most consistently for boredom, confusion, and frustration, while concentration remained closely aligned. The high DET values indicate that recurrent points generally belonged to repeated sequences rather than isolated matches. Mean ENTR was nearly identical between groups through approximately 20 observations. After that point, highLearn increased more than lowLearn across all four states, with the clearest differences for boredom, concentration, and frustration. RR therefore identified differences in the amount of repetition earlier, DET described whether matches formed sequences, and ENTR identified later differences in the range of sequence lengths.

The differences that appeared after more observations were included are relevant to affective chronometry. Confusion and frustration are difficult to interpret because either can accompany productive persistence or unresolved difficulty [3, 4]. Values for these states were similar between groups when only the first 10–20 observations were included and differed more after additional observations were added. Boredom differed earlier. Although the results do not explain why these states differ, they show that conclusions about affect can depend on how much of the session is analyzed and whether a state returns in repeated sequences.

Recurrence extends the temporal questions available to QE. ENA and ONA preserve relationships among constructs within a specified window of context, whereas aRQA summarizes repetitions across the full sequence being analyzed. Future comparisons using the same data can examine whether these methods provide different information. Beyond affect, recurrence could characterize discourse, collaborative problem solving, regulation, identity work, or strategy use. Two groups may show similar connections among strategies while differing in whether a strategy occurs in one sustained episode or returns throughout an activity.

Applying aRQA to binary codes also identifies useful methodological extensions. Two observations count as matching when the affective state is present at both time points or absent at both time points. For infrequent states such as confusion and frustration, most matches involve their absence. Their high absolute RR therefore describes repetition across the complete binary sequence and should not be read as frequent reappearance of confusion or frustration alone. Comparisons with shuffled or prevalence-matched sequences can estimate how much recurrence exceeds the amount expected from the state's overall prevalence. Inferential tests and continuous learning outcomes would further evaluate the descriptive group patterns. Cross- and multidimensional RQA could then examine whether confusion and frustration recur independently or together. These extensions would allow QE researchers to compare relationships among codes with how individual codes repeat over time.

References

1. Botelho, A.F., Baker, R.S., Ocumpaugh, J., Heffernan, N.T.: Studying affect dynamics and chronometry using sensor-free detectors. International Educational Data Mining Society (2018)
2. Cloude, E.B., Munshi, A., Andres, J.A., Ocumpaugh, J., Baker, R.S., Biswas, G.: Exploring confusion and frustration as non-linear dynamical systems. In: Proceedings of the 14th learning analytics and knowledge conference. pp. 241–252 (2024)
3. D'Mello, S., Graesser, A.: Dynamics of affective states during complex learning. Learning and Instruction **22**(2), 145–157 (2012)
4. Lehman, B., D'Mello, S., Graesser, A.: Who benefits from confusion induction during learning? an individual differences cluster analysis. In: International conference on artificial intelligence in education. pp. 51–60. Springer (2013)
5. Marwan, N., Romano, M.C., Thiel, M., Kurths, J.: Recurrence plots for the analysis of complex systems. Physics reports **438**(5-6), 237–329 (2007)
6. Ocumpaugh, J.L.: Baker rodrigo ocumpaugh monitoring protocol (bromp) 2.0 technical and training manual. (2015)
7. Shaffer, D., Ruis, A.: Epistemic network analysis: A worked example of theory-based learning analytics. Handbook of learning analytics (2017)
8. Tan, Y., Swiecki, Z., Ruis, A.R., Shaffer, D.: Epistemic network analysis and ordered network analysis in learning analytics. In: Learning analytics methods and tutorials: A practical guide using R, pp. 569–636. Springer (2024)
9. Webber Jr, C.L., Marwan, N., Facchini, A., Giuliani, A.: Simpler methods do it better: Success of recurrence quantification analysis as a general purpose data analysis tool. Physics Letters A **373**(41), 3753–3756 (2009)

Comparing STEM Identity Connections Among Students in the USA and Kenya

Nastassia N. Jones[1], Anna Wilson[1], Danielle Espino[2], Seung Lee[2], Eric Hamilton[2]
[1]Southern University, Baton Rouge, LA 70807, USA
[2]Pepperdine University, Malibu, CA 90263, USA
nastassia.jones@sus.edu

Abstract: This paper examines how adolescents' STEM experiences and relationships shape their interest, competence, recognition, and future STEM aspirations. This study analyzes interviews with 11 learners in the USA and Kenya participating in an informal STEM learning program. For USA students, doing STEM appeared central to seeing themselves in STEM while for Kenya students, STEM experiences were tied to enjoyment and future career direction. Across both groups, STEM pathways were strengthened when interest was supported by active participation and external STEM validation. These findings highlight the need for STEM programs that offer hands-on experiences, opportunities to demonstrate competence, leadership roles, and affirming feedback from adults, peers, and STEM professionals.

Keywords: STEM Identity, Informal Learning, Epistemic Network Analysis, Cultural nuances

1 Introduction

A redefinition of the sense of self is a hallmark of the adolescent years, a foundational period for identity development, shaped by momentous cognitive, biological, and social transitions [1]. In terms of STEM (Science, Technology, Engineering, and Mathematics) identity development, competence, performance, and recognition as a science person have been identified as central domains [2], while more recent work highlights that intrinsic interest and a sense of belonging are critical components [3, 4]. Within informal learning environments, youth can extend academic content beyond formal classroom settings while also exploring diverse roles, developing agency, and building resilience through peer interaction [5, 6]. This study seeks to explore the role of an informal STEM learning program in shaping the STEM identity of the participants. Specifically, it examines STEM identity development as a relational and experiential process, exploring how adolescents' STEM experiences and interactions with peers, mentors, and educators shape their interest, competence, recognition, sense of belonging, and future STEM aspirations.

2 Methods

2.1 Data Collection

This study analyzes data collected from pre- and post-semi-structured interviews of 11 participants in the USA and Kenya. Interviews focused on students' STEM experiences, interests, confidence, participation, and future goals. Transcripts were reviewed for accuracy across two rounds by researchers and segmented by turn-of-talk [7]. Interview excerpts were coded using a STEM identity-focused codebook in Table 1, which included six codes capturing students' STEM interests, aspirations, self-recognition, competence, performance, and external recognition.

Table 1. Codebook.

Code	Definition	Example
STEM Interest	Expressions of enjoyment, curiosity, excitement, or intrinsic motivation toward STEM subjects, topics, or activities.	"I really enjoy doing math." (Eunice)
STEM Aspirations	A desired future in STEM, including plans for further study, a STEM career, or a STEM-related life path.	"I want to pursue nuclear engineering…and robotics." (Eunice)
STEM Self-Recognition	Statements in which students identify themselves as a STEM/science/math/technology person or position STEM as central to who they are.	"I'm a science student so I'm interested in science." (Paxton)
External STEM Recognition	Recognition by teachers, peers, family, or institutions that positions the student as capable, talented, or belonging in STEM.	"I'm part of the quiz team in my school…and if you are part, it means you are very good." (Paxton)
STEM Competence	Perceived understanding, skill, confidence, or proficiency in STEM content, practices, or problem solving.	"Robotics and chemistry." (Eunice, naming the STEM areas where she feels strongest)
STEM Performance	Descriptions of doing STEM through visible participation, practice, projects, clubs, competitions, teaching, or applied problem solving.	"Also, I'm a peer teacher in science. I teach when the teacher is not there." (Eunice)

2.2 Data Analysis

This study used epistemic network analysis (ENA) to model patterns of co-occurrences among STEM identity-related codes [8]. Data were formatted as a binary spreadsheet for use with the ENA Web Tool, version 1.9.0 [9]. Constructs spanned columns and participant utterances were displayed in a row alongside metadata such as country, speaker, interview type, and row ID. Units of analysis included all coded participant lines associated with a single country, interview type, and speaker, and a one-line moving window was used to capture construct co-occurrence within individual utterances. Interview utterances were excluded from the dataset, and two raters independently coded the data before using social moderation to reach consensus on final codes [10].

3. Results

3.1 Network Models

Figure 1 displays ENA models of STEM identity-related discourse patterns for all 11 participants: USA participants in 1a, Kenyan participants in 1b, and the different network map in 1c. The STEM Self-Recognition and STEM Performance connection (Edge 1, Figure 1c) appears only in the USA map. In the post interview, Raliyah stated, *"I'm president of STEM club…"* as she described why leadership is important to her, positioning herself as both a STEM participant and leader. This connection is also clear when Ashanti responds with *"me being in stem and being active with other people, who want to learn new things and create stuff"* when asked what event impacted how she sees herself as a STEM student, linking STEM self-recognition to active STEM participation.

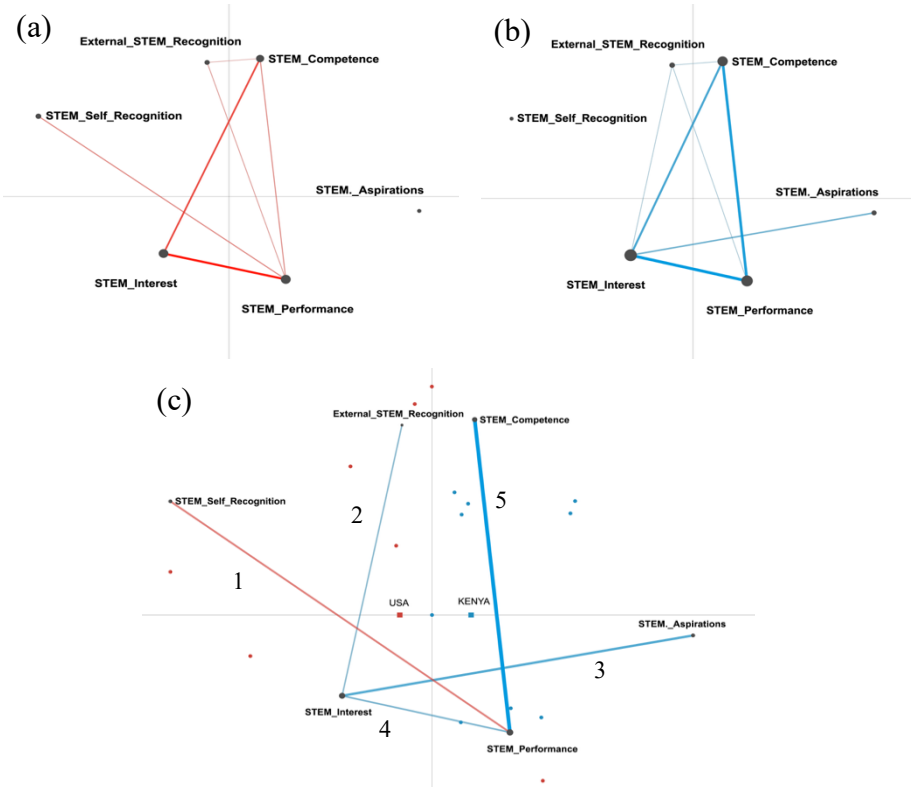


Fig. 1. ENA model showing (a) mean network graph and points for USA (red), (b) mean network graph and points for Kenya (blue), and (c) difference network.

Alternatively, the STEM Interest and External STEM Recognition connection (Edge 2, Figure 1c) as well as STEM Interest and STEM Aspirations connection (Edge 3, Figure 1c) appear only in the Kenya map. Participant discourse suggests that intrinsic STEM interest may lead students to seek out opportunities or groups where external recognition from can occur. One participant explains:

> *Yeah. I'm really into science into science and mathematics and. Okay. Also, I'm a peer teacher in science. I teach when the teacher is not there. And for math classes, I really try my best. So, math is fun. I really enjoy doing math. (Eunice)*

STEM Interest and STEM Aspirations are closely linked because students often describe their future STEM goals as growing out of their fascination with STEM topics. Interest appears to be the motivation that helps students imagine a future in STEM. Liana explains it like this:

> *Because I love sciences. And the reason why I study science is because I want to, like, be, a forensic psychologist, which actually involves science and criminology and everything. So, it interests me. It has always interested me since I was a child. And look at how plants grow, since my mum is like, a farmer. So I've been interested in sciences, since childhood. (Liana)*

While the STEM Interest and STEM Performance connection (Edge 4, Figure 1c) and the STEM Performance and STEM Competence connection (Edge 5, Figure 1c) are seen in both the USA and

Kenya maps, the difference map indicates that there may be a stronger connection between these constructs within the Kenya data. As expected, students' enjoyment and interest in STEM often lead them to actively participate in STEM activities, and at the same time, doing STEM can make STEM feel more enjoyable. Danielle explains:

> *I really enjoyed. I had a lot of fun. I, I was able to interact with my members, with my group members. I learned new tricks. I learned how to make a rocket. But apparently mine backfired. But I learned, and I really enjoyed, I had fun. Apart from the differences, I wouldn't be able to make it because of the technical issues. Apart from that, I think I had the best time with my members. Yeah. (Danielle)*

Similarly, as expected, STEM Competence and STEM Performance are strongly connected because students often develop confidence through active STEM participation. This connection is exemplified by Raven describing participating in math contests, winning trophies, and becoming number one in an algebra contest.

> *Maths....Maths... Maths, I have gone to some contests, and I've won some trophies. Yep. And they have recently impacted me. So they always push me to continue with what I'm doing. Yeah. I'm going to a robotics contest this year. And yeah, I had a good time there and it always motivates me to go on with what I like….We were learning about algebra and we were doing it with some other high schools. And in that, I became number one. (Raven)*

4. Discussion

Overall, these findings suggest that STEM identity development is both experiential and relational. Students develop strong STEM interest, confidence, aspirations, and self-recognition when they have opportunities to actively participate in STEM experiences and when others recognize them as capable. Interestingly, the comparison between the USA and Kenyan participants suggest that context may shape which aspects of STEM identity are most visible. For USA students, it appears as if doing STEM is what allows them to see themselves as a STEM person, being especially tied to belonging, leadership, and opportunity. For Kenyan students, STEM experiences appeared connected to enjoyment, skill-building, competitions, and future career direction. Across both groups, students' STEM pathways were strengthened when interest was supported by active participation and external STEM validation. These findings point to the importance of designing STEM programs that provide meaningful hands-on experiences, opportunities to demonstrate competence, leadership roles, and affirming feedback from peers, adults, and STEM professionals.

Acknowledgments

This material is based upon work supported by the US National Science Foundation under Award No. 2215613. Any opinions, findings and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation.

References

1. Branje, S., de Moor, E. L., Spitzer, J., & Becht, A. I. (2021). Dynamics of identity development in adolescence: A decade in review. Journal of Research on Adolescence, 31(4), 908–927. doi: 10.1111/jora.12678
2. Carlone, H. B., & Johnson, A. (2007). Understanding the science experiences of successful women of color: Science identity as an analytic lens. *Journal of Research in Science Teaching*, *44*(8), 1187–1218.

3. Hazari, Z., Chari, D., Potvin, G., & Brewe, E. (2020). The context dependence of physics identity: Examining the role of performance/competence, recognition, interest, and sense of belonging for lower and upper female physics undergraduates. *Journal of Research in Science Teaching*, *57*(10), 1–25.
4. Jiang, Z., & Wei, B. (2023). Understanding science identity development among college students: A systematic literature review. *Science & Education*. Advance online publication.
5. Albarello, F., Crocetti, E., & Rubini, M. (2018). I and us: A longitudinal study on the interplay of personal and social identity in adolescence. *Journal of Youth and Adolescence*, *47*(4), 689–702.
6. Jones, N.N., Wilson, A.D., Gaskin, L.N., Rayford, D. and Owusu, K.A. (2025). Using Quantitative Ethnography to explore the cultural nuances of identity of secondary students in Ghana, Kenya and the United States of America in an informal learning program. *Journal of International Cooperation in Education.* 27(2). pp.104-120, https://doi.org/10.1108/JICE-04-2024-0016
7. Shaffer, D. W., & Ruis, A. R. (2017). Epistemic network analysis: A worked example of theory-based learning analytics. In C. Lang, G. Siemens, A. F. Wise, & D. Gasevic (Eds.), Handbook of learning analytics (pp. 175–187). Society for Learning Analytics Research.
8. Shaffer, D. W. (2017). *Quantitative ethnography*. Madison, WI: Cathcart Press.
9. Marquart, C. L., Hinojosa, C., Swiecki, Z., Eagan, B., & Shaffer, D. W. (2021). Epistemic network analysis (Version 1.9.0) [Software]. Available from http://app.epistemicnetwork.org
10. Herrenkohl, L. R., & Cornelius, L. (2013). Investigating elementary students' scientific and historical argumentation. The Journal of the Learning Sciences, 22(3), 413–461. doi: 10.1080/10508406.2013.799475

Disciplinary Differences in Conceptions of Laboratory Instruction among Preservice Science Teachers: An Epistemic Network Analysis with Multigroup Statistical Comparisons

Yung-Hsiang Huang[1[0000-0003-4558-8517]] and Chia-Ching Lin[1[0000-0002-8951-2584]]

[1] Graduate Institute of Science Education and Environmental Education, National Kaohsiung Normal University, Taiwan
Codyemichaelyhh@gmail.com
cclin.edu@gmail.com

1 Introduction

Laboratory instruction integrates conceptual understanding with scientific inquiry [1], whereas digital technologies can expand opportunities for experimentation, representation, and reflection [2]. Our previous study [3] used drawing-based interviews and developed an 18-code framework comprising content presentation (CP), experimental process optimization (EPO), increasing interest and engagement (IEE), collaborative learning (CL), experimental skills (SKL), trial-and-error learning (TNE), knowledge application (KA), problem solving (PS), experimental principle understanding (EPU), experimental design (ED), reflection (REF), and self-directed exploration (SDE); .E and .T denote conceptions without and with digital technology, respectively. The study then used epistemic network analysis (ENA) to identify distinct conceptions of digital technology-integrated laboratory instruction among preservice science teachers: student-centered orientations (SC) emphasized knowledge construction through inquiry, whereas teacher-centered orientations (TC) emphasized the acquisition and application of established knowledge. SC conceptions emphasized SDE, PS, CL, and REF. In contrast, TC conceptions emphasized SKL, KA, and TNE (Fig. 1) [3]. The SC–TC contrast was represented primarily along MR1, which was derived through mean rotation to maximize the separation between the SC and TC groups. Based on the dominant code connections at each end of this dimension, MR1 was interpreted as reflecting a continuum from knowledge construction to knowledge acquisition. The orthogonal SVD2 dimension was labeled authentic versus generic teaching characteristics in the original study; however, because SC and TC did not differ significantly along this dimension, this interpretation should be treated cautiously [3].

Since scientific inquiry is partly discipline-specific, and teachers' conceptions and enactment of inquiry differ across science disciplines, conceptions of laboratory instruction may also vary by disciplinary background [4]. However, little is known about whether preservice teachers in Physics, Chemistry, and Biotechnology differ in how they conceptualize laboratory instruction incorporating digital technology. This exploratory secondary analysis reexamines the original coded data by disciplinary background

and projects the three groups into the original SC–TC ENA model to investigate such disciplinary differences.

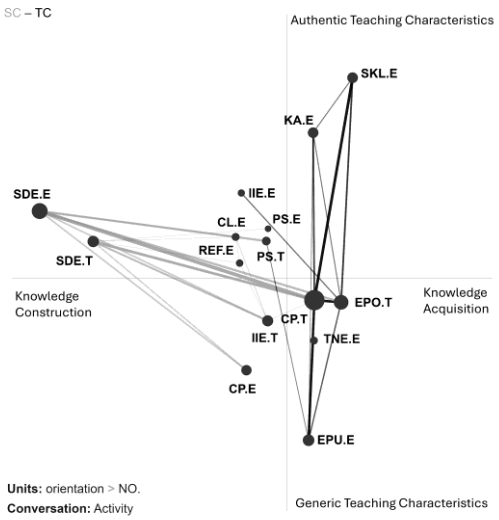


Fig. 1. Subtraction network of epistemic network models between SC (grey) and TC (black) orientations in the original study [3], and the model was retained as the fixed ENA reference space for the present analysis.

2 Methods

This study conducted a secondary analysis using the coded data from [3], reorganizing 54 participants by discipline: physics (n = 19), chemistry (n = 19), and biotechnology (n = 16). The original ENA model was constructed with individual participants as the unit of analysis, interview content as the conversation boundary, a window stanza of 1, and means rotation using TC and SC groups as rotation referents. The model explained a total of 15.93% of the variance, with co-registration Pearson and Spearman correlations all above .95, indicating good model fit. The present study adopted this model as a fixed reference space and used the rENA package [5] to project the three disciplinary groups into the same ENA space for comparison.

3 Results

Prior to group comparisons, normality (Shapiro–Wilk) and homogeneity of variances (Levene's test) were assessed for each dimension, and appropriate tests were selected accordingly. For MR1, a Kruskal–Wallis test revealed a significant difference among the three groups, χ^2 (2)=9.946, p = .007. Dunn's post hoc test with Holm correction indicated significant differences between Physics and Biotechnology (z = 2.352, p_{adj}

= .019, r = .320) and between Physics and Chemistry (z = 2.970, p_{adj} = .005, r = .404), both with medium effect sizes, whereas Biotechnology and Chemistry did not differ significantly (z = −0.488, p_{adj} = .313). For SVD2, Welch's ANOVA revealed no significant difference among the three groups, $F(2, 31.225) = 0.922$, p = .408 (see Fig. 2).

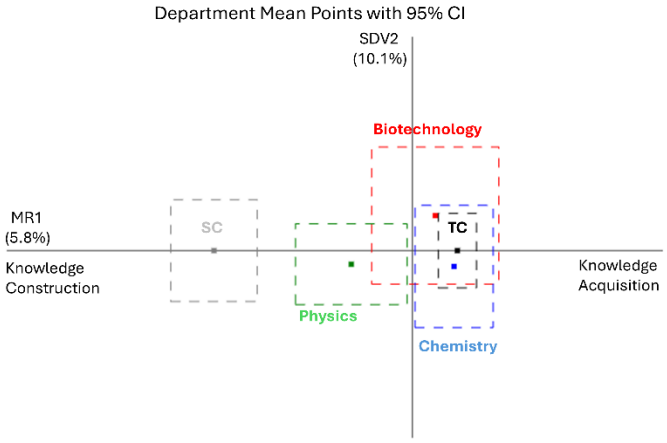


Fig. 2. ENA mean points of conceptions of laboratory instruction for Physics (green), Chemistry (blue), and Biotechnology (red) preservice teachers, projected into the fixed reference space [4]. SC and TC indicate the mean points of the original SC and TC groups, respectively.

Subtraction network analyses, a minimum edge weight threshold of 0.04 was applied, revealed that physics preservice teachers consistently emphasized knowledge construction-oriented conceptions across both comparisons, including SDE.E/T, CL.E, and PS.E/T. In contrast, both chemistry and biotechnology preservice teachers' conceptions of laboratory instruction were more oriented toward knowledge acquisition, such as SKL.E, EPO.T, and EPU.E.

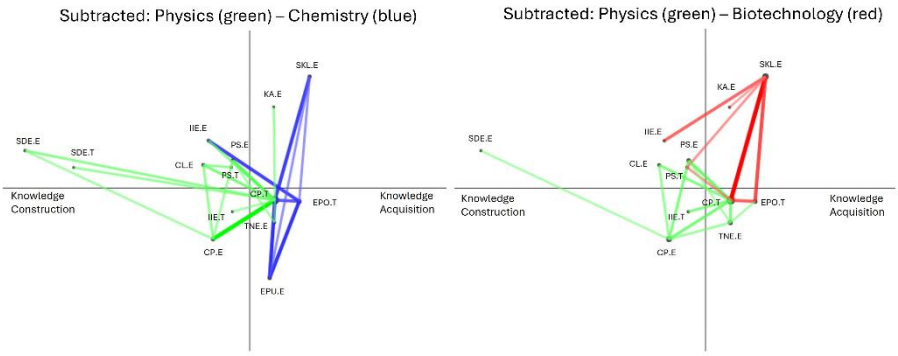


Fig. 3. Subtraction networks comparing conceptions of laboratory instruction across disciplines. Left: Physics (green) minus Chemistry (blue); Right: Physics (green) minus Biotechnology (red).

4 Conclusion and Discussion

This exploratory secondary analysis found that physics preservice teachers were positioned more strongly toward the knowledge-construction end of MR1 than their chemistry and biotechnology counterparts. This pattern is consistent with prior research showing that science teachers' conceptions and enactment of inquiry can vary across disciplines [6] and that scientific inquiry may be influenced by disciplinary context [4]. One possible explanation is that laboratory activities across disciplines differ in their degree of structure and the student independence they afford [7]. However, because the present study did not examine participants' prior coursework or laboratory experiences, this interpretation remains tentative. The findings nevertheless suggest that teacher education may benefit from discipline-responsive preparation for laboratory instruction. Future research should use larger, new samples and examine whether these patterns generalize across contexts.

Acknowledgments. This study was funded by National Science and Technology Council (NSTC), Taiwan (grant number NSTC114-2410-H017-014-MY3).

Disclosure of Interests. The authors have no competing interests to declare that are relevant to the content of this article.

References

1. Hofstein, A., Lunetta, V.N.: The laboratory in science education: Foundations for the twenty-first century. Science Education. 88, 28-54 (2004). https://doi.org/10.1002/sce.10106.
2. Wang, C.-Y., Wu, H.-K., Lee, S.W.-Y., Hwang, F.-K., Chang, H.-Y., Wu, Y.-T., Chiou, G.-L., Chen, S., Liang, J.-C., Lin, J.-W., Tsai, C.-C.: A review of research on technology-assisted school science laboratories. Journal of Educational Technology & Society. 17, 307-320 (2014).
3. Huang, Y.H., Wang, P.L., Lin, C.C.: Employing epistemic network analysis to explore preservice science teachers' conceptions of laboratory instruction incorporating digital technology. Contemporary Journal of Science Education. 33, 373-405 (2025). https://doi.org/10.6173/CJSE.202512_33(4).0003.
4. Chi, S., Wang, Z., Liu, X., Yang, Y.: Investigating disciplinary context effect on student scientific inquiry competence. International Journal of Science Education. 41, 2736-2764 (2019). https://doi.org/10.1080/09500693.2019.1697837
5. Marquart, C.L., Swiecki, Z., Collier, W., Eagan, B., Woodward, R., Shaffer, D.W.: rENA: Epistemic Network Analysis, https://cran.r-project.org/package=rENA, (2023).
6. Breslyn, W., McGinnis, J.R.: A comparison of exemplary biology, chemistry, earth science, and physics teachers' conceptions and enactment of inquiry. Science Education. 96, 48-77 (2012). https://doi.org/10.1002/sce.20469.
7. Buck, L.B., Bretz, S.L., Towns, M.H.: Characterizing the level of inquiry in the undergraduate laboratory. Journal of College Science Teaching. 38, 52-58 (2008).

Epistemic Network Analysis of Self-Concept Articulation in a Dialogue-Based Identity Exploration Program for Adolescents

Go Shoji[1], Naoya Yoshida[2], Kaoru Hatano[2], Keita Suzuki[3]

[1] Waseda University, Tokyo, Japan
[2] Axis Discovery School, Tokyo, Japan
[3] The University of Tokyo, Tokyo, Japan
go.shoji@akane.waseda.jp

Abstract. How do adolescents articulate a personal identity axis through structured dialogue? This study applies Epistemic Network Analysis (ENA) to examine how the co-occurrence structure of epistemic elements shifts across three temporal phases of a dialogue-based self-exploration program for Japanese middle and 8 high school students. Eight codes were applied to segmented transcripts: past experience (PAST), self-evaluation (SELF), emotional state (EMO), values (VAL), other-relatedness (OTH), future projection (FUT), concrete action (ACT), and metacognition (META). Results show that the most significant network reorganization occurred not in the final presentation phase, as program design might imply, but in the middle phase, where SELF, VAL, and META formed a newly consolidated cluster. In the final phase, FUT became connected to this cluster rather than emerging independently. These findings suggest that self-concept articulation in structured programs is better characterized as a reconfiguration of epistemic connections than as a staged linear progression from past to future.

Keywords: Epistemic Network Analysis, identity exploration, self-concept articulation, adolescent career development, dialogue-based learning.

1 Introduction

Evaluating self exploration and career programs is hard: change following the designed activity sequence is difficult to separate from genuine epistemic reorganization. A program that opens with retrospection, turns to values, and closes with future planning yields data in which past talk dominates early and prospective talk rises late; confirming that pattern documents compliance, not learning. This is acute for jibunjiku, a Japanese construct for the provisional yet personally meaningful articulation of one's values, tendencies, and life direction, refined through dialogue rather than discovered as a fixed label. The useful question is not which themes appeared when but how the relational structure among epistemic elements shifted.

Narrative identity theory frames self concept formation as relational and temporal (McAdams and McLean, 2013; Savickas, 2015). ENA models discourse as a network in which codes are nodes and codes occurring together within a moving window form weighted edges; Fan, Barany, and Foster (2023) used it to show how future selves and exploration concepts combine differently across stages. We extend this to a Japanese adolescent program not tied to any single domain.

- RQ1: How the network structure of self concept articulation shifts across the early, middle, and late phases?
- RQ2: How those shifts appear in participants' discourse ?

2 Method

2.1 Program and Phases

The program comprised nine coached sessions (Day 1 through Day 5, with intermediate half steps) for Japanese students in grades 7 through 12. The analytic sample consisted of eight students. Sessions used life timeline reflection, free association keywords, small experiments, image based projective dialogue, role play in future work, and a final oral presentation of one's axis; coaches asked Socratic questions and peers gave feedback. Sessions were grouped into three phases: early (T1: Days 1 to 1.5), middle (T2: Days 2.5 to 3.5), and late (T3: Days 4 to 5). The phases do not assume a past, values, future progression; each held varied activities and all three kinds of talk throughout, providing comparable temporal units rather than a thematic order.

2.2 Data and Codes

One author iteratively reviewed a subset of the transcripts and considered potential additions, mergers, and refinements to the code definitions. The final codebook was established through discussion among the authors. Eight binary codes captured the epistemic elements of self concept articulation (Table 1); multiple codes could apply to one segment.

Table 1. Codebook.

Code	Description
PAST	Reference to a specific past event, episode, or time period
SELF	Evaluation or characterization of own abilities, traits, or tendencies
EMO	Verbalization of an emotional or affective state
VAL	Expression of a value, evaluative criterion, or normative judgment
OTH	Reference to a specific other person, relationship, comparison, or social influence

FUT	Projection toward future states, plans, aspirations, or imagined scenarios
ACT	Description of a specific behavioral choice or concrete action
META	Metacognitive awareness, self-reflection on one's own thinking or change process

2.3 ENA Configuration

Analysis used webENA. The unit of analysis was participant by phase, and the conversation unit was participant by session day, so the moving window created no edges across sessions. The eight codes were nodes, with a moving window of four lines. Participant utterances were aggregated at the unit level; coach and peer utterances remained only for window calculation, so the networks represent participant structure. Mean and difference networks were compared and trajectories plotted. Given the small sample, inferential statistics are supplementary, and interpretation rests on the fit between structure and discourse.

3 Results

3.1 Phase-Level Network Structure (RQ1)

In T1 the mean network centered on SELF, with strongest links to PAST and EMO and a separate strong OTH and ACT link (Fig. 1, left); early self descriptions rested on specific episodes and emotional memory, with others as context, while VAL and META carried little weight. The sharpest change came in T2 (Fig. 1, center): links among SELF, VAL, and META strengthened markedly and PAST and OTH connected into VAL, as experiential material became evaluative and reflective connection. In T3 (Fig. 1, right) that cluster was preserved while FUT connected to its three nodes; FUT formed no strong links of its own to PAST, EMO, or ACT but entered as a terminal node mediated by the T2 cluster.

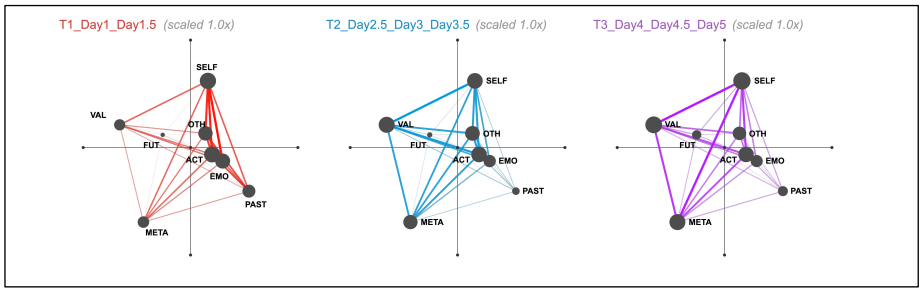


Fig. 1. Schematic of the mean ENA networks for the three phases. Node positions are held constant; line thickness indicates edge weight. The cluster linking SELF, VAL, and META forms in T2 and stabilizes in T3, where FUT enters as a terminal node mediated by that cluster rather than as an independent hub.

3.2 Discourse Evidence for Network Shifts (RQ2)

In T1, participants described past events through emotion and relationship without tying them to value principles: one narrated a relay race in which a self assessed failure of leadership led to team breakdown, voiced as confusion at the gap between her competence and others' responses, carrying PAST, EMO, SELF, and OTH but neither VAL nor META. In T2 the same kind of material appeared differently. A participant revisited her procrastination and reframed it not as a weakness but as a tendency that revealed the conditions under which she works well, adding an evaluative dimension (VAL) and explicit reflection (META) to self characterization (SELF); the reframing arose inside a keyword task, not a direct question about values. In T3, future statements sat inside a frame of self understanding and values: one account of joining the family business began with childhood feeling toward the workplace, articulated as belonging (VAL), which then grounded a projection of continued involvement (FUT), so that VAL mediated between PAST and FUT. T3 also showed frequent hedging and revision; one participant moved from calling herself a leader to describing leadership as a quality she exercises apart from any formal role, a move that added META while sustaining VAL.

4 Discussion

The central result is that the largest reorganization occurred in T2, the middle phase, not in T3, the final presentation. This unsettles the assumption that presenting one's jibunjiku is where self concept forms: the presentation drew on a SELF, VAL, and META framework already assembled in earlier sessions, consistent with Savickas but locating the shift before the closing moment. That FUT was mediated rather than autonomous suggests that what matters may be less a future image's vividness than how far it is anchored in a coherent evaluative and reflective frame. Future studies should examine whether the middle-phase network reorganization observed here recurs across larger samples, multiple cohorts, and diverse program settings.

Disclosure of Interests. The fourth author (K. Suzuki) received a donation from Axis Discovery School. The remaining authors have no competing interests to declare.

References

1. Fan, Y. K., Barany, A., & Foster, A. (2023). Possible future selves in STEM: An epistemic network analysis of identity exploration in minoritized students and alumni. *International Journal of STEM Education,* 10, Article 22. https://doi.org/10.1186/s40594-023-00412-z
2. McAdams, D. P., & McLean, K. C. (2013). Narrative Identity. *Current Directions in Psychological Science*, 22(3), 233-238.
3. Savickas, M. L. (2015). Career counseling paradigms: Guiding, developing, and designing. In P. J. Hartung, M. L. Savickas, & W. B. Walsh (Eds.), APA handbook of career intervention, Vol. 1. Foundations (pp. 129–143). American Psychological Association. https://doi.org/10.1037/14438-008

Evaluating the Impact of a Global Collaborative STEM Project: A Pre–Post Analysis of Teen Participants

Lu Liang[1[0009-0002-3848-3506]], Danielle P. Espino[1[0000-0002-6885-2125]], Seung B. Lee[1[0000-0003-0149-0681]], Julia S. Gibson[2[0009-0002-4841-1805]], Anmol Kaur[1[0009-0001-6680-3813]], Nastassia N. Jones[3[0009-0000-6042-6415]], Eric R. Hamilton[1[0000-0002-2337-1328]]

[1] Pepperdine University, Malibu, CA 90263, USA
[2] American University, Washington, DC 20016, USA
[3] Southern University, Baton Rouge, LA 70807
lu.liang@pepperdine.edu

Abstract. The study is based on an online STEM project designed to provide adolescents with access to global collaboration within an informal learning environment. The research analyzed learners' pre- and post-interview responses and connections among their social and STEM-related identities. Findings highlight the age-based development and meaningful transformation of their increased awareness across the project, contributing to the growing discussion on STEM learning design that cultivates adolescents' social connection and identity formation in the computer-supported collaborative learning process.

Keywords: Computer-supported Collaborative Learning, Social Environment, STEM Education, Epistemic Network Analysis.

1 Introduction

Computer-supported collaborative learning (CSCL) has emerged as a pathway for learners to access information beyond the traditional classroom and connect with people worldwide. As a platform for learners to form communities and engage in social exchanges, virtual collaboration increasingly forms a learner mindset and cultural awareness that foster their personal growth [1, 2]. Research also suggested that, although social presence significantly correlated with the learning experience and satisfaction of both younger and more experienced learners, age differences play an important role in shaping their online behavior patterns [3].

The study conducts a longitudinal comparative analysis of participants' discourse patterns in pre- and post-interview for an online collaborative STEM learning project to (1) compare the overall outcomes for the students in two school sites, and (2) highlight the differences observed in each school site before and after the project cycle. The study aims to add evidence to inform systemic research on collaborative STEM learning and effective design in a virtual environment that supports adolescent learners' interaction and achievement.

2 Methods

Learners participating in this project were middle- and high-school students, grouped into teams that met weekly online to collaboratively complete STEM-related projects. The pre-interviews were conducted from August to September 2024, and the post-interviews in May and June 2025. Table 1 presents the breakdown of the interview data (see Table 1). A codebook of 14 codes was used to conduct social moderation to reach the final codes (see Table 2) [4]. Grounded in the complexity of qualitative data, Epistemic Network Analysis (ENA) was used to visualize connections among the constructs [5]. The current study used interviews and school sites as units, a moving window of 1 line to examine the interconnected constructs within participants' discourse.

Table 1. Participant data.

School site	# of Participants	Grade	Male	Female
Bears Academy	16	9-12	4	12
Lions Academy	15	6-9	9	6

Table 2. Table captions should be placed above the tables.

Code	Definition
Social Awareness	Identifying the importance of interpersonal communication and relationships.
Math Orientation	Focus on math as a subject or concept.
Science Orientation	Focus on science as a subject or concept.
Global/Cultural Awareness	Recognizing differences could arise based on culture.
Collaborative Openness	Seeking assistance from others in finding a solution
Aspirational Orientation	Future-oriented, motivated goals.
Personality Attributes	How a student behaves as an individual/person.
Social Attributes	Identifying as part of a particular social grouping.
Perseverance	Pushing through challenges/difficulties; resilience.
Self-efficacy	Recognizing strength in something.
Challenge/Deficiency Awareness	Recognizing weakness in something.
Uncertainty	Not knowing about something or having ambivalence.
Encountering Challenges/ Differences	Identifying problems encountered when working with others or in external situations.
External Validation	External factors about why they are good at something.

3 Results

Figure 1 demonstrates the subtracted network model for learners' pre- and post-interviews. In the pre-interview (red), the most prominent connection is between Social Awareness and Personality Attributes, whereas in the post-interview (blue), strong relationships are observed between Social Awareness and Global/Cultural Awareness

and Uncertainty, and between Science Orientation and Aspirational Orientation and Self-efficacy. Throughout the project, learners developed more interconnected relationships among the constructs in the post-interviews, showcasing more comprehensive discourse patterns, interactive communication, and information exchanges at the individual and group levels.

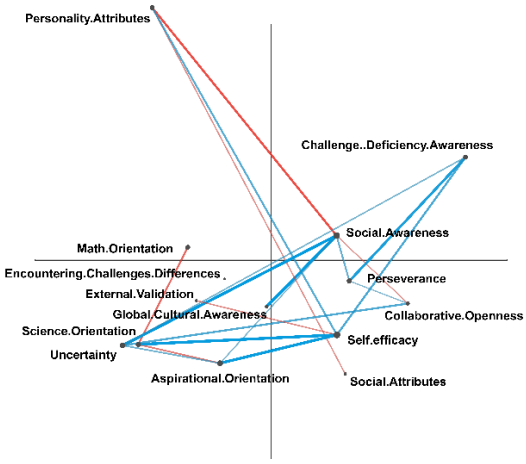


Fig. 1. Subtracted model for all learners in pre-interview (red) and post-interview (blue)

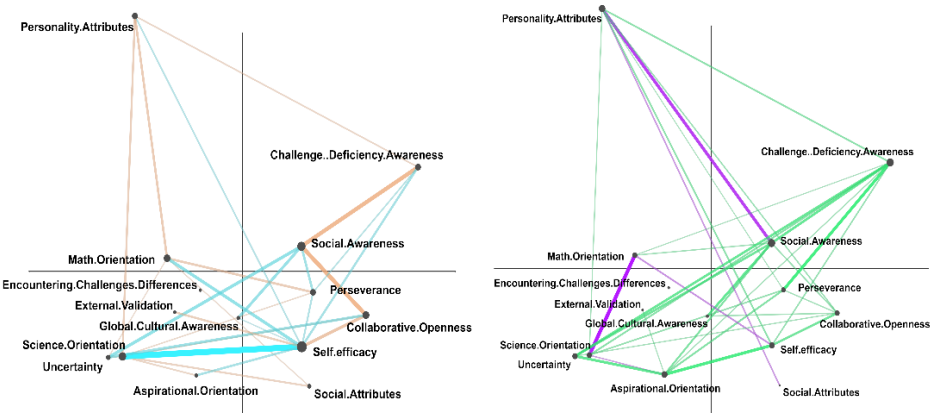


Fig. 2. Subtracted model for Bears Academy pre-interview (orange) and post-interview (cyan) on the left; and Lions Academy pre-interview (purple) and post-interview (green) on the right

Researchers further examined the subtracted ENA model for Bears Academy and Lions Academy (see Fig. 2). For Bears Academy, the post-interview indicated stronger connections between Self-Efficacy and both Math Orientation and Science Orientation, as one learner described science as a challenge but coupled it with an optimistic

mindset: *Science… it's kind of challenging for me. I enjoy the subject, but I feel like there's a bunch of, like, a lot of stuff that is hard to remember… [but] the fundamentals are there, and I feel like I can manage.*

For Lions Academy, the post-interview emphasized more on aspirational and self-efficacy. One student noted: *I'm very outgoing and I like talking to people and I have conversations with people. I enjoy doing things like coding, robotics, and I'm an extroverted person in general*. Additionally, Lions Academy participants demonstrated greater recognition of the importance of overcoming adversity through effort and persistence. As reflected by a seventh grader: *So it's very important to me to be a good collaborator because I really want to know better to do it. And I was struggling a bit more with it in the beginning of the year, but I feel like I've improved a lot and I'm going to continue to improve.*

4 Conclusion

This paper analyzes how learners' individual perspectives and social interactions within CSCL environments evolve over time. Findings align with prior CSCL research emphasizing the role of collaborative environments in fostering learners' self-reflection and social awareness and suggest that an informal online learning experience can strengthen self and social awareness, perseverance, and reflective thinking across age groups [6]. The comparative network models illustrate that learners' discourse patterns shift from more self-oriented to broader social, collaborative, and future-oriented thinking over the project cycle, and emphasize the importance of creating virtual learning environments that support cross-generational communication and collaboration at an early age.

Acknowledgments. The authors gratefully acknowledge funding support from the US National Science Foundation (Award #2215613) for the work reported in this paper. Views appearing in this paper do not reflect those of the funding agency.

References

1. Barnes, M., Lafferty, K., Li, B.: Assessing twenty-first century competencies: Can students lead and facilitate the co-construction process? Educ. Rev. **76**(4), 691–709 (2024)
2. Galdonez, D.P.: Impact of alternative learning activity program on student leadership development in a science, technology, engineering, and mathematics school. Sci. Educ. Int. **35**(3), 250–260 (2024)
3. Siriaraya, P., Ang, C.S.: Age differences in the perception of social presence in the use of 3D virtual world for social interaction. Interact. Comput. **24**(4), 280–291 (2012).
4. Herrenkohl, L.R., Cornelius, L.: Investigating elementary students' scientific and historical argumentation. J. Learn. Sci. **22**(3), 413–461 (2013).
5. Shaffer, D.W.: Quantitative Ethnography. Cathcart Press, Madison (2017)
6. Kreijns, K., Yau, J., Weidlich, J., Weinberger, A.: Towards a comprehensive framework of social presence for online, hybrid, and blended learning. Front. Educ. 8 (2024).

Exploring Discourse Changes in a Facilitator-Supported Novice Learner Community

Mao Morita[1], Hideki Kondo[2][0009-0001-8775-3712] and Hiroyuki Masukawa[1][0000-0001-8840-8905]

[1] Aoyama Gakuin University, 4-4-25 Shibuya, Shibuya-ku, Tokyo 150-8366, Japan
[2] Kanda University of International Studies, 1-4-1 Wakaba, Mihama-ku, Chiba 261-0014, Japan
c9125006@aoyama.jp

Abstract. Improving ideas requires comparing and connecting them, which may challenge novice learners without appropriate support. This study examines a learning community in which a facilitator supported learners in comparing and connecting ideas. Epistemic Network Analysis (ENA) and a complementary analysis of learner participation suggested greater involvement by learners in comparing and connecting ideas during the later meeting.

Keywords: Facilitator-Supported Learning Community, Novice Learners, Epistemic Network Analysis.

1 Introduction

Improving ideas requires comparing and connecting them. However, in communities of novice learners without appropriate support, discourse may remain limited to sharing ideas or providing simple responses. Prior ENA research has examined patterns of questioning and explanation in discourse related to improving ideas [1], whereas less attention has been paid to how learners participate in the process of comparing and connecting ideas within facilitator-supported discourse. Accordingly, this study examines a community in which a facilitator supported learners in comparing and connecting ideas. Using the twelve Knowledge Building principles [2] as analytic tags, we use ENA and a complementary analysis of learner participation to explore how learners' participation in this process differed between the early and later meetings.

2 Method

This study examined a learning community comprising two undergraduate students who were novices in AI and computer science and one faculty facilitator. Learners conducted interest-based inquiries, documented their progress on a shared wiki, and discussed their work in weekly online meetings. To compare discourse at relatively early and later time points while holding participant composition constant, we selected two meetings in which all three participants were present throughout: the third meeting,

held on December 16, 2025 (early), which was the earliest eligible meeting in the early period, and the tenth meeting, held on February 24, 2026 (later), which was the latest eligible meeting in the later period. The early and later meetings differed in duration and discussion topic. The early meeting lasted 53 minutes and 17 seconds and focused on the development of an AI tool for supporting critical essay reading and computer component purchases, whereas the later meeting lasted 1 hour, 58 minutes, and 44 seconds and focused on the continued development of the same AI tool and on reading academic papers. This study followed standard reporting guidelines and ethical procedures.

The recordings were transcribed, and utterances were segmented at the sentence level according to their meaning. A codebook (Table 1) was developed from the twelve Knowledge Building principles [2] and Teo's examples [3]. The first author applied the same coding criteria to both meetings, assigning each applicable utterance to the best-matching principle. Utterances that did not match any principle, such as greetings, backchannels, and administrative remarks, were left uncoded and denoted as N only in the examples. Of the 696 utterances in the early meeting and the 1,084 utterances in the later meeting, 213 and 309, respectively, were assigned to one of the 12 principles.

Table 1. Codebook

Code	Definition
Real Ideas, Authentic Problems (RI)	Utterances expressing ideas based on one's interests or authentic problems.
Improvable Ideas (II)	Utterances treating ideas as improvable objects.
Idea Diversity (ID)	Utterances presenting, comparing, or connecting different ideas.
Rise Above (RA)	Utterances synthesizing ideas at a higher level.
Epistemic Agency (EA)	Utterances directing inquiry processes or goals.
Community Knowledge, Collective Responsibility (CK)	Utterances promoting shared understanding or the advancement of community knowledge.
Democratizing Knowledge (DK)	Utterances encouraging contributions from all members.
Symmetric Knowledge Advancement (SK)	Utterances in which members mutually present ideas and contribute to one another's understanding.
Pervasive Knowledge Building (PK)	Utterances connecting the discussion to everyday contexts.
Constructive Use of Authoritative Sources (CU)	Utterances that use authoritative sources constructively.
Knowledge Building Discourse (KD)	Utterances oriented toward the advancement of community knowledge.
Embedded and Transformative Assessment (ET)	Utterances reflecting on and improving the community's understanding or discourse.

ENA [4] was used to explore co-occurrences among the 12 principles in the early and later meetings. The model used the moving stanza window method, with a window size of 4, and included utterances without a principle code. To examine

speaker-specific patterns over time, Units were Month–Speaker combinations, Conversations were topic-shift segments, and the Codes were the 12 principles listed in Table 1. Although the sample was small, little is known about how learners participate in discourse in which a facilitator supports learners in comparing and connecting ideas. We therefore used ENA exploratorily to characterize speaker- and time-specific connection patterns within this community. The ENA plots were interpreted by comparing the strength of specific connections across the two meetings. A subtraction network was used to identify connections that were relatively stronger in either meeting. We considered the RI (Real Ideas, Authentic Problems: utterances expressing ideas based on one's interests or authentic problems)–ID (Idea Diversity: utterances presenting, comparing, or connecting different ideas) connection potentially important because it links learners' expression of their own ideas with their comparison and connection of ideas. Speaker-specific utterance proportions and relevant discourse were then examined to interpret learner participation in the focal connection.

3 Results

Figure 1 presents the ENA results. In the later meeting, RI–ID was the strongest connection (connection weight = 0.35). The subtraction network showed that II (Improvable Ideas: Utterances treating ideas as improvable objects)–CU (Constructive Use of Authoritative Sources: Utterances that use authoritative sources constructively) had the largest later-weighted difference (0.26), followed closely by RI–ID (0.25). Although II–CU showed a substantial difference, the later meeting's focus on academic papers may have provided more opportunities for CU. We therefore focused on RI–ID, which was also prominent and was less directly related to the topic specific to the later meeting, to examine learner participation in comparing and connecting ideas.

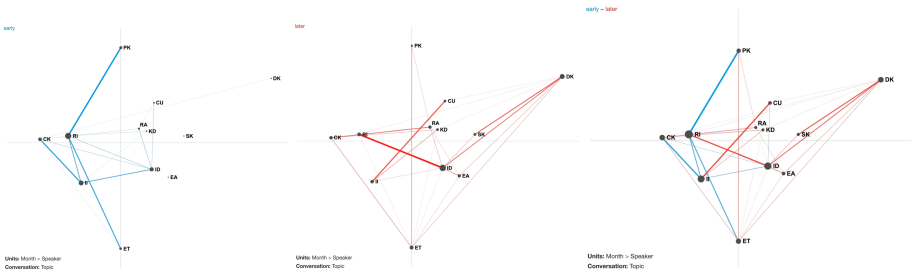


Fig. 1. ENA networks and subtraction plot for early (blue) and later (red) meetings.

Learners continued to account for most RI utterances across both meetings (early: 86 of 89, 96.6%; later: 36 of 47, 76.6%). In contrast, their share of ID utterances increased from 1 of 21 (4.8%) in the early meeting to 11 of 59 (18.6%) in the later meeting, with corresponding increases for both Learner A (1 to 7 utterances) and Learner B (0 to 4 utterances). Examination of the discourse contributing to RI–ID further showed a difference in participant involvement within the connection. All three

RI–ID instances identified in the early meeting involved the facilitator, whereas one of five instances in the later meeting occurred between the learners without facilitator involvement. The following excerpt from the later meeting illustrates this pattern in a discussion of the development of the AI tool for supporting critical essay reading. Learner A introduced a different idea and connected it to an authentic problem, after which Learner B related the discussion to their own experience.

Learner A: Then, isn't there something like curriculum guidelines? For example, guidelines on how to teach critical essays? (ID)
Learner B: That's true. (N)
Learner A: I wonder if people who have difficulty reading critical essays may have gotten stuck in elementary or junior high school classes and moved on without overcoming it. (RI)
[...]
Learner B: Thinking about it, I had assumed that there were no curriculum guidelines because I myself was never taught how to read critical essays in elementary, junior high, or high school. (RI)

4 Discussion

In this study, the facilitator supported learners in a small learning community by presenting ideas and helping learners compare and connect ideas. ENA and the analysis of learner participation showed that learners accounted for most RI utterances in both meetings, while both learners contributed more to ID utterances in the later meeting. In addition, one RI–ID instance in the later meeting occurred between learners without facilitator involvement, a pattern not observed in the early meeting. These findings suggest greater learner involvement in comparing and connecting ideas in the later meeting. However, given the exploratory two-time-point design, these differences should be interpreted cautiously. Changes in topic and meeting duration may reflect changes in the community over time, but they may also have provided different opportunities for comparing and connecting ideas. Two coding limitations remain: coding was conducted only by the first author, and each utterance was assigned a single label. Future work should examine inter-rater reliability and use multi-label coding.

References

1. Ong, A., Teo, C.-L., Lee, A.-V.-Y., and Yuan, G.: Investigating shared agency in student Knowledge Building discourse using epistemic network analysis. Research and Practice in Technology Enhanced Learning, 21, 006 (2026). https://doi.org/10.58459/rptel.2026.21006
2. Scardamalia, M.: Collective Cognitive Responsibility for the Advancement of Knowledge. In: Smith, B. (ed.) Liberal Education in a Knowledge Society, pp. 67–98. Open Court, Chicago, IL (2002)
3. Teo, C.L.: A Teacher's Playkit for Knowledge Building: Creative Ways to Make Students' Ideas Come Alive. Routledge, London (2023). https://doi.org/10.4324/9781003149163
4. Shaffer, D.W.: Quantitative Ethnography. Cathcart Press, Madison, WI (2017)

How Does Ambiguity Tolerance Relate to Creativity-Fostering Teaching Practices among Art Teachers? An Exploration through Epistemic Network Analysis

Jimpei Hitsuwari[1,2], Yoshitaka Ito[3], and Bunichi Otaki[4]

[1] Helmut Schmidt University, Hamburg, Germany
[2] Japan Society for the Promotion of Science, Tokyo, Japan
[3] Higashi-Mozu Junior High School, Sakai, Japan
[4] Graduate School of Science and Technology, Shizuoka University, Hamamatsu, Japan-
hitsuwari@hsu-hh.de

Abstract. This study used Epistemic Network Analysis (ENA) to explore how teachers' ambiguity tolerance (AT) relates to creativity-fostering teaching. Art teachers were split at the tails of the Absolutism subscale of the Japanese Multidimensional Attitude toward Ambiguity Scale into lower-absolutism (higher-AT; n = 8) and higher-absolutism (lower-AT; n = 9) groups, and their focus-group talk was coded on the nine Creativity Fostering Teacher Index dimensions. In the higher-AB group Integration was a hub tied to Independence and Frustration, with small dispersion; the lower-AB group showed weaker links across many code pairs. For the same response to student difficulty (Frustration), pathways differed: the higher-AB group supported students through collaboration (Integration), the lower-AB group through motivation for foundational skills (Motivation). As the groups came from the distribution's tails, we report these as exploratory contrasts suggesting creativity-fostering styles that differ with AT.

Keywords: Ambiguity Tolerance, Creativity-Fostering Teaching, Epistemic Network Analysis, Art Education, Group Interview.

1 Introduction

Research on creativity-fostering teaching asks how teachers' own cognitive characteristics shape the practices through which they nurture students' creativity. Cropley [1] organized the conditions for fostering creativity, operationalized in Soh's [5] Creativity Fostering Teacher Index (CFTIndex), a measure of teacher behaviors that support student creativity. Its nine dimensions, used here as codes, are Independence (letting students work on their own), Integration (organizing cooperation and cross-reference among students), Motivation (encouraging persistence and interest), Judgement (withholding premature evaluation), Flexibility (adapting plans to students), Evaluation (involving students in judging work), Questioning (encouraging students' questions), Opportunities (open-ended choices of material and theme), and Frustration (responding to students' setbacks). Because creative work requires holding a task open, ambiguity

tolerance (AT)—in Frenkel-Brunswik's [2] sense, the disposition to accept ambiguous situations rather than force premature closure—is in turn seen as a foundation for creativity [8], and teachers high in AT show positive associations with creativity and playfulness [7]. Yet how teachers' own AT surfaces in practice has rarely been examined.

Epistemic Network Analysis (ENA; [4]) visualizes the co-occurrence structure of codes as a network, bridging quantitative indices with qualitative content. We ask how the network structure of teachers' talk about creativity-fostering practice differs between lower- and higher-absolutism teachers. For Quantitative Ethnography the question is also methodological: showing that a measured person-level trait surfaces as the connective structure of discourse would link person-level measurement with discourse-level practice.

2 Method

2.1 Participants and Grouping

Recruitment had three steps: an online survey yielded 54 valid responses; 30 consented to a follow-up interview; 17 of these joined the focus groups reported here. Respondents came from the second author's professional network—convenience sampling, so they are plausibly more engaged with art education than the population. The survey used the Japanese MAAS [3] (Discomfort with Ambiguity, Absolutism [AB], Need for Complexity and Novelty). We grouped participants post hoc on AB (black-and-white thinking), chosen for its largest variance, bipolar distribution, and closest fit to Frenkel-Brunswik's [2] conception of AT [6]; lower AB indicates higher AT. Taking the two tails and setting the middle aside gave a higher-AB group (n = 9; AB $M = 3.75$, $SD = 0.33$) and a lower-AB group (n = 8; AB $M = 2.00$, $SD = 0.29$), with completely separated scores (Mann-Whitney U, $W = 72$, $p < .001$). The 17 were 12 women, 3 men, and 2 unreported; age bands ran from the 20s to 60s or above (40s and 50s most frequent) and teaching experience from 1–5 to over 31 years (most often 11–15). Specializations were painting (6), design (4), art education (2), and one each in four other areas (1 unreported). The whole-sample mean AB (2.99) fell below the seven-point midpoint, so these labels are relative to this sample rather than absolute.

2.2 Data, Coding, and Analysis

Four focus-group interviews were run on Zoom—two higher-AB (L1, L2) and two lower-AB (H1, H2) groups of four to five—facilitated by the second author. Utterances were coded on the nine CFTIndex dimensions [5], multiple codes allowed; interviewer turns were excluded. Two coders independently coded L1; after jointly refining the definitions, re-coding reached 73.9% exact match, mean Jaccard 0.798, 95.8% cell agreement, and pooled Cohen's $\kappa = .64$ (per-code κ median .66, lowest where a code appeared fewer than three times). Shaffer's rho reached the conventional .05 threshold for no code (lowest rho = .10), as the double-coded set is small (69 utterances) and most codes rare; reliability is therefore provisional. The other groups were single-coded with

the agreed codebook (H2: Coder 1; L2 and H1: Coder 2), so coder is partly confounded with condition. In ENA (web tool) the stanza window was 4, spanning a typical question-and-answer exchange, and was not varied in a sensitivity check. Group (L/H) was the comparison unit, with first-dimension differences tested by Mann-Whitney *U*.

3 Results

Code frequencies (participant utterances only; higher-AB = 115 lines, lower-AB = 111), as higher-AB/lower-AB: Independence 15/11, Integration 16/3, Motivation 8/9, Judgement 2/1, Flexibility 10/11, Evaluation 4/1, Questioning 2/2, Opportunities 5/13, Frustration 11/9. The sharpest contrasts were Integration and, in the other direction, Opportunities.

In the comparison plot (Fig. 1), the higher-AB group showed Integration as a hub, thickly tied to Independence and Frustration, with members clustered tightly; its Integration–Evaluation edge rests on too few utterances (4 vs. 1) to be stable and is not interpreted. The lower-AB group showed no comparably strong link, but moderate-to-weak connections across pairs such as Motivation–Frustration and Flexibility–Opportunities, with members widely dispersed. On the first dimension the groups separated completely (higher-AB $Mdn = -0.26$; lower-AB $Mdn = 0.34$; $U = 0$, $p < .001$, $r = 1.00$), but since they came from the tails of the AB distribution, near-maximal separation was built into the design; we treat this as an exploratory contrast, not an inferential test. Because the conversation unit was the focus group, connections form across speakers within the four-line window: no single utterance carried both Frustration and Integration, so that edge arises entirely from cross-speaker sequences (Integration and Independence co-occurred within six single utterances).

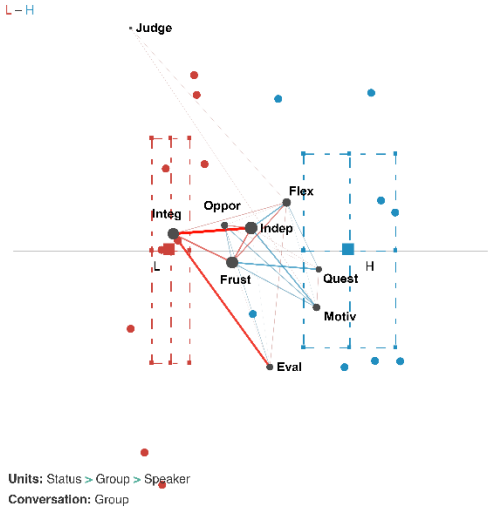


Fig. 1. ENA comparison plot of the higher-absolutism group (L, red) and the lower-absolutism group (H, blue). L and H denote lower and higher AT, respectively.

4 Discussion

The higher-AB group's Integration-centered pattern corresponds to teachers interweaving students' autonomous work with organized collaboration. One teacher described this interweaving:

> With a theme where motivated kids can keep going, I cover only the key points and leave the rest somewhat up to them, or I introduce another student who is good at it. This year I spent a lot of time supporting the students who cannot manage on their own. (Speaker E)

The teacher leaves capable students to proceed (Independence) while brokering peers for mutual reference (Integration), so as to focus on those needing help. Integration in this group meant proceduralized cooperation rather than free dialogue: staged interim presentations, and group exchange of works and worksheets. The teacher builds occasions on which students must look at one another's work, then uses them to redistribute their own attention toward students who cannot yet manage alone. This suggests that higher-AB teachers structure lessons around goals and evaluation.

The lower-AB group instead emphasized Opportunities and a dispersed, multi-pathway pattern with greater within-group diversity. One teacher described offering open-ended material choices:

> If we are doing woodwork, it could be a pen tray, a plate, a shoehorn—anything. By specifying "a paper knife," we crush the children's possibilities. (Speaker N)

Valuing an open process over a guaranteed outcome—even when a work fails partway—comes more readily to teachers who tolerate ambiguity, and the dispersion echoes Zenasni et al.'s [8] finding that AT predicts originality more than fluency. Frustration was handled differently in the two groups (Frustration–Integration in the higher-AB group vs. Frustration–Motivation in the lower-AB group). For the lower-AB group a setback was read as accumulated self-doubt and became the occasion for motivation: one teacher traced a student's difficulty to a long-standing sense of being poor at drawing, then offered concrete technique so the student could feel "I did it," praising each small step afterwards. Higher-AB teachers instead scaffolded setbacks through collaboration, placing a struggling student alongside peers. AT thus appears reflected in how far teaching relies on structure.

The study has some limitations. Each condition comprises two focus groups, so group membership is confounded with focus-group membership: participants shape one another's talk, and the observed difference could reflect a conversational dynamic in L1/L2 versus H1/H2 rather than an individual trait. Participants are therefore not independent units in the test above. Questioning and Judgement were too infrequent (≤ 2) to interpret, and reliability rests on one group (L1). School level (elementary vs. lower secondary) was not recorded, although it shapes what art teaching involves. Finally, the sample's mean AB fell below the scale midpoint, so the group labels are relative.

Even so, ambiguity tolerance was reflected in how art teachers talk about creativity-fostering practice—most clearly Integration as a hub and the divergent Frustration pathways—which future work should link to classroom observation and student outcomes.

References

1. Cropley, A. J.: Fostering creativity in the classroom: general principles. In: Runco, M. A. (ed.) The Creativity Research Handbook, vol. 1, pp. 83–114. Hampton Press, Cresskill (1997)
2. Frenkel-Brunswik, E.: Intolerance of ambiguity as an emotional and perceptual personality variable. J. Pers. **18**(1), 108–143 (1949)
3. Hitsuwari, J., Nomura, M.: Developing and validating a Japanese version of the Multidimensional Attitude toward Ambiguity Scale (MAAS). Psychology **12**(4), 477–497 (2021)
4. Shaffer, D. W.: Quantitative Ethnography. Cathcart Press, Madison, WI (2017)
5. Soh, K.: Fostering student creativity through teacher behaviors. Think. Skills Creat. **23**, 58–66 (2017)
6. Steger, D., Schütz, A.: Disambiguating tolerance of ambiguity: introducing the MAAS-9 as a short, reliable, and multidimensional measure of attitudes toward ambiguity. J. Pers. Assess. (2026). https://doi.org/10.1080/00223891.2026.2645707
7. Tegano, D. W.: Relationship of tolerance of ambiguity and playfulness to creativity. Psychol. Rep. **66**(3), 1047–1056 (1990)
8. Zenasni, F., Besançon, M., Lubart, T.: Creativity and tolerance of ambiguity: an empirical study. J. Creat. Behav. **42**(1), 61–73 (2008)

Inclusion as Contested Discourse: Perspectives from State, District, and School-based Practitioners

Lindi R. Shepard[1[0009-0003-5629-575X]], Rebecca A. Cruz[1[0000-0002-9163-1312]], Alexandra E. Shelton[1[0000-0003-3659-347X]], Rachel S. McClam[1[000-0002-4745-6094]], Sarah A. Caroleo[2[0000-0002-0041-9616]], Irene Ramirez[1[0009-0004-7522-6079]], and Katy Mullins[1[0009-0001-5849-7967]]

[1] Johns Hopkins University, Baltimore MD 21218, USA
[2] Annenberg Institute, Brown University, Providence, RI 02912, USA
lshepar6@jhu.edu

Abstract. In the United States, the Individuals with Disabilities Education Improvement Act (IDEA, 2004) mandates that students receiving special education services have access to the general education setting "to the maximum extent appropriate." However, without objective guidelines on what is appropriate, IDEA can be used to rationalize enacting *or* withholding inclusion. Enactments of inclusion are thus reliant on how state, district, and school-based actors interpret inclusion and translate these understandings into local policy, placement decisions, and everyday practice. Drawing on Critical Inclusion, this study examines how educators and other partners perceive and approach inclusion and how these understandings differ across levels of the education system. We used Epistemic Network Analysis (ENA) to analyze discourses about inclusion from state working group transcripts and interviews with district and school-based educators with attention to how conceptualizations of inclusion intersect with normative conceptions of ability. Across groups, participants linked inclusion to a normative center, but while school practitioners showed greater variation in the forms of inclusion they described, district leaders more commonly discussed inclusion as contingent on a student meeting normative academic and behavioral standards. Meanwhile, state actors emphasized inclusion as a means of fostering belonging for students across cultural differences. Our results show that inclusion is not a singular concept but a contested discourse shaped by assumptions around who and what is normal and accepted in educational settings. We argue that meaningful inclusion requires moving beyond placement-based logics toward a rightful presence framework that centers rigorous learning, belonging, and access for all students.

Keywords: Special Education, Inclusion, Epistemic Network Analysis, Quantitative Ethnography.

1 Background

In the United States, debates about disability inclusion have persisted for decades despite the principle established in *Brown v. Board of Education* [1] that separate educational opportunities are inherently unequal. These debates are shaped by IDEA

[2], the primary U.S. special education law, which requires students to be educated in the LRE "to the maximum extent appropriate" while also maintaining a continuum of placement options ranging from general education classrooms to fully segregated settings. Because the law leaves concepts such as “appropriate” and “satisfactory” educational progress undefined, IDEA can be used to rationalize enacting or withholding inclusion via state interpretations of the policy. Consequently, separate educational placements remain both common and legally permissible for many students with disabilities, and rates of inclusion vary widely between states [3]. Although racial segregation is illegal, racialized patterns of exclusion continue to operate through special education systems [4-6]. Currently, the field possesses few tools to challenge district and school norms that perpetuate the exclusion of historically marginalized students from meaningful learning opportunities [7].

2 Theoretical Framing

Ableism is embedded within the normative center of U.S. schools, where non-disabled ways of being are privileged [8] and students with disabilities, particularly those at the intersection of multiple forms of oppression (e.g., racism) [9], experience limited opportunities for meaningful inclusion [10] despite policies intended to promote access. Our work is theoretically guided by Critical Inclusion [11-15], a framework that challenges the discourse of inclusion situated in such normative assumptions. Critical Inclusion problematizes the idea of “normal” learners [12] and emphasizes all students’ fundamental right to classroom membership and rigorous learning [16-17].

3 Purpose

We examined state interpretation and district enactment of IDEA, and how actors across different levels of the educational system define and enact inclusive education. Our research seeks to answer the following questions:

1) What are prominent discourses related to inclusion?
2) How does this discourse differ between state, district, and school practitioners?

4 Data Sources

Data for this study are drawn from a research-practice partnership with a large school district in the Mid-Atlantic region of the U.S. Our partner district is situated in a state that recently committed to improving outcomes for students receiving special education. The state’s superintendent organized a working group that convened 16 times, discussing topics such as improving LRE placement rates. These meetings, which included contributions from 50 individuals, were recorded, transcribed, and made public. Data included transcripts from these convenings in addition to 12 interviews with district (n = 3) and school-based (n = 9) educators.

5 Methods

We used ENA [18] to analyze connections within discourse around inclusion. ENA models the connections among codes by quantifying their co-occurrence within conversations, producing a weighted network visualization for each unit of analysis (e.g., state, district, and school). We inductively and deductively coded transcripts by categorizing types of inclusion. In particular, we conceptualized eight types of inclusion in our data. These include: *segregated*, feeling welcomed in an exclusionary placement; *disadvantaged*, inclusion of students from low-income backgrounds; *cultural*, including students with diverse racial, ethnic, or linguistic backgrounds; *social*, including for the purpose of socializing with peers; *physical*, inclusion through proximity; *contingent,* requiring certain conditions to be met; *differential*, inclusion with adjustments to instruction that reduce rigor; and *critical*, inclusion in which all students access learning opportunities aligned with the general education curriculum in ways that are responsive, relevant, and rigorous [13]. For a full list of codes with definitions and examples, see Table 1.

Table 1. Coding Categories, Definitions and Examples

Category	Definition	Codes	Examples
Types of Inclusion	Participant describes how they or others conceptualize and/or implement "inclusion"	segregated disadvantaged cultural social physical contingent differential critical / meaningful	"There have been students who...if they, according to their maybe academic achievement and testing...they were coming from those programs which were totally independent classrooms to the inclusive class." Subcode: *contingent*
Normative Center	Discussion of normalized or norm-referenced ability in schools	normative center	"The data was the data right? Data speaks. But I think sometimes perception or...observed behavior sometimes influenced my colleagues willingness to open up their doors."

6 Results

Across groups, participants frequently connected inclusion to norm-referenced assumptions about ability, behavior, and participation. ENA results showing connections between types of inclusion and the normative center are displayed in Figure 1. State leaders most often framed inclusion through cultural or systems-level

rationales, while district leaders described inclusion as contingent upon students meeting normative expectations (e.g., verbal communication or behavioral compliance). School practitioners emphasized social inclusion, though their responses showed wide variation, describing many formulations of inclusion enacted within their settings. Further, discussions often positioned non-disabled ways of learning and behaving as the implicit standard against which inclusion decisions were made. Taken together, these findings suggest that inclusion was commonly framed not as a right, but as conditional participation within existing normative structures.

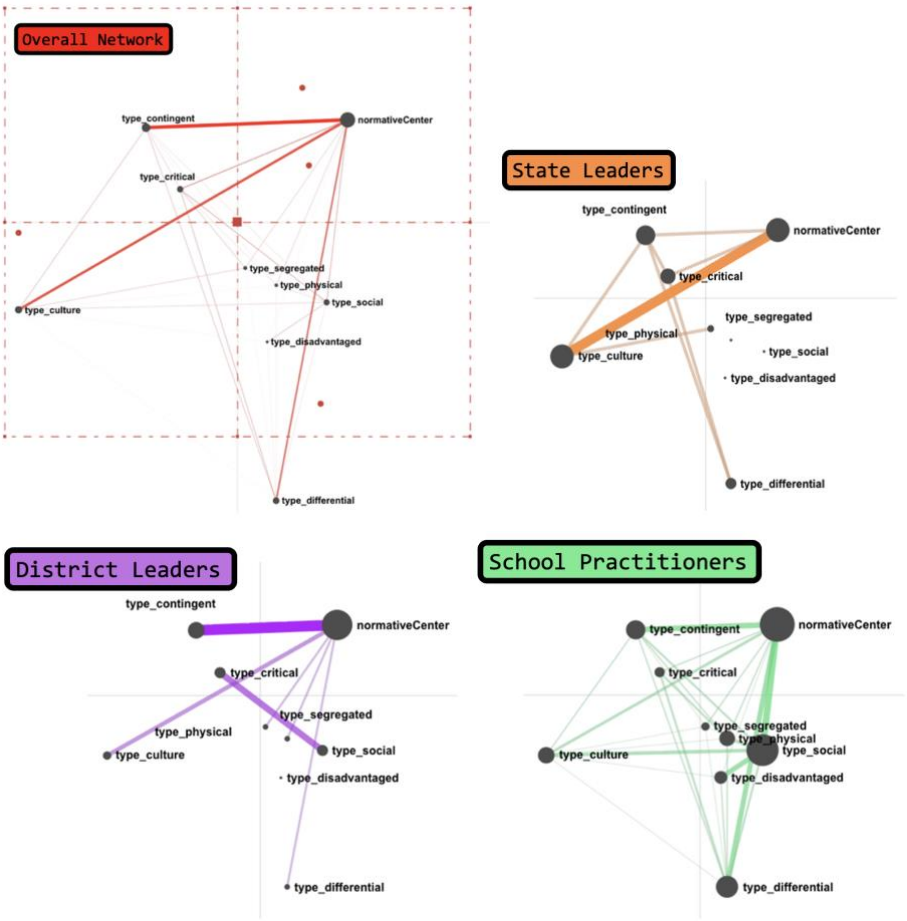


Fig. 1. Logics of inclusion for overall network and within state, district, and school levels of discourse.

7 Discussion and Conclusion

Our findings suggest that inclusion is not a singular, stable educational practice but a contested concept that takes on different meanings across system levels. State actors tended to frame inclusion through cultural belonging, district actors more often treated inclusion as contingent on students meeting normative expectations, and school practitioners varied widely in their understandings of inclusion and approaches to enactment. These competing interpretations were rooted in a normative center that positioned non-disabled ways of being as the standard, often leading educators to view segregation as an appropriate response when students deviated from those expectations. Thus, inclusion was frequently discussed in terms of conditional access rather than an unconditional right. By modeling these discourses through ENA, we show how inclusion is discursively produced and differentiated across institutional contexts. Although IDEA offers flexibility in how students receive special education services, that flexibility also permits decisions shaped by subjective perceptions of student ability rather than a shared vision of meaningful inclusion. Reimagining inclusion may therefore require adopting a framework of rightful presence, in which schools are responsible for creating responsive, rigorous, and relevant learning opportunities for all students. Until policy and practice challenge the normative assumptions that define who belongs, efforts to expand inclusion are unlikely to produce meaningful change.

Acknowledgments. This work was supported by the Maryland State Department of Education.

Disclosure of Interests. All authors declare that they have no competing interests.

References

1. Brown v. Board of Education, 347 U.S. 483 (1954).
2. Individuals with Disabilities Education Improvement Act of 2004 (IDEA), 20 U.S.C. § 1400 (2004).
3. McLeskey, J., Landers, E., Hoppey, D., Williamson, P.: Learning disabilities and the LRE mandate: An examination of national and state trends. Learn. Disabil. Res. Pract. 26(2), 60–66 (2011). https://doi.org/10.1111/j.1540-5826.2011.00326.x
4. Bell, N.S., Collier, Z., Vélez, V.N., Ford, D.Y.: CritSEM: Advancing QuantCrit to examine racialized resegregation in special education. J. Res. Educ. Eff. 1–33 (2024). https://doi.org/10.1080/19345747.2024.2408538
5. Cooc, N.: Disparities in general education inclusion for students of color with disabilities: Understanding when and why. J. Sch. Psychol. 90, 43–59 (2022). https://doi.org/10.1016/j.jsp.2021.10.002
6. United States Department of Education: Racial and ethnic disparities in special education. Office of Special Education and Rehabilitative Services (2016). Available at: https://www.ed.gov/sites/ed/files/programs/osepidea/618-data/LEA-racial-ethnic-disparities-tables/disproportionality-analysis-by-state-analysis-category.pdf
7. Miller, A.L., Mansouri, M.C., Ruhter, L.C., Kurth, J.A., Morningstar, M.E., Toews, S.G., Wilt, C.L., Andreoli, M.: Rightful presence? Teacher candidates navigating inclusive

education for youth with complex support needs. Except. Child. 91(3), 260–280 (2025). https://doi.org/10.1177/00144029251317029
8. Hehir, T.: Eliminating ableism in education. Harv. Educ. Rev. 72(1), 1–33 (2002). https://doi.org/10.17763/haer.72.1.03866528702g2105
9. Annamma, S.A., Ferri, B.A., Connor, D.J.: Disability critical race theory: Exploring the intersectional lineage, emergence, and potential futures of DisCrit in education. Rev. Res. Educ. 42(1), 46–71 (2018). https://doi.org/10.3102/0091732X18759041
10. Connor, D.J., Danforth, S., Gallagher, D.: An open letter to the field: Contemplating special education's collaborative role in developing inclusive education. Except. Child. (2024). https://doi.org/10.1177/00144029241303051
11. Artiles, A.J., Harris-Murri, N., Rostenberg, D.: Inclusion as social justice: Critical notes on discourses, assumptions, and the road ahead. In: Inclusive Schooling Practices, pp. 260–268. Routledge (2021).
12. Baglieri, S., Bejoian, L.M., Broderick, A.A., Connor, D.J., Valle, J.: [Re]claiming "inclusive education" toward cohesion in educational reform: Disability studies unravels the myth of the normal child. Teach. Coll. Rec. 113(10), 2122–2154 (2011). https://doi.org/10.1177/016146811111301001
13. Cruz, R.A., Firestone, A.R., Love, M.: Beyond a seat at the table: Imagining educational equity through critical inclusion. Educ. Rev. 76(1), 69–95 (2024). https://doi.org/10.1080/00131911.2023.2173726
14. Kulkarni, S.S., Miller, A.L., Nusbaum, E.A., Pearson, H., Brown, L.X.: Toward disability-centered, culturally sustaining pedagogies in teacher education. Crit. Stud. Educ. 65(2), 107–127 (2024). https://doi.org/10.1080/17508487.2023.2234952
15. Waitoller, F.R., Artiles, A.J.: A decade of professional development research for inclusive education: A critical review and notes for a research program. Rev. Educ. Res. 83(3), 319–356 (2013). https://doi.org/10.3102/0034654313483905
16. Calabrese Barton, A., Tan, E.: Beyond equity as inclusion: A framework of "rightful presence" for guiding justice-oriented studies in teaching and learning. Educ. Res. 49(6), 433–440 (2020). https://doi.org/10.3102/0013189X20927363
17. Kozleski, E.B.: Disrupting what passes as inclusive education: Predicating educational equity on schools designed for all. Educ. Forum 84(4), 340–355 (2020).
18. Shaffer, D.W.: Quantitative ethnography. Cathcart Press (2017).

Productive Failure in Generative-AI-Mediated App Development: A Quantitative Ethnography of Student Reflections and Test Score Change

Shuntaro Iseri [1][0009-0000-9651-1179], Hideki Kondo[1][0009-0001-8775-3712], and Masaaki Ishii[1][0000-0002-4074-3226]

[1] Kanda University of International Studies, 261-0014 Chiba, Japan
iseri-s@kanda.kuis.ac.jp

Abstract. The Generative Failure Method (GFM) adapts productive-failure principles to generative-AI learning. This exploratory study characterized GFM-related processes in post-course reflections and compared patterns by test-score change. Twenty-four students provided 144 responses; 15 had paired tests (score change ≥50 points, n=9; <50, n=6). Eight theory-informed codes were analyzed in response-level co-occurrence networks. Code reading and verification were common in both groups. Agency/ownership was more prevalent in the ≥50 group, whereas conceptual uptake was more prevalent in the <50 group; difference edges linked comparison and verification with agency in the ≥50 group. An exact global permutation test was nonsignificant (pseudo-F=0.97, R^2=.069, p=.462). Associations useful for generating hypotheses were observed, but they should not be read as showing that the coded practices produced or predicted gains.

Keywords: Generative AI, Productive failure, Programming education, Quantitative ethnography, Co-occurrence networks.

1 Background and Purpose

Productive Failure (PF) is more than error or unguided trial and error. In PF, generation and exploration precede consolidation and knowledge assembly [1, 2]. Mechanisms include activating prior knowledge, identifying critical features and knowledge gaps, explaining them, and organizing them into canonical concepts; comparing canonical solutions can reveal gaps [2].

GFM adapts these principles to generative-AI programming [3, 4] without reproducing a strict failure-before-instruction sequence. Because a large language model supplies code, explanations, and feedback during generation, exploration and consolidation may overlap. Prior work distinguishes passively using AI-generated code from understanding it and actively modifying or extending it [5]. We therefore examined how GFM-related processes varied across learners, rather than GFM effects or predictors of gain:

RQ1. How were GFM-related processes represented in students' post-course reflections?

RQ2. What differences in code prevalence and co-occurrence structure were observed between students whose test scores improved and those whose did not?

2 Method

Students in a Japanese non-STEM undergraduate course developed p5.js applications with generative AI. Six post-course prompts addressed AI interaction, output quality, code comparison and revision, project difficulty, programming understanding, and new questions. Twenty-four students supplied six responses each (144 fixed windows); for duplicates, we retained the latest. Fifteen students with paired programming-understanding tests formed the analytic sample.

The week 11 and 15 tests each had six questions and 28 scorable blanks. Blanks received 100 points if correct, 50 if partially acceptable, 5 for an incorrect attempt, and 0 if empty (theoretical maximum: 2,800). A 50-point gain—one partial-credit step—was the exploratory threshold (≥50, n=9; <50, n=6). Membership was unchanged at thresholds of 25–200 points. The forms were closely matched but not psychometrically equated.

We used primarily deductive, theory-informed coding with data-informed refinement (Table 1). PF/GFM categories, including agency/ownership and conceptual uptake, were added or revised while reading. Dictionary/regex rules generated binary prelabels. One researcher, blinded to score-change group, audited all windows and assigned final codes. Of 1,296 code-by-window decisions (144 × nine), 1,234 matched (95.2%); this single-auditor pre-code retention rate is not interrater reliability. The ninth category, affective shift, was excluded from the eight-code network.

Table 1. Eight-code PF/GFM-informed scheme used in the main analysis.

Code	Definition	Phase
AI_output_insufficient	AI output was incomplete or inappropriate as-is	Impasse
Prompt_specification	Specifying goals, constraints, or context	Repair
Iterative_dialogue	Repeatedly asking, refining, or following up	Repair
Debugging_repair	Fixing errors or unintended behavior	Repair
Code_reading_comparison	Reading, comparing, or interpreting code	Repair
Verification	Checking output, behavior, or correctness	Verification
Agency_ownership	Making own/team judgments about AI output	Agency
Conceptual_uptake	Understanding, noticing, or forming questions	Uptake

Students were independent units; responses were fixed windows. Each unordered code pair within a window counted as a co-occurrence; student counts were divided by six, yielding response-opportunity proportions. An exact global permutation test compared 28-edge vectors across all 5,005 nine-versus-six assignments. Exploratory Fisher's exact prevalence and exact edgewise permutation tests used within-family Benjamini–Hochberg adjustment. Post hoc, excerpts were purposively selected for the

largest signed edge contrasts and a contrasting repair case; they contextualize, not validate, the networks. Japanese translations were lightly edited.

3 Results

Code reading/comparison and verification appeared for 8/9 (88.9%) and 6/6 (100%) students in the ≥50 and <50 groups, respectively. Agency/ownership appeared for 8/9 (88.9%) versus 2/6 (33.3%), and conceptual uptake for 5/9 (55.6%) versus 6/6 (100%). None of the eight prevalence contrasts was significant after BH adjustment; we report them descriptively.

The largest positive differences were code reading/comparison–agency/ownership (.130 vs .000), debugging/repair–code reading/comparison (.241 vs .139), and verification–agency/ownership (.093 vs .000). The largest negative differences were code reading/comparison–conceptual uptake (.093 vs .222) and verification–conceptual uptake (.037 vs .139). The global difference was nonsignificant, pseudo-F(1,13)=0.97, R^2=.069, exact p=.462; no edge contrast survived FDR adjustment.

One ≥50-group student described interpreting and revising AI-generated code: "I could roughly tell what each part did, so I rewrote the code based on that understanding." A <50-group student described completion-oriented repair: "When the proposed code did not work, I changed the question slightly and asked again. It eventually worked." Another <50-group student called an error "a valuable learning opportunity," showing rich PF-related reflection in both groups.

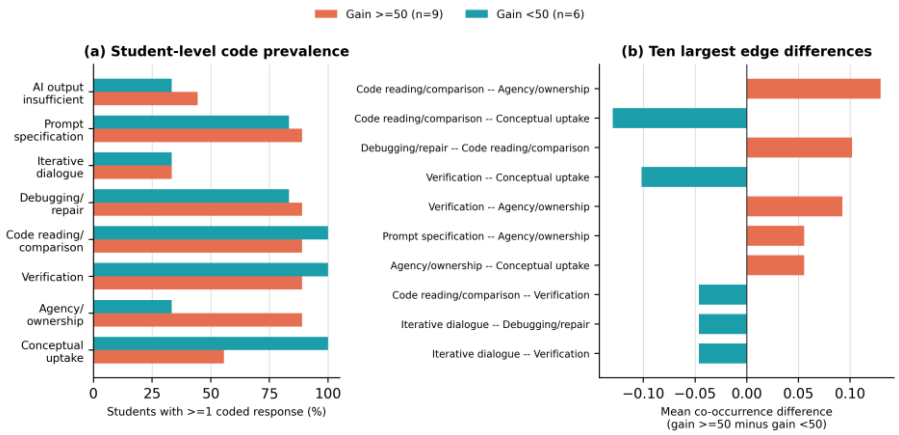


Fig. 1. Code prevalence and co-occurrence differences by score-change group. (a) Percentage of students with each final code (≥50: n=9; <50: n=6). (b) Ten largest differences in response-window-normalized mean edge weights (≥50 − <50); positive values favor the ≥50 group.

4 Discussion and Limitations

Reflections represented impasse, repair, verification, agency, and uptake (RQ1). Both groups described reading and checking code, but connections differed (RQ2): repair and verification connected more with agency/ownership in the ≥50 group, whereas conceptual-uptake links were stronger in the <50 group. This aligns with work distinguishing working AI-generated code from understanding how to actively modify it [5], but the nonsignificant global test and post hoc grouping preclude causal or predictive claims.

The findings also qualify GFM's relation to PF. In classical PF, consolidation follows initial generation; in LLM-mediated work, code, explanation, exploration, and repair may occur within one episode. The codes thus trace selected PF-related meaning-making processes, not every PF mechanism. Linking network contrasts to excerpts supports the QE interpretation [6].

The following limitations should be noted. The outcome-group comparison involved 15 students in one course. The 50-point threshold was an exploratory rubric-based rule, not a validated educational benchmark or reliable-change criterion. The test forms lacked item-level reliability and equating evidence. Coding was finalized by one researcher; the 95.2% pre-code retention rate is not interrater reliability. Fixed retrospective windows do not establish temporal sequence and may reflect prompt structure or response length rather than interaction. These constraints limit power and transferability. Future work should combine interaction logs and artifacts, independent coding, and larger samples to examine how repair, verification, and agency support understanding and transfer.

Acknowledgments. This research was supported by Kanda University of International Studies Research Grant(B) and JSPS-KAKENHI (Grant Number 24K06334 and 26K06358).

Disclosure of Interests. The authors have no competing interests to declare.

References

1. Kapur, M. (2008). Productive failure. Cognition and Instruction, 26(3), 379–424.
2. Kapur, M., & Bielaczyc, K. (2012). Designing for productive failure. Journal of the Learning Sciences, 21(1), 45–83.
3. Kondo, H., Tohyama, S., Iseri, S., & Ishii, M. (2024). Exploring the use of generative AI to realize productive failure in programming education (in Japanese). Information Education Symposium 2024, 176–183.
4. Kondo, H., Tohyama, S., Iseri, S., & Ishii, M. (2025). Generative failure: Learning design in programming with generative AI. CSCL 2025 Proceedings, 625–627.
5. Kazemitabaar, M., Hou, X., Henley, A.Z., Ericson, B.J., Weintrop, D., Grossman, T.: How novices use LLM-based code generators to solve CS1 coding tasks in a self-paced learning environment. Koli Calling '23, Article 3, 12 pp. (2023).
6. Shaffer, D. W., Collier, W., & Ruis, A. R. (2016). A tutorial on epistemic network analysis: Analyzing the structure of connections in cognitive, social, and interaction data. Journal of Learning Analytics, 3(3), 9–45.

Temporal Dynamics of Interactional Meaning-making of Topics in Interprofessional Clinical Education

Jun Lu[1[0009-0001-5698-2889]], Crina Damsa[2[0000-0001-7382-4164]], Rachelle Esterhazy[3[0000-0003-0494-1417]], Jun Oshima[1[0000-0002-9885-448X]] and Ritsuko Oshima[1[0000-0002-2499-6573]]

[1] Graduate School of Science and Technology, Shizuoka University, Shizuoka, Japan
ro.takashi.18@shizuoka.ac.jp
[2] Department of Education, University of Oslo, Oslo, Norway
[3] Centre for the Studies of Professions, Oslo Metropolitan University, Oslo, Norway

Abstract. In the study, we propose an integrated methodology and investigate the temporal dynamics of topics in a group dialogue in interprofessional clinical education practice. Identification of discussed topics is essential to understand students' meaning-making of knowledge from other professions. Previous studies have shown the effectiveness of topic modeling. Integrating temporal network analysis provides additional capability to capture temporal dynamics of topics. We analyzed a 60-minute discussion from an interprofessional simulation in a clinical education context. Students from different disciplines assessed a patient and proposed a clinical assessment plan based on disciplinary course documents. The course documents describe important topics students must discuss. We employed sapBERT to identify which topics were discussed in one group consisting of four students. Integration of temporal network analysis allowed us to capture the dynamic trajectory of topics. The results suggest that our proposed methodology identified the temporal dynamics of students' discussion, shedding light on the process of interprofessional, interactional meaning-making.

Keywords: Temporal Network Analysis, Topic Modeling, Interprofessional Education

1 Introduction

In this study, we address a methodological challenge in natural language processing (NLP) and QE: How can we capture the dynamics of topics in collaborative discourse? Along with qualitative analysis, computational methodologies have been discussed for decades to reduce the burden of manual coding and introduce quantitative interpretation. A major approach is topic modeling. In addition to conventional methods like Latent Dirichlet Allocation (LDA) [1], recent studies have proposed a BERT-based vector-embedding approach, enabling more flexible topic identification [2].

However, an important limitation of topic modeling is the reduced capability of detecting the structure of topics [3]. Topic modeling techniques detect *what* topics are discussed, but they cannot fully identify *how* topics are discussed. It becomes problematic when topics stem from different disciplines and are intertwined. When students

mention different topics in the same conversation, topic modeling extracts them as a single mixed topic, making it difficult to interpret *how* these topics are connected. One cannot interpret if the students are simply juxtaposing two different topics, or they are creating a meaningful connection between them and developing shared understanding.

Network analysis is a method to capture the relational structure by modeling the connection of topics [3, 4]. This is particularly effective in interdisciplinary discourse, because it provides an opportunity to reveal the trajectory of *meaning-making*. As per the example above. If it is a simple juxtaposition, these concepts will not be connected to other concepts. In contrast, when students are making meaningful connections between different concepts, other related concepts would also be connected, and the network will become dense. With network analysis, one can reach an interpretation that helps students build shared understanding. In this study, we integrate topic modeling and network analysis to capture the structure of topics.

Recent studies have proposed temporal network analysis [4]. Temporal network analysis expands usual network analysis by introducing lifetime into network structure. It increases sensitivity to temporal dynamics such as interactions of meaning-making. In this study, we apply temporal network analysis to a dialog from an IPE course.

As the need for interprofessional collaboration is recognized, IPE has attracted more attention over the past decade [1]. IPE prepares for interprofessional collaboration between students from nursing, medical, audiological, psychological, and other professions. The core process of IPE practice is interprofessional meaning-making [5]. Students discuss various topics, such as medical conditions and medication, mental health, social insurance status, and daily life. Through interprofessional discussion, students will develop their interprofessional understanding by interconnecting these topics [5]. It is essential to capture the dynamics of topics to understand and support IPE [1, 5].

To the best of our knowledge, no studies have combined topic modeling and temporal network analysis to investigate topics in interprofessional discourse. This integration could capture the dynamics of topics that are difficult to identify using single methodology [3]. In this study, we propose an integrated framework to investigate the temporal dynamics of topics in IPE practice. The research questions are as follows:

RQ1. What topics are discussed in the session?
RQ2. What are the temporal trajectories of topics interconnected by each student?
RQ3. How do students interconnect the topics?

2 Methods

In this study, we analyzed one session selected from a larger IPE program in Norway. This program is designed for interprofessional collaboration, and students from multiple professions participate in a clinical simulation. We analyzed one 60-minute discussion session of a group consisting of students from nursing, medical, pharmacy, and psychology education. They are anonymized as “Nurse”, “Doctor”, “Pharmacist”, and “Psychologist”. Students were asked to discuss and prepare a clinical assessment plan for a patient visit. Disciplinary course documents were provided as references, including a course guide and patient journal. The discussion was recorded and transcribed.

Our integrated methodology is as follows. First, the transcription was segmented by conversation turns as the analysis unit. To address RQ1, we employed sapBERT [6] to identify the topics of each conversation turn in students' discussion. Due to the semi-structured nature of the simulation, we conducted embedding-based topic assignment. Based on the disciplinary course documents, we qualitatively extracted the topics and created a topic corpus. Each conversation turn was then embedded and assigned one or several topics based on the cosine similarity. To address RQ2, we conducted temporal network analysis of topics in the students' discourse. Since topics can be modeled as a cluster of words [4], we analyzed the co-occurrence network of words. To capture the temporal dynamics of topics, we appended the identified topics to the end of each word, e.g., speech difficulty_CLI_A_audio. Total degree centrality (TDC) was employed as the metric of topic interconnection [4]. When topics interconnected within a conversation turn, edges are added between words, increasing TDC. Lastly, to address RQ3, a significant segment of topic interconnection was extracted and analyzed qualitatively.

3 Results and Discussion

Regarding RQ1, we identified 3 categories and 14 topics based on the disciplinary course documents. Category "clinical assessment" [CLI_A] included 10 topics: personal, social, situational, resource, cognitive, psychological, previous illness, audio, nutritional, and medication. Category "clinical examination" [CLI_E] included 3 topics: oral, blood test, and other exams. Category "ethical concerns" [ETH] included: ethical concerns. Topic modelling revealed that the most frequent topics were "medication" (99 times) and "previous illness" (95 times) in the category CLI_A, and the least one were "situational" in CLI_A (1 time) and "oral" in CLI_E (2 times). These results suggest that students discussed various topics beyond their own profession.

Regarding RQ2, the temporal network analysis revealed the group dynamics of topic interconnection. Fig. 1 shows the temporal trajectories. Nurse and Doctor have several spikes, indicating that they primarily interconnected the topics. The Psychologist also has spikes, and the Pharmacist has small spikes in several sequences. Most topics were interconnected in the first half, and the rest of the session was about elaboration.

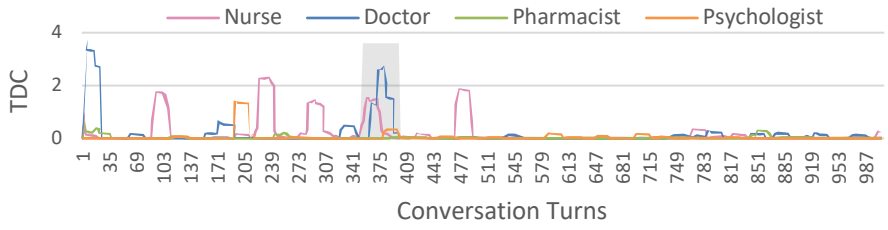
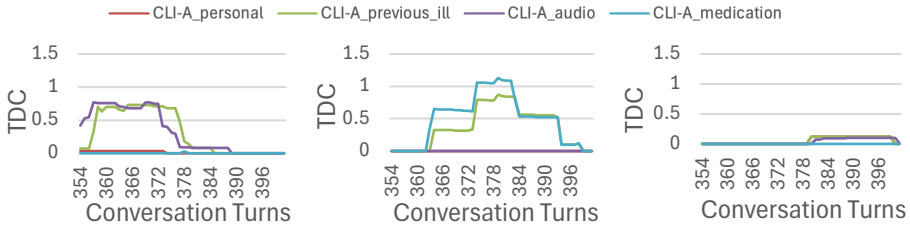


Fig. 1. Overall temporal trajectories of the topic interconnection by students. Conversation turns highlighted in grey were analyzed in the following section.

Regarding RQ3, the temporal trajectories of topics and in-depth qualitative analysis showed the process of interprofessional meaning-making. Fig. 2 suggests the temporal process of topic interconnection in conversation turns 354 to 401 by Nurse, Doctor, and

Psychologist. In this segment, the Nurse starts making meaning by connecting topics of previous illness and auditory status with the context of personal information, by stating "So those are really his problem, the speech difficulty and the impaired walking function." (turn 355). The Doctor took over and connected the previous illness and medication, discussing speech difficulty and Transient Ischemic Attack (TIA). In turn 380, the Psychologist states "I think there's a lot of interesting stuff here. What's the fear regarding those falls?", and "The speech difficulty in relation to family member." This illustrates interconnections of topics, and how students try to co-construct a shared understanding of previous illness, medical condition, and personal situation.



(a) Nurse (b) Doctor (c) Psychologist

Fig. 2. Temporal trajectories of topics from conversation turns 354 to 401 by students. Colored lines indicate each topic. Other topics not mentioned in this segment are omitted from the plots.

4 Conclusion

In this study, we integrated topic modeling and temporal network analysis and investigated the dynamic structure of topics in an IPE practice. The results suggested that (1) 14 topics were discussed, (2) collective trajectories of topic interconnection were identified, and (3) the temporal process of interprofessional meaning-making was revealed. Because this study has limitations of small sample size and insufficient qualitative interpretation, further analyses are necessary. Validation on sapBERT topic-assignment step against human coding is also required. Nevertheless, we contribute to the NLP and QE field by proposing an integrated methodology to capture the temporal dynamics of topics and shed light on the process of meaning-making in IPE.

Acknowledgments. This study was funded by JST SPRING (grant number JPMJSP2167) and supported by the TeamLearn project, grant number 324826.

Disclosure of Interests. The authors have no competing interests to declare that are relevant to the content of this article.

References

1. Black, E.W., Zurek-Ost, M., Strekalova, Y., Blue, A.: Natural language processing as a program evaluation tool in interprofessional education. Journal of Interprofessional Education & Practice **36**, 100712 (2024)

2. Abdulnazar, A., Kreuzthaler, M., Roller, R., Schulz, S.: SapBERT-based medical concept normalization using SNOMED CT. Studies in Health Technology and Informatics **302**, 825–826 (2023)
3. Barnett, G.A., Calabrese, C., Ruiz, J.B.: A comparison of three methods to determine the subject matter in textual data. Frontiers in Research Metrics and Analytics **8**, 1104691 (2023)
4. Oshima, J., Oshima, R., Kawakubo, A.J.T.: Multi-layered temporal network analysis: Idea improvement processes in knowledge-building practice leading to high learning performance. Journal of Computer Assisted Learning **41**(3), e70063 (2025)
5. Jakobsen, H.F., Filstad, C.: Crossing knowledge boundaries in interprofessional collaborations knowledge sharing in the child's best interest. Nordic Journal of Studies in Policing **7**(2), 121–136 (2020)
6. Liu, F., Shareghi, E., Meng, Z., Basaldella, M., Collier, N.: Self-alignment pre-training for biomedical entity representations. In: Proceedings of NAACL-HLT 2021, pp. 4228–4238. Association for Computational Linguistics (2021)

The Changing Structure of Student Reflection in AI-Supported Inquiry Learning

Ran Nakao[1][0000−0002−5772−972X], Go Sorano[2], Yizhou Fan[3][0009−0004−5690−1227], and Kaixiang Kang[3,4][0009−0001−9124−8073]

[1] Ehime University, Matsuyama, Ehime, Japan nakao.ran.fq@ehime-u.ac.jp
[2] Ehime Prefectural Iyo High School, Ehime, Japan
[3] Hiroshima University, Hiroshima, Japan
[4] Prefectural University of Hiroshima, Hiroshima, Japan

Abstract. Reflection is a critical component of learning, yet most existing studies evaluate it using a single outcome measure. This study examines the longitudinal structure of student reflection in an AI-supported inquiry-based learning environment, conceptualizing reflection as a network of interconnected epistemic practices rather than a single score. Data were collected from fifteen weekly inquiry-learning classes at a Japanese high school, in which students submitted reflections and received automated feedback from a generative AI system. Reflection and feedback were coded along six dimensions each using a theory-informed, LLM-assisted procedure. Preliminary results indicate that students' reflections became increasingly organized around reflective depth and metacognitive awareness over time, suggesting that reflection developed as a process of structural reorganization rather than simple improvement on isolated dimensions. Future analyses will employ epistemic network analysis (ENA) and student-level panel models to examine how AI-generated feedback characteristics are associated with subsequent changes in reflective structure.

Keywords: Quantitative Ethnography, Generative AI, Reflection, Learning Analytics

1 Introduction

Reflection has long been recognized as a critical component of learning, enabling learners to make sense of their experiences, monitor their understanding, and connect learning activities to future actions [4,2]. Generative AI systems have increasingly been introduced into educational settings as tools for feedback and reflective support, with previous studies suggesting that large language models (LLMs) can scaffold learning and promote reflective thinking [3,1].

However, most existing studies evaluate reflection using a single outcome measure, and relatively little is known about how different dimensions of reflection are structurally related, how these relationships evolve over time, or whether AI-generated feedback is associated with such structural changes. From a quantitative ethnography perspective, reflection can be understood as a network of

epistemic practices [5], and understanding how its dimensions become integrated over time provides insight into students' meaning-making processes.

Therefore, this study examines (1) how students' reflective structures are organized, (2) how these structures evolve across a semester-long program, and (3) how different forms of AI-generated feedback are associated with subsequent changes in reflective development.

2 Data

The data were collected from an inquiry-based learning program conducted at a Japanese high school between November 2025 and March 2026. Students participated in fifteen weekly sessions, submitting written reflections after each one and receiving automated feedback from a generative AI system. The dataset comprises 2,983 reflection–feedback pairs from 41 students collected throughout the semester.

The coding scheme was developed by synthesizing theories of reflective learning, metacognition, inquiry-based learning, and educational feedback [4,2,8,6,7,3,1]. Reflection texts were evaluated along six dimensions: concreteness, reflective depth, inquiry orientation, metacognition, forward-looking orientation, and connection to learning content. AI-generated feedback was evaluated along six corresponding dimensions: specificity, accuracy, feedforward orientation, scaffolding, inquiry promotion, and affective support. Each dimension was scored on a three-point scale (0 = absent, 1 = partially present, 2 = clearly present) using a common rubric derived from the theoretical framework described above.

To increase robustness, all texts were independently evaluated by three LLMs (Claude Haiku 4.5, Gemini 3.5 Flash, and DeepSeek V4 Flash), with final scores computed by averaging across models. Inter-model agreement was acceptable (mean pairwise $r = .64$ across the six reflection dimensions). To validate the LLM-assisted coding, all reflections were independently rated by a human coder blind to the model scores; human and LLM ratings were positively correlated (pooled $r = .68$), supporting the validity of the automated scores. These dimensions were used to construct longitudinal reflection networks and examine how feedback characteristics were associated with subsequent changes in reflective development.

3 Methods

This study employs a two-stage analytical strategy: first examining how reflective structure evolves over time, and second investigating whether AI-generated feedback is associated with subsequent changes in reflection.

We calculate wave-average scores for each reflection dimension to trace overall trends, and estimate wave-specific correlation matrices among the six dimensions to characterize the structural evolution of reflection across the fifteen-week program. Note that these correlation-based networks constitute a preliminary struc-

tural analysis and are distinct from formal epistemic network analysis (ENA), which we plan to apply in future work.

To examine feedback effects, we estimate lagged regression models using student-level panel data, where the outcome is the wave-to-wave change in each reflection dimension ($\Delta r_{it} = r_{it} - r_{i,t-1}$):

$$\Delta r_{it} = \beta_1 f_{\text{specific},i,t-1} + \beta_2 f_{\text{accurate},i,t-1} + \beta_3 f_{\text{forward},i,t-1} + \beta_4 f_{\text{scaffold},i,t-1} + \beta_5 f_{\text{inquiry},i,t-1} + \beta_6 f_{\text{affect},i,t-1} + \alpha_i + \gamma_t + \varepsilon_{it} \quad (1)$$

where α_i and γ_t denote student and wave fixed effects, respectively. Estimated coefficients should be interpreted as longitudinal associations rather than causal effects.

4 Preliminary Results

Figure 1 presents the average scores of the six reflection dimensions across waves (left) and correlation-based networks in Wave 1 and Wave 15 (right). Overall, reflective depth increased from 0.30 to 0.71, metacognition from 0.58 to 0.88, and inquiry orientation from 0.40 to 0.68. Forward-looking orientation exhibited relatively little change, and connection to learning content remained comparatively low throughout the program.

Turning to the correlation-based network structure, reflective depth and metacognition were only moderately associated in Wave 1 ($r = .27$), and the relationship between metacognition and inquiry orientation was almost absent ($r = .04$). By Wave 15, these associations increased substantially to $r = .50$ and $r = .26$, respectively. These findings indicate that students' reflections evolved from relatively descriptive accounts toward more integrated forms of reflective and inquiry-oriented thinking, with the internal structure increasingly organized around metacognitive awareness.

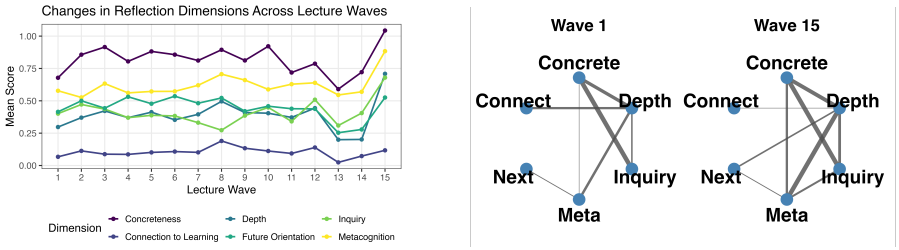


Fig. 1: Evolution of reflective structure. Left: mean scores of the six reflection dimensions across waves. Right: correlation-based networks of reflection dimensions in Wave 1 and Wave 15, where edges represent Pearson correlations between dimension scores with $|r| \geq .10$ (not ENA). Node positions are fixed across waves.

5 Discussion and Future Research

The preliminary results suggest that students' reflections became increasingly organized around reflective depth and metacognitive awareness, with associations among these dimensions and inquiry orientation strengthening throughout the program. Students thus gradually moved beyond describing classroom activities toward more integrated forms of reflective thinking. From a quantitative ethnography perspective, reflective development is better understood as structural reorganization than as simple improvement on isolated dimensions.

Several limitations should be noted. First, the present results are based on correlation-based network analysis, and the association between feedback characteristics and reflective development (RQ3) has not yet been examined; the panel models described above will be estimated in future analyses. Second, although the LLM-assisted coding was validated against human ratings, the validation relied on a single coder, and agreement varied across dimensions; validation with multiple expert raters is needed.

Future work will also employ epistemic network analysis (ENA) to model how reflective structures evolve throughout the learning process. Unlike correlation matrices, which capture only pairwise linear associations, ENA models the co-occurrence structure of reflective practices within students' discourse and enables statistical comparison of network configurations across time points and student groups.

References

1. Carless, D., Boud, D.: The development of student feedback literacy: Enabling uptake of feedback. Assessment Evaluation in Higher Education **43**(8), 1315–1325 (2018)
2. Flavell, J.H.: Metacognition and cognitive monitoring: A new area of cognitive-developmental inquiry. American Psychologist **34**(10), 906–911 (1979)
3. Hattie, J., Timperley, H.: The power of feedback. Review of Educational Research **77**(1), 81–112 (2007)
4. Sch"on, D.A.: The Reflective Practitioner: How Professionals Think in Action. Basic Books, New York (1983)
5. Shaffer, D.W.: Quantitative ethnography. Cathcart Press Working Paper (2016)
6. Wald, H.S., Borkan, J.M., Taylor, J.S., Anthony, D., Reis, S.P.: Fostering and evaluating reflective capacity in medical education: Developing the reflect rubric for assessing reflective writing. Academic Medicine **87**(1), 41–50 (2012). `https://doi.org/10.1097/ACM.0b013e31823b55fa`
7. Wood, D., Bruner, J.S., Ross, G.: The role of tutoring in problem solving. Journal of Child Psychology and Psychiatry **17**(2), 89–100 (1976)
8. Zimmerman, B.J.: Becoming a self-regulated learner: An overview. Theory Into Practice **41**(2), 64–70 (2002)

The Social Dimension in Global Virtual STEM Teams: A Comparative Network Analysis

Lu Liang[1[0009-0002-3848-3506]], Danielle P. Espino[1[0000-0002-6885-2125]], Seung B. Lee[1[0000-0003-0149-0681]], Julia S. Gibson[2[0009-0002-4841-1805]], Anmol Kaur[1[0009-0001-6680-3813]], Nastassia N. Jones[3[0009-0000-6042-6415]], Eric R. Hamilton[1[0000-0002-2337-1328]]

[1] Pepperdine University, Malibu, CA 90263, USA
[2] American University, Washington, DC 20016, USA
[3] Southern University, Baton Rouge, LA 70807
lu.liang@pepperdine.edu

Abstract. This study conducts a discourse analysis of students' project team discussions. In the computer-supported collaborative learning (CSCL) environment, the study explores adolescents' social aspects in completing STEM tasks. Findings reveal network patterns and suggest that learners evolved from initially uncertain participants into more engaged collaborators, with significant implications for supporting digital STEM education that enhances learners' experiences through meaningful teamwork formation and identity development.

Keywords: Computer-assisted Collaborative Learning, Socialization, STEM Education, Epistemic Network Analysis.

1 Introduction

Online learning opportunities have expanded and become increasingly accessible worldwide, enabling learners from diverse cultural backgrounds to participate in virtual interactions and socialization [1, 2]. Research on computer-supported collaborative learning (CSCL) has focused on understanding social interactions in online learning communities and claimed that in online interactions, the level of engagement, the purpose of the task, and the relationship among group members are significant factors in the learner's experience and development outcome. Furthermore, evidence shows that informal perspectives, such as casual conversations, greetings, and back-and-forth curiosity in off-task conversations, actually boost the interaction and have a positive impact on engagement and social outcomes [3, 4].

This study aims to understand learner groups' online conversations and interactions in the CSCL setting through a social lens. Specifically, the study conducts comparison network analysis on project teams' discourse throughout the two cycles, aiming to answer the questions: (1) What are the significant social factors being captured within these group interactions? (2) What are the notable changes and shifts from the two different cycles over time? The study provides a robust interpretation of learners' social aspects in the CSCL environment and informs researchers and educators in developing

systematic online project designs that value the social and emotional aspects, fostering more engaging, interactive online learning environments that improve the learner experience and outcomes.

2 Methods

Participants in this project were middle and high school students from the U.S. and around the world. The project lasted 1 year and was divided into 2 cycles: cycle A, from September to December 2024, and cycle B, from January to May 2025. Participants were grouped into diverse international teams and met weekly via synchronous Zoom meetings. These meetings lasted about 60 minutes, providing teams with opportunities to share ideas, develop plans, and collaborate to solve STEM problems. The current analysis focuses on meeting conversations from 4 project teams. Each team consisted of 5 to 9 learners and provided 2 sessions per project cycle, for a total of 16 included in the analysis. Researchers examined the data and identified meaningful constructs to develop the codebook that consists of 15 codes: Task Negotiation; Task Commitment; Task Accountability; Initiative/Topic Change; Inclusion; Encountering Problem/Challenge; Suggesting; Informal Talk; Explanation; STEM Content; Facilitator Effort; Collaborative Language; Social Disposition; Questioning; Media Making (see Table 1).

Table 1. Codebook of constructs.

Code	Definition
Task Negotiation	Steering the conversation toward completing a task.
Task Commitment	Agree to complete a task without having actually done it yet.
Task Accountability	Checking to see if anyone has completed their task.
Initiative/Topic Change	Speaking up on a new topic.
Inclusion	Mentioning or bringing another student into the conversation.
Encountering Problem/Challenge	Describing a challenge or issue with either the project, group, or communication in the context of working together.
Suggesting	Team members actively sharing ideas or offering advice
Informal Talk	Casual conversations and chats outside the project.
Explanation	Describing the task, defining, and clarifying the idea.
STEM Content	Focus on STEM topics, concepts, tools, and information.
Facilitator Effort	Talks or actions taken by the facilitator to drive inclusion.
Collaborative Language	The use of inclusive pronouns to emphasize team identity.
Social Disposition	Recognizing, acknowledging, and appreciating help or guidance.
Questioning	Asking questions for clarification or to express uncertainty.
Media Making	Discussions focused on selecting and using digital tools.

Each meeting session transcript was organized and independently coded by two researchers, followed by social moderation of the final coding [5]. Furthermore, the current study used Epistemic Network Analysis (ENA), a prominent tool in quantitative ethnography, to analyze learners' discussion conversations. ENA is widely used in educational research to examine discourse patterns and relationships among constructs, offering effective perspectives for researchers to examine network visualizations that support the analysis of conversational changes over time [6].

3 Results

The network model for learners' discourse in project cycles A and B is shown in Figure 1 below (see Fig. 1). Each project team represented the unit of analysis, and each meeting session was defined as the conversation where the connections were examined. In cycle A, dominant relationships are observed between Task Negotiation and Questioning (0.12); Explanation and Questioning (0.11); and Suggesting and Questioning (0.10). At the initial stage of the project, group members have less prior collaboration experience, they spend time getting to know each other, asking questions, and seeking advice and feedback: *"I was looking to the Canva. I like the. I like all the design that you all have made. I like the light purple and blue combined. Okay, so before we can finish, I want to ask you all to start a better communication. What are your interests, by the way?"* Early-stage discussions indicate that learners were unfamiliar with the collaborative STEM virtual environment, were trying to navigate and seek advice from peers, and needed to identify their self-position and direction within the project team, while also trying to understand the expectations at the team and project levels.

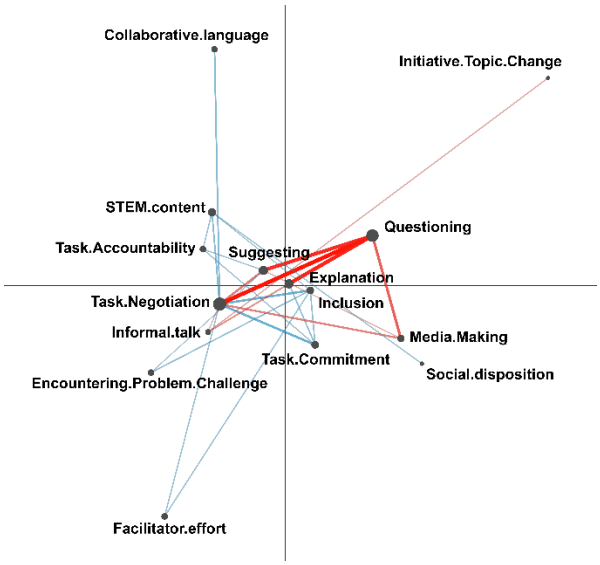


Fig. 1. ENA model for discourse in project cycle A (red) and cycle B (blue).

In cycle B, the center of learners' communication shifts from uncertainty about the project to a more focused discussion of the group task. Stronger connections are observed between Task Negotiation and Task Commitment (0.07); Task Negotiation and Inclusion (0.06); and Task Negotiation- Collaborative Language (0.05). The participants became increasingly task-oriented and emphasized collaborative problem solving. More specifically, the sense of teamwork and increased social aspects emerged. Compared to interactions in earlier cycles, learners' discourse patterns in cycle B reflected higher levels of collaborative language, inclusion, and shared commitment. For example, towards the end of the project, the learner facilitator expresses greater concerns

about ensuring group members' commitment to working collaboratively towards the project goal: *"But yeah, maybe what we can do today, since we have to make the video, I'll create a document pretty soon. And then we can type out like a script that we can give an AI, and then that AI can be like an avatar for our video to explain our website. So yeah, I'll create a doc right now and then we just work on that"*. The change indicates that learners developed a stronger awareness as active social members of the virtual community, taking on greater responsibility in group dynamics and contributing to the collective outcomes of the learning process. Learners demonstrate more frequent and stronger patterns in supporting task negotiation and commitment.

4 Conclusion

The study examines adolescent learners' social interactions and how they evolve over time in a STEM-focused CSCL environment. Findings reveal meaningful shifts in participants' sociability from an early stage of distant and ambiguous role to a more interactive, collaborative team, further confirming the effectiveness of facilitating positive social dynamics in digital education [1]. Throughout the project cycles, learners demonstrate substantial growth from the initial stage of unfamiliarity, uncertainty, and reliance on guidance to a more mature, responsible team player in virtual social learning. Furthermore, the study provides contextual evidence on informal online STEM education for adolescents and offers insights to address the unique challenges young learners face in cultivating digital literacy and online collaborative socialization, thereby supporting the design of more effective CSCL curricula. Further research should investigate additional aspects of learners' social and cultural backgrounds across more diverse groups and scaffold the various factors contributing to the development of a sustained social learning experience in the digital age.

Acknowledgments. The authors gratefully acknowledge funding support from the US National Science Foundation (Award #2215613) for the work reported in this paper. Views appearing in this paper do not reflect those of the funding agency.

References

1. Kreijns, K., Kirschner, P., Vermeulen, M.: Social aspects of CSCL environments: A research framework. Educ. Psychol. **48**(4), 229–242 (2013).
2. Mäkelä, T., Fenyvesi, K., Kankaanranta, M., Pnevmatikos, D., Christodoulou, P.: Co-designing a pedagogical framework and principles for a hybrid STEM learning environment design. Educ. Technol. Res. Dev. **70**(4), 1329–1357 (2022).
3. Abedin, B., Daneshgar, F., D'Ambra, J.: Do nontask interactions matter? The relationship between nontask sociability of computer supported collaborative learning and learning outcomes. Br. J. Educ. Technol. **43**(3), 385–397 (2012).
4. Preece, J.: Sociability and usability in online communities: Determining and measuring success. Behav. Inf. Technol. **20**(5), 347–356 (2001).
5. Herrenkohl, L.R., Cornelius, L.: Investigating elementary students' scientific and historical argumentation. J. Learn. Sci. **22**(3), 413–461 (2013).
6. Shaffer, D.W.: Quantitative Ethnography. Cathcart Press, Madison (2017)

Under Pressure to Excel: Epistemic Network Analysis of Academic Identity and Wellbeing in a Malaysian High-Achieving School

Raihan Munira Moh Sani[1,2], Aini Marina Ma'rof[1],

Mohd Mokhtar Bin Muhamad[1]

[1] Universiti Putra Malaysia 43400 Selangor, Malaysia
[2] Asia Pacific University of Technology and Innovation 57000 Kuala Lumpur, Malaysia
raihan.munira@apu.edu.my

Abstract. Students at high-achieving schools (HAS) face elevated psychological risks globally, yet evidence from Malaysian contexts remains absent. This study examines how academic identity and social relationships shape emotional wellbeing among female students at a Sekolah Berprestasi Tinggi (SBT) residential school in Selangor, Malaysia. Using Quantitative Ethnography (QE) and Epistemic Network Analysis (ENA), we model the co-occurrence structure of four constructs including autonomy, competence, relatedness and wellbeing in interview, diary and fieldnotes discourse. Model One Pilot results (n=8) reveal a preliminary structural difference: lower secondary students show stronger COMPETENCE–WELLBEING and RELATEDNESS–WELLBEING connections, while upper secondary students show a stronger AUTONOMY–COMPETENCE connection. The ENA model demonstrated excellent fit (Pearson r=1.00; Spearman r=1.00) and explained 94% of variance (X=40%; Y=54%). Findings suggest a developmental shift in how HAS students construct academic identity and emotional wellbeing across grade levels, with implications for school-based support.

Keywords: *High-Achieving Schools*, *Emotional Wellbeing*, *Malaysia, Quantitative Ethnography, Epistemic Network Analysis*.

1 Introduction

A persistent problem in educational psychology is that students in high-achieving schools (HAS) face elevated risk for internalising problems (Luthar, Kumar, & Zillmer, 2020). Studies across the United States, Hong Kong, and Türkiye have consistently documented that rates of clinically significant anxiety and depression in these settings reach six to seven times national benchmarks (Luthar et al., 2020; Tilley & Farver, 2023; Haspolat & Yalçin, 2023). In Malaysia, 134 schools hold formal Sekolah Berprestasi Tinggi (SBT) designation under the Ministry of Education's High Performing Schools framework, yet no empirical study has examined the psychological experiences of students within them. This poster uses data from one fully residential, all-girls HAS in Selangor to investigate how academic identity and social relationships

connect structurally to emotional wellbeing, and whether this structure differs between lower secondary and upper secondary students.

2 Theory

This study is grounded in Self-Determination Theory (SDT; Ryan & Deci, 2017), which holds that wellbeing depends on the satisfaction of three universal psychological needs: competence, autonomy, and relatedness. In high-achieving school (HAS) environments, academic demands, peer competition, and parental overinvolvement place systematic pressure on all three needs. SDT thus predicts that the degree to which these needs are blocked or fulfilled rather than academic performance alone, will explain students' emotional wellbeing outcomes.

3 Methods

3.1 Data

Data were collected at one fully residential, all-girls HAS secondary school in Selangor. Participants were recruited through purposive homogeneous sampling across four grade levels (Form 1–4). The Model One Pilot comprises eight participants: four from lower secondary (Form 1–2, ages 13–14) and four from upper secondary (Form 3–4, ages 15–16). Participants completed semi-structured interviews (30–45 minutes each), three weekly diary entries and two years of ethnographic observation. All interviews were audio-recorded, transcribed verbatim, and verified manually. Participants wrote diary entries in either English or Bahasa Malaysia.

3.2 Codebook

Codes were developed using a combined deductive–inductive approach, grounded in SDT and refined through Model Zero. Table 1 presents the final codebook with inter-rater reliability (Krippendorff's alpha, minimum threshold: 0.70).

Code	Definition	Example	κ (R1 vs R2)
AUTONOMY	Agency or direction in learning	*" I control my own academics."*	0.82
COMPETENCE	Evaluates academic ability relative to standards or peers	*" I feel insecure with people higher in ranking."*	0.85
RELATEDNESS	Quality of connection with friends, teachers, or parents	*" Even the friend who helps me trusts in me."*	0.80
WELLBEING	Positive or negative emotional state	*" I feel extremely stressed."*	0.83

3.3 Analysis

Data were segmented into lines (single coherent utterances) and grouped into windows by interview section (Autonomy, Competence, Relatedness, Wellbeing) and diary entry. Each line was coded using binary coding (1=present, 0=absent). The ENA model used the following parameters: unit of analysis = individual participant; horizon = interview section / diary entry (Shaffer & Ruis, 2026); moving window = 4 lines. Co-occurrences were accumulated per participant, normalised, and projected into a shared ENA space using means rotation (MR), with the primary dimension (MR1) defined to maximise separation between lower and upper secondary groups. The two groups were compared on projected positions along MR1 and SVD2 using Mann-Whitney U tests. Goodness of fit was assessed using Pearson and Spearman correlations between projected and original co-occurrence values. All analyses were conducted using the ENA Web Tool (epistemicnetwork.org; Shaffer et al., 2016).

4 Results

4.1 Qualitative Analysis

The interview and diary data revealed distinct patterns in how lower secondary (Form 1–2) and upper secondary (Form 3–4) students connected AUTONOMY, COMPETENCE, RELATEDNESS, and WELLBEING in their discourse.

Lower secondary students predominantly linked COMPETENCE and RELATEDNESS as inseparable, with WELLBEING closely tied to both: *"I do feel insecure because there are people who are higher [in achievement]. My friends don't feel any difference, but I feel like there is a deficiency in me. [COMPETENCE, RELATEDNESS]"* Positive wellbeing surfaced when peer relationships felt supportive, while negative wellbeing appeared when competition intensified.

In contrast, upper secondary students showed a stronger AUTONOMY–COMPETENCE connection, articulating academic identity through self-directed effort rather than social comparison: *"For me now, it is as though I myself am the one who will control my academics. Because it is like I know what I want. I know I want to get 9A+. [AUTONOMY, COMPETENCE]"* Here, WELLBEING was less bound to RELATEDNESS and more tied to personal progress and goal attainment.

4.2 ENA Analysis

Figures 1a and Figure 1b show the mean network graph of lower secondary students respectively. Lower secondary students' networks were dominated by COMPETENCE–WELLBEING and RELATEDNESS–WELLBEING; upper secondary students displayed a relatively stronger AUTONOMY–COMPETENCE connection. Both groups shared a prominent COMPETENCE–RELATEDNESS connection, consistent with a sustained centrality of peer dynamics across grade levels.

The difference network (Figure 1c), identifies the x-axis (MR1) as capturing the primary structural distinction between the two groups. Along the x-axis, a Mann-Whitney U test showed that lower secondary students (Mdn=0.08, N=4) were not statistically significantly different from upper secondary students (Mdn=−0.03, N=4; U=13.00, p=0.20, r=−0.62). Along the y-axis (SVD2), no significant difference was found between lower secondary (Mdn=−0.04, N=4) and upper secondary students (Mdn=0.04, N=4; U=7.00, p=0.89, r=0.13). The ENA model demonstrated excellent goodness of fit (Pearson r=1.00, Spearman r=1.00 for both axes), and the two dimensions together explained 94% of the variance in the network space (X axis: 40%, Y axis: 54%).

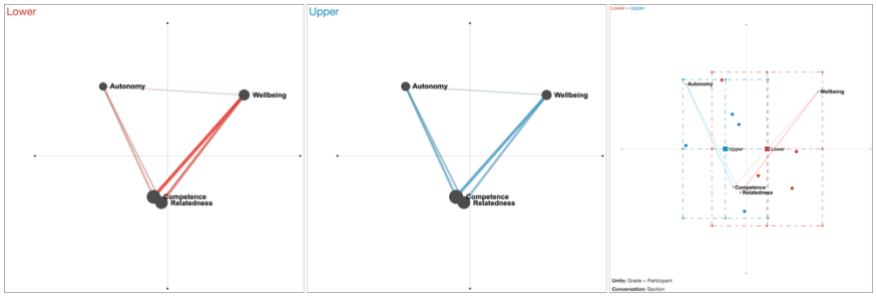


Fig. 1. ENA *model (a) lower secondary (red), (b) upper secondary (blue), and (c) difference network.*

5 Discussion

These results are broadly consistent with SDT: lower secondary students' stronger COMPETENCE–WELLBEING and RELATEDNESS–WELLBEING connections align with prior findings that relational quality buffers academic pressure in high-achieving settings (Tilley & Farver, 2023), while the relatively stronger AUTONOMY–COMPETENCE connection among upper secondary students reflects a gradual internalisation of achievement culture (Luthar et al., 2020). The large effect size (r=−0.62) alongside non-significance (p=0.20) is expected at the Model One Pilot stage (n=8) and is consistent with this study's progressive model-building protocol (Shaffer, 2017). This study applies QE to a Malaysian HAS school, showing how academic identity, social relationships, and wellbeing connect structurally in students' discourse, with implications for school-based support as data collection continues toward the full study.

References

1. Haspolat, N. K., & Yalçın, İ. (2023). Psychological symptoms in high achieving students: The multiple mediating effects of parental achievement pressure, perfectionism, and academic expectation stress. Psychology in the Schools, 60, 4721–4739. https://doi.org/10.1002/pits.23012

2. Luthar, S. S., Kumar, N. L., & Zillmer, N. (2020). High-achieving schools connote risks for adolescents: Problems documented, processes implicated, and directions for interventions. American Psychologist, 75(7), 983–995. https://doi.org/10.1037/amp0000556
3. Luthar, S. S., Suh, B. C., Ebbert, A. M., & Kumar, N. L. (2020). Students in high-achieving schools: Perils of pressures to be "standouts". Adversity and Resilience Science, 1, 135–147. https://doi.org/10.1007/s42844-020-00009-3
4. Ryan, R. M., & Deci, E. L. (2017). Self-determination theory: Basic psychological needs in motivation, development, and wellness. Guilford Press.
5. Shaffer, D. W. (2017). Quantitative ethnography. Cathcart Press.
6. Shaffer, D. W., Collier, W., & Ruis, A. R. (2016). A tutorial on epistemic network analysis. Journal of Learning Analytics, 3(3), 9–45. https://doi.org/10.18608/jla.2016.33.3
7. Shaffer, D. W., Hatfield, D., Svarovsky, G. N., Nash, P., Nulty, A., Bagley, E., Frank, K., Rupp, A. A., & Mislevy, R. (2009). Epistemic network analysis: A prototype for 21st-century assessment of learning. *International Journal of Learning and Media*, *1*(2), 33–53. https://doi.org/10.1162/ijlm.2009.0013
8. Shaffer, D. W., & Ruis, A. R. (2026). Models all the way down: Discourse and other models in quantitative ethnography. In G. Carmona et al. (Eds.), *Advances in Quantitative Ethnography: 7th International Conference, ICQE 2025* (Communications in Computer and Information Science, Vol. 2677, pp. 3–20). Springer. https://doi.org/10.1007/978-3-032-12229-2_1
9. Tilley, J. L., & Farver, J. M. (2023). Contextual predictors of internalizing and externalizing problems among adolescents from high-achieving private schools in Hong Kong. School Psychology International, 44(3), 344–363. https://doi.org/10.1177/01430343221133499

Understanding Social Presence in Hybrid Learning: Teachers' Perspectives Across Educational Levels

Azusa Nakata[1][0000-0002-2845-3920], Mervi Heikkinen[1][0000-0002-8416-2319], Hanna Järvenoja[1][0000-0003-4816-5392], and Karel Kreijns[2][0000-0002-8714-4415]

[1] Faculty of Education and Psychology, University of Oulu, Oulu, Finland
[2] Faculty of Educational Sciences, Open Universiteit, Heerlen, Netherlands
azusa.nakata@oulu.fi

Abstract. This small-scale exploratory study investigates (pre-service) teachers' perspectives on the role of social presence in hybrid learning, which integrates onsite and online participation. Participants included higher education students enrolled in pedagogical courses preparing them for teaching at the university level ($N = 12$) and the K–12 level ($N = 13$). Focus group interviews and Epistemic Network Analysis were employed to explore conceptual connections of emerged themes around hybrid learning and teacher and student presence across the two groups. Preliminary findings suggest that participants preparing for different educational levels value students' and teacher's presence in hybrid settings differently, particularly in relation to pedagogical, cognitive, and relational dimensions of learning. These findings indicate that teachers should intentionally design hybrid learning scenarios that account for diverse learner needs and educational contexts.

Keywords: Epistemic Network Analysis, Hybrid learning, Teacher training, Social presence

1 Background

Educational landscape has reshaped after the pandemic across all educational levels, and hybrid learning has emerged as a transformative solution to enhance educational opportunities [1, 2, 3, 4]. By bridging physical and virtual spaces, hybrid learning blurs the boundaries between spaces [1], enabling students to participate simultaneously in shared learning activities regardless of location [2].

Despite its potential, hybrid learning remains an emerging practice that requires further empirical investigation across diverse contexts [2]. A key challenge lies in the limited interaction across physical and virtual spaces, often due to a lack of *social presence*—defined as "the psychological phenomenon of perceiving other persons in technology mediated communication as "real" and with whom one can connect" [3, p.1]. Existing research point that social presence plays a crucial role in supporting both cognitive and social processes in hybrid learning [4, 5]. However, the optimal level and forms of its manifestation remain unclear. To address this gap, the present small-scale

exploratory study examines how (pre-service) teachers preparing for different educational levels (higher education and K–12) conceptualize the role of social presence in hybrid learning environments.

2 Methods

2.1 Participants and context

The study was conducted within two university-level courses designed to prepare future teachers in Finland. The first course targeted university staff members, including PhD students, who currently teach or intend to teach in higher education (N = 12; M_{age} = 35.6 years; 7 men, 4 women, 1 other). The second course was for pre-service teachers to become K–12 classroom teachers (N = 13; M_{age} = 22.0 years; 1 men, 9 women, 3 other). Participation in the study was voluntary. A data privacy notice was provided, and written informed consent was obtained prior to the study.

The researchers, who also served as instructors, delivered lectures and facilitated hands-on learning activities in simulated hybrid environments by dividing students into two locations. The main teacher provided a lecture on hybrid learning and the concept of social presence to ensure participants' shared understanding for responding to subsequent questions in the focus group interview.

2.2 Data collection and analysis

Focus group interviews were conducted in groups of 6–7 participants (a total of four groups) and facilitated by the researchers using semi-structured questions about teacher and student presence such as "*How do you see the presence of other students in the distant classroom as a student?*". Audio data were collected using a Røde wireless microphone and automatically transcribed using the dictation function in Microsoft Word. A research assistant checked the transcription accuracy and made some corrections if necessary.

The resulting text data (167 segments, excluding interviewer speech) were first analysed using thematic analysis [6]. Two independent coders conducted the analysis. The initial coding scheme was developed inductively by the first coder, after which the second coder reviewed the data using this scheme and suggested refinements. Following discussion and agreement, iterative refinement resulted in eight themes (see Results). Interrater agreement ranged from moderate to substantial across codes.

Further analysis was conducted using Epistemic Network Analysis (ENA) [7] with the ENA web tool (version 1.7.0). ENA was selected as an appropriate method for examining the relationships between the identified hybrid learning themes and teacher and student presence across the two groups. Groups and individual speakers were defined as the units of analysis, and each interview transcript was treated as a conversation with a window size of five. A Mann–Whitney U test was used to assess whether the two groups differed significantly in their network structures. To focus the analysis on the most relevant relationships, only connections involving teacher and student presence were selected, following the masking technique described in [8].

3 Results

Thematic analysis identified eight themes in participants' responses regarding teacher and student presence in hybrid learning settings: *Technological affordance* (TECH), *Privacy and security* (PRIV), *Pedagogical choice* (PEDA), *Inclusivity and equality* (INCL), *Cognitive focus* (COG), *Social interaction and relations* (SOCL), *Emotional responses* (EMO), and *Motivation and self-regulation* (MOTI). These themes served as the analytical framework for further analysis.

ENA further illustrated how (pre-service) teachers connected these themes to teacher and student presence in hybrid learning (Figure 1). A Mann–Whitney U test confirmed that there is a significant difference between the higher education (HE) group (Mdn = −.26, N = 12) and the K–12 group (Mdn = .28, N = 12), U = 144.00, p < .001.

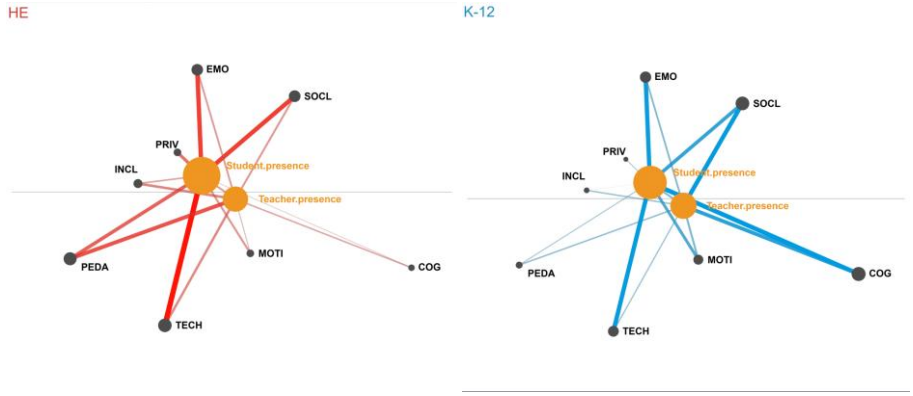


Fig. 1. ENA models for the HE group and the K-12 group. The node size indicates code frequency, while edge thickness represents the strength of co-occurrence between codes.

The ENA results revealed distinct patterns between the groups. The HE group more often linked teacher and student presence to *pedagogical choices* referring to instructional decisions and teaching practices, whereas the K–12 group more strongly associated both forms of presence with *cognitive focus* referring to students' attention and concentration during learning activities. In addition, the K–12 group emphasized *social interaction and relationships*, particularly in relation to teacher presence.

These differences are illustrated in participants' reflections. For example, one participant in the HE group highlighted the pedagogical alignment with learning needs:

> "I think alignment with content matters a lot here (…) If the class is mostly about lecturing and discussions, then physical presence isn't so necessary for me, and I can do hybrid or remote teaching where I can go into different breakout rooms." (29 years, man, HE group)

Another participant in the K-12 group emphasised the relational dimension of teacher–student interactions:

> "My point was that face to face is always kind of intimate, let's say in zoom, you can't really talk one-on-one, but in class you can go closer have discussion with

kids and see where they go through the books and materials. Whereas in online, you're just delivering it and hoping for the best." (24 years, woman, K-12 group)

4 Discussion and Conclusion

This study explored the role of social presence in hybrid learning from (pre-service) teachers' perspectives on hybrid learning. ENA provided an initial understanding of how different aspects of hybrid learning relate to teacher and student social presence across educational contexts. The findings build on prior studies highlighting the role of social presence in students' perceived learning [4] and instruction [5]. Future research could further examine gender- and age-related differences in how social presence is conceptualised across pedagogical and technological competencies [9]. To conclude, effective hybrid learning requires educators to flexibly support social interaction and cognitive engagement tailored to varying educational contexts and learning needs.

Disclosure of Interests. The authors have no competing interests to declare that are relevant to the content of this article.

References

1. Xiao, J., Sun-Lin, H., Lin, T., Li, M., Pan, Z., Cheng, H.: What makes learners a good fit for hybrid learning? Learning competences as predictors of experience and satisfaction in hybrid learning space, British Journal of Educational Technology, **51**(4), 1203–1219 (2020)
2. Raes, A.: Toward an Integrated Understanding of Designing for Hybrid Learning and Teaching. In Hyb-IT-up! Designing Effective Hybrid Teaching and Learning (pp. 235–246). Cham: Springer Nature Switzerland (2026).
3. Kreijns, K., Yau, J., Weidlich, J., Weinberger, A.: Towards a comprehensive framework of social presence for online, hybrid, and blended learning. Frontiers in Education, **8**, p. 1286594. Frontiers Media SA (2024)
4. Wagner, M., Pishtari, G., Ley, T.: The role of learning activities and social presence in synchronous hybrid classrooms: results from a natural experiment. Interactive Learning Environments. **33**(3), 1–19 (2024)
5. Szeto, E., Cheng, A. Y.: Towards a framework of interactions in a blended synchronous learning environment: what effects are there on students' social presence experience?. Interactive Learning Environments. **24**(3), 487–503 (2016)
6. Braun, V., Clarke, V.: Conceptual and design thinking for thematic analysis. Qualitative psychology, **9**(1), (2022)
7. Shaffer, D. W., Collier, W. Ruis, A. R.: A Tutorial on Epistemic Network Analysis: Analyzing the Structure of Connections in Cognitive, Social, and Interaction Data. Journal of Learning Analytics. **3**(3), 9–45 (2016)
8. Liu, Z., Li, X., Tan, Y.: Code Masking: K-Partite ENA with Multi-Mode Epistemic Codes. In proceedings on International Conference on Quantitative Ethnography, pp. 95–99. Springer, Mexico (2025)
9. Scherer, R., Siddiq, F., Howard, S. K., Tondeur, J.: Gender divides in teachers' readiness for online teaching and learning in higher education: Do women and men consider themselves equally prepared?. Computers & Education, **199**, p.104774 (2023)

Using Epistemic Network Analysis to Examine How Identity Feedback Can Foster Transformative Learning

Neta Sheffer[1] [0009-0004-5006-4333] and Yotam Hod[2][0000-0001-7353-4052]

[1] University of Haifa, Israel, netahe06@gmail.com
[2] University of Haifa, Israel, yotamhod24@gmail.com

Abstract. This study examines how identity feedback fosters transformative learning by integrating the identity artifacts framework with Epistemic Network Analysis. Using a case study of a doctoral student in a Humanistic Learning Community, we coded 228 utterances across 14 reflective diary entries and modeled five relational dimensions of identity feedback using ENA. Results reveal three patterns of identity feedback categories that co-occur with student transformation, patterns invisible through narrative analysis or descriptive statistics alone. This work responds to Shaffer's challenge of operationalizing identity empirically, offering both theoretical and methodological contributions to the learning sciences.

Keywords: Epistemic Network Analysis; Identity Artifacts; Identity Feedback; Humanistic Learning Communities; Higher Education.

1 Introduction

Fostering identity transformation is a central goal in sociocultural research, yet operationalizing it empirically remains a challenge because identity is simultaneously performative, subjective, and continually reconfigured in discourse (Shaffer, 2021). Shaffer (2021) argues that overcoming this requires strict alignment among identity theory, a supportive learning environment, and appropriate data-collection methods. This study responds to that challenge by integrating the identity artifacts framework (Hod & Dvir, 2022) with Epistemic Network Analysis (ENA). Building on prior work on identity feedback (Singer, 2022), we ask: (1) *How can ENA reveal relationships among identity feedback types, categories, and themes that descriptive statistics and narrative analysis miss*? (2) *What network patterns underlie learner identity transformation*?

2 Background

Hod and Dvir (2022) developed the identity artifacts framework to describe a learner's identity story across contexts and over time. The framework tracks five artifact types: S1 (stories from outside the community; "there-and-then"), S2 (stories from within community activities; "here-and-now"), and S3 (identity feedback intended to support growth). Singer (2022) categorized S3 feedback into three types—Support/Encourage, Raise Reflection, and Suggest a Different Path—expressed through eight specific categories (e.g., positive regard, empathy,

questioning). Finally, the framework captures outcomes through transformations enacted either within the community (S4) or in everyday life outside it (S5).

Humanistic Learning Communities (HLCs) extend the traditional knowledge building framework (Scardamalia & Bereiter, 2014) with a person-centered layer, creating a rich qualitative reservoir of these identity artifacts (Hod & Ben-Zvi, 2018). ENA provides the ideal methodological tool for this environment by modeling temporal co-occurrences of coded discourse into weighted network graphs. This allows researchers to visualize, quantify, and compare complex relational structures that traditional frequency counts and narrative analyses cannot capture (Csanadi et al., 2018).

3 Methods

This study took place in a required first-year PhD seminar (N=19, ages 29–53) at the University of Haifa, designed as an HLC combining weekly Zoom sessions with asynchronous Google Drive reflective diaries. This paper focuses on Robert (pseudonym), a student selected via purposeful sampling for exhibiting significant identity transformation. The data corpus consists of all 14 of Robert's diary entries plus inline comments from seven peers and moderators, parsed into 228 meaningful utterances. Using the identity artifacts framework, 212 utterances were coded into five artifact stages (IRR = .94), and S3 artifacts were sub-coded by feedback type and category (IRR = .96), resolving disagreements through consensus.

ENA was applied using the ENA Web Tool to model five relational dimensions across Robert's trajectory, focusing on the interplay between feedback types, categories, themes, and transformations (S4/S5). Units of analysis were individual diary notes indexed by date, stage, theme, and authorship, with conversations defined as all notes within a given identity stage. The analysis compared an Initial phase (Weeks 1–5) with a Transformed phase (Weeks 6–13), utilizing singular value decomposition for dimensional reduction to visualize the shifting structural connections in the discourse.

4 Results

Our analysis revealed six distinct identity themes regarding Robert. Three of these themes exhibited clear evidence of transformation (S4/S5), reflecting fundamental shifts in how he relates to others: (1) *Supportive Person* theme shifted from giving-only to also receiving community support; (2) *Rationalist* theme moved from emotional suppression to openness and emotional sharing; and (3) *Yearning but Insecure* theme transformed from a fear of online isolation to a deep sense of belonging. Conversely, the remaining three themes (*Passion for Learning*, *Persistence*, and *Receiving Feedback*) occupied less discourse time and showed no signs of transformation.

Of the 212 coded artifacts, 95 were identified as instances of identity feedback (S3). The co-moderator, Rachel, provided 73% of these instances, rendering the dataset largely a structured dialogue between Robert and a single facilitator. This feedback was distributed across three primary categories: 66% focused on supporting and encouraging, 27% aimed to raise reflection, and 6% suggested a different path. Critically, an inverse relationship emerged between feedback volume and transformation evidence. While the *Supportive Person* theme received the highest volume of feedback (n = 25) it yielded few transformative moments (n = 4), whereas *the Rationalist* theme (n = 16 feedback instances, n = 7 transformations) and the *Yearning but Insecure* theme (n = 13 feedback instances, n = 15 transformations) showed progressively higher concentrations of change despite lower feedback counts. Because these category distributions across themes lacked a consistent, predictive pattern for transformation, ENA was deployed to model the underlying relational structures.

Figure 1 points to the interplay between identity feedback categories that lead to an identity transformation. The connection between the identity feedback categories and Robert's identity transformation was a new finding that was not present through the two other types of analyses (narrative nor descriptive). In the initial part, we can see that the feedback category *provide positive regard* was given along with the category *ask a question*, suggesting that this connection set up the conditions for Robert's transformation later on. In the transformed period, it can be seen that *provide positive regard*, *show empathy*, *ask a question* and *illumination and clarification* had large co-occurances as they related to identity transformation inside the community (S4). Figure 1 not only shows the frequencies of the identity feedback categories, but also offers a new perspective about the way they should be given to lead to transformation and growth of the student.

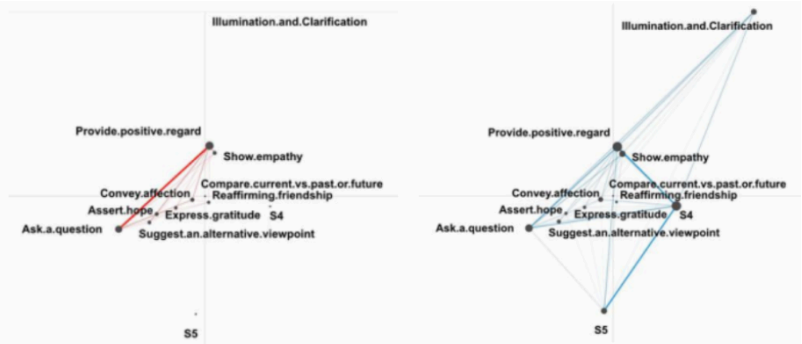


Figure 1. Categories and identity transformation (left: initial; right: transformed).

5 Discussion and Conclusion

While only three of Robert's six identity themes transformed, ENA revealed that these changing themes shared a common structural characteristic: they represented interconnected, active tensions (e.g., his rational disposition reinforcing support

avoidance, deepening his online isolation). The HLC participants recognized these internal tensions and directed identity feedback to bridge the gap between his intended goals and actual outcomes (Nicol & Macfarlane-Dick, 2006). Methodologically, this study demonstrates that the inverse relationship between feedback volume and transformation evidence cannot be explained by frequency data alone; networks, rather than sheer counts, drive transformation (Csanadi et al., 2018). ENA illuminated the specific co-occurrence patterns—such as providing positive regard anchoring reflective and empathic probes—that characterized these transformative moments.

This dynamic underscores how the HLC environment shapes the broader feedback ecology, supporting the view that feedback efficacy is inseparable from its context (Hattie & Timperley, 2007). The HLC design provided a safe space that built the qualitative reservoir of identity artifacts necessary for rich feedback (Hod & Ben-Zvi, 2018). The high volume of supportive feedback in the initial phase established the deliberate conditions needed to prevent defensiveness when challenging, reflective feedback arrived later (Rogers, 1961/1995). Ultimately, this work demonstrates that Shaffer's (2021) three-part challenge can be met. By aligning discursive identity theory, a purposefully designed environment, and reflective diary data, ENA provides the statistical validity required to uncover the microgenetic processes of identity transformation that narrative and descriptive statistics miss.

References

1. Csanadi, A., Eagan, B., Kollar, I., Shaffer, D. W., & Fischer, F. (2018). When coding-and-counting is not enough: Using epistemic network analysis (ENA) to analyze verbal data in CSCL research. *International Journal of Computer-Supported Collaborative Learning*, *13*(4), 419–438.
2. Hattie, J., & Timperley, H. (2007). The power of feedback. *Review of Educational Research*, *77*(1), 81–112.
3. Hod, Y., & Ben-Zvi, D. (2018). Co-development patterns of knowledge, experience, and self in Humanistic Learning Communities. *Instructional Science*, *46*(4), 593–619.
4. Hod, Y., & Dvir, M. (2022). Identity artifacts: Resources that facilitate transforming participation in blended learning communities. *Internet and Higher Education*, 54, 100846.
5. Nicol, D. J., & Macfarlane-Dick, D. (2006). Formative assessment and self-regulated learning: A model and seven principles of good feedback practice. *Studies in Higher Education*, *31*(2), 199–218.
6. Rogers, C. R. (1961/1995). On becoming a person: A therapist's point of view of psychotherapy. Houghton Mifflin.
7. Scardamalia, M., & Bereiter, C. (2014). Knowledge building and knowledge creation: Theory, pedagogy, and technology. In R. K. Sawyer (Ed.), *The Cambridge handbook of the learning sciences* (2nd ed., pp. 397–417). Cambridge University Press.
8. Shaffer, D. W. (2021). Operationalizing identity: Studying changing selves in experimental learning environments. *The Journal of Experimental Education*, *89*(3), 560–567.
9. Singer, M. (2022). The role of identity feedback in fostering doctoral students' transformations in research communities [*Unpublished master's thesis*]. University of Haifa.

When Does GenAI Mediate Knowledge Building? Tracing Uptake Patterns in Collaborative Learning Discourse

Bunichi Otaki, Ritsuko Oshima, and Jun Oshima
Shizuoka University, 3-5-1 Johoku, Chuo-ku, Hamamatsu 432-8011, Japan
otaki.bunichi.23@shizuoka.ac.jp

Abstract. Generative AI produces fluent text on demand, but fluency does not by itself indicate knowledge-building learning. The question is when and how AI output is taken up as an improvable idea rather than received as a finished product. We trace GenAI uptake in ten undergraduate pairs writing on genetically modified foods. Using Knowledge Building Discourse eXplorer (KBDeX), we trace vocabulary-network degree-centrality trajectories, excluding AI utterances to isolate student-mediated integration, and analyze three cases through knowledge-creating dialogue. Three uptake patterns emerge: AI output is received with little engagement, partially appropriated as a writing resource, or critiqued, compared, and reformulated. GenAI mediates knowledge building when peers act on its output through chained discourse moves.

Keywords: Generative AI, Collaborative Learning Discourse, Knowledge Building, Quantitative Ethnography

1 Introduction

When students work collaboratively with generative AI (GenAI), the tool supplies polished text just when they would otherwise need to work out a shared position. Yet knowledge building advances through discourse in which learners treat ideas as improvable objects of inquiry, not finished pronouncements [1]; fluent AI output is not itself such discourse. It may stimulate inquiry, or quietly replace the discourse it appears to support. Offloading to GenAI is associated with fewer metacognitive processes and weaker knowledge transfer, despite gains in immediate performance [2]. The question is therefore not whether students used GenAI, but *when* and *how* its output entered peer discourse, as an improvable idea or a finished product.

This requires separating two layers. Counting AI utterances as equivalent units inflates indicators, since AI responses are longer and lexically richer than student talk; excluding them entirely hides the mediation we seek. Multi-layered temporal network analysis [3] shows that idea improvement stays hidden unless meaningful layers are distinguished over time. We therefore trace uptake by treating peer discourse and AI contribution as related but separable layers, with two conceptual moves.

First, rather than analyzing GenAI output in its own right, we center students' peer discourse and treat each AI response as material it works on, an epistemic object constituted by how students engage with it. Second, we use Good Moves in knowledge-creating dialogue [4], the turn-level moves by which knowledge-building communities

advance ideas, namely Problem Definition (PD), New Ideas (NI), Promisingness Evaluation (PE), Comparison (CP), Critical Discourse (CD), and Meta-dialogue (MD), to characterize what students do following AI responses. We address two questions:

RQ1: How do temporal vocabulary-network trajectories differ across student pairs in GenAI-supported collaborative writing?

RQ2: How do students differently take up GenAI outputs into knowledge-building discourse?

2 Methods

2.1 Participants and data collection

Twenty undergraduate students at a national university in Japan, in ten pairs, completed a 60-minute collaborative writing task: proposing problems and solutions about genetically modified foods (approx. 800 Japanese characters). Because students lacked sufficient prior knowledge, the task required information exploration and dialogic negotiation. All pairs used ChatGPT in Google Docs. Half were also offered optional Good Moves cards as part of a broader scaffolding design, treated here as a contextual resource rather than an experimental variable.

2.2 Data analysis

Stage 1: Trajectory analysis of peer discourse. Using KBDeX [5], we built a temporal social-semantic network of how the content-word keywords from each pair's final product co-occur within students' peer utterances. Keyword sets were pair-specific and varied in size (Groups 1–10: 50, 57, 59, 89, 94, 55, 82, 73, 98, and 83 keywords; M = 74). Figure 1 displays raw cumulative degree-centrality trajectories to preserve the temporal pattern of change within each pair. Because cumulative values can vary with keyword-set size, we additionally standardized each trajectory by dividing cumulative degree centrality at each turn by the number of keywords in that pair's network. Standardization rescales the vertical magnitude but does not alter the within-pair trajectory shape. Accordingly, the raw trajectory bands shown in Figure 1 are treated descriptively rather than as directly comparable standardized categories across pairs. AI utterances were excluded because our question was whether students integrated these keywords through peer dialogue rather than whether the keywords appeared in AI output.

Stage 2: Trajectory-informed episode analysis. We selected three contrasting cases that occupied progressively higher positions in both the raw trajectories and the standardized sensitivity check: Group 1 (.042), Group 4 (.055), and Group 9 (.070). All three pairs had access to the Good Moves cards. Based on the Stage 1 trajectories, we identified GenAI-use episodes around salient temporal features, including stagnation, moderate increase, and sharper transition. Each episode comprised pre-prompt peer discussion, the prompt to GenAI, the GenAI response, post-response peer discussion, and incorporation into the final written product. A single coder systematically applied Good Moves codes to post-response peer discussion within the selected episodes.

3 Results

3.1 Trajectory patterns (RQ1)

KBDeX trajectories showed three descriptive bands (Figure 1): lower (<3), intermediate (3–6), and higher (>6) final values. Because keyword-set sizes differed across pairs, these raw magnitudes are interpreted as within-pair temporal patterns rather than directly comparable between-pair differences. A keyword-adjusted sensitivity check changed some pairs' relative positions but preserved the ordering of the three focal cases. Because AI utterances were excluded, the trajectories reflect student-mediated integration of product vocabulary rather than AI lexical contribution.

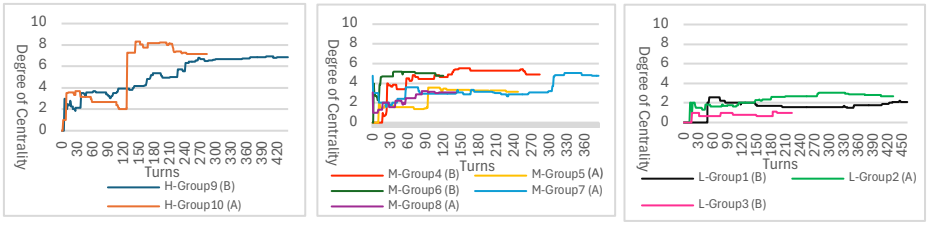


Fig. 1. Degree-centrality trajectories in peer discourse by group; AI utterances excluded. (A) without Good Moves cards, (B) with.

3.2 Uptake patterns (RQ2)

Episode analysis revealed three qualitatively distinct uptake patterns.

Lower focal case (Group 1). Students entered multiple card-derived questions into GenAI in sequence (cf. "Are there new perspectives to add?"). GenAI returned dense terms: traceability, blockchain, genome editing, co-creative governance. Post-response talk did not engage them: students commented that the output "doesn't sink in". NI were externalized to AI; CP, CD, and PE were absent. The flat trajectory tracks this absence of student-mediated integration.

Intermediate focal case (Group 4). Prompting GenAI for new perspectives, students received concepts including coexistence resilience. After initially calling it "formulaic," they treated it as a candidate organizing frame, asking GenAI for elaboration and incorporating it into writing. NI and partial PE/MD appear, but CD questioning the concept's basis or adequacy, is limited. AI output entered peer discourse as a writing resource yet was only partly transformed.

Higher focal case (Group 9). Students had already organized the problem space (biodiversity, foreign corporate seed control, food safety, labeling) before prompting. When GenAI proposed that labeling improves social transparency, they questioned the empirical basis ("but is that really happening yet?") and requested concrete examples, that is, CD on AI output. Prompted later about integrating opposing positions, they engaged in chained CP, CD, and PE on GenAI's suggestions and reformulated them into a novel proposal, benefit labeling as more consumer-accessible alternative to technical transparency. The sustained high trajectory tracks this dense Good Moves activity.

4 Discussion and Conclusion

The three focal cases illustrate how GenAI may mediate knowledge building differently depending on how peers engage with its output. In the lower case, students tended to use GenAI to generate ideas without substantially reorganizing or evaluating them through peer dialogue. In the higher case, students had already begun structuring the problem space before prompting and subsequently compared, questioned, and refined AI responses in relation to their emerging position. These cases suggest that GenAI becomes more consequential for knowledge building when its output is treated as material for further inquiry rather than as an endpoint. At the same time, the observed trajectories should not be interpreted as direct standardized differences in network integration across pairs, because keyword-set sizes varied and pre-existing differences in problem framing may also have shaped both trajectory development and uptake.

Methodologically, separating peer discourse from AI contributions and linking temporal network patterns to episode-level analysis provides a way to examine not only whether GenAI enters collaborative discourse, but how its contributions are taken up and transformed. Overall, the findings suggest that the educational value of GenAI lies less in the quality of its output alone than in the dialogic work through which learners evaluate, connect, and improve that output.

Limitations. This study examined a small number of pairs and used qualitative episode analysis for three focal cases, limiting generalizability. In addition, raw cumulative degree-centrality values should not be interpreted as directly comparable measures of network integration across pairs; keyword-adjusted values were therefore used as a sensitivity check. The Good Moves analysis was conducted by a single coder, and pre-existing differences between pairs may also have shaped the observed trajectories. In particular, the higher-trajectory case had already begun organizing the problem space before prompting, so the differences observed cannot be attributed solely to how students took up GenAI output.

References

1. Scardamalia, M., Bereiter, C.: Knowledge building and knowledge creation. In: Sawyer, R.K. (ed.) The Cambridge Handbook of the Learning Sciences, 3rd edn., pp. 385–405. Cambridge University Press, Cambridge (2022)
2. Fan, Y., Tang, L., Le, H., Shen, K., Tan, S., Zhao, Y., et al.: Beware of metacognitive laziness: Effects of generative artificial intelligence on learning motivation, processes, and performance. British Journal of Educational Technology **56**(2), 489–530 (2025)
3. Oshima, J., Oshima, R., Kawakubo, A.J.T.: Multi-layered temporal network analysis: Idea improvement processes in knowledge-building practice leading to high learning performance. *Journal of Computer Assisted Learning* **41**(3), e70063 (2025)
4. Bereiter, C., Scardamalia, M.: "Good Moves" in knowledge-creating dialogue. QWERTY-Interdisciplinary Journal of Technology, Culture and Education **11**(2), 12–26 (2016)
5. Oshima, J., Oshima, R., Matsuzawa, Y.: Knowledge building discourse explorer: A social network analysis application for knowledge building discourse. Educational Technology Research and Development **60**(5), 903–921 (2012)

"Who Are We to the Cosmos?" Epistemic Network Analysis of Identity and Global Citizenship in Middle-School Students' Civilisation-Reboot Discourse

Pak Ho Chan[1][0009-0000-9700-5489] and Chen Chen[1][0000-0002-6065-8889]

Faculty of Education, The University of Hong Kong, Hong Kong SAR
logancc@hku.hk

Abstract. We report a formative study of whether a single fifty-minute astrosociology lesson reshapes middle-school students' connected reasoning about identity, community, and civilisation. Forty Grade 7 students at a public middle school in eastern China worked in eight five-student groups (four experimental, four control) across two parallel classes with a shared protocol. Each group discussed a common prompt: if Earth's civilisation were destroyed, which five short messages should be left to help "new humans" rebuild. Translated transcripts (508 sentences) were coded against a twelve-code identity-and-citizenship codebook, and Epistemic Network Analysis (ENA) compared mean networks. The experimental network shows a humanistic–cosmic connective structure absent from the control network even where individual codes appear in both. The poster invites discussion of the lesson, codebook, and use of an nCoder-style automated second rater on a small Mandarin–English translated corpus.

Keywords: Epistemic Network Analysis · global citizenship education · astrosociology · middle school · automated coding.

1 Introduction

Chinese middle-school education foregrounds national identity and patriotic citizenship; supranational, planetary, and species-level frames are rarely surfaced explicitly in everyday classroom discourse [3]. The global-citizenship-education literature warns that "global citizen" cannot be a slogan layered on top of local identity but must be evidenced as a *connected* way of reasoning that links the local to the planetary [5,10]. Astrosociology, the study of human social life in cosmic context [6], offers a candidate pedagogical resource: by reframing humanity at the scale of the universe, it can foreground the Cosmic Perspective and Overview Effect [11] and align with Sagan's "Pale Blue Dot" stance [7].

We ask whether a single, low-resource fifty-minute astrosociology lesson, inspired by existing university-level astrosociology teaching, produces a measurable shift in the epistemic frame through which Grade 7 students reason about who "we" are. We use Epistemic Network Analysis (ENA) [8,9] because the question

is about which identity codes co-occur in connected reasoning, not which are individually present. The contribution is formative and methodological: (a) a compact lesson-and-prompt design eliciting identity-relevant discourse in 35–40 min of small-group talk; (b) a twelve-code codebook spanning *identity-at-scale*, *survival-and-stewardship*, *meaning-making*, and *integration*; and (c) a workflow for using an nCoder-style automated second rater [1,4] when only one human coder and a translated corpus are available.

2 Method

Participants and design. Forty Grade 7 students (ages 12–13) at a public middle school in eastern China were assigned to eight groups of five (four experimental, four control) across two parallel classes taught by two facilitators using a shared protocol. Experimental groups received the astrosociology lesson; control groups received a generic astronomy lesson of equivalent length on the Solar System and nearby galactic regions, without any explicit framing of human identity, society, or civilisation at the cosmic scale.

Task and corpus. All eight groups responded to the same prompt: *"If Earth's civilisation were destroyed and only five short messages could be left for the new humans, what would you write?"* Discussions were video-recorded, transcribed in Mandarin, and translated to English by the lead author. The corpus comprises 508 utterance-level sentences (320 experimental, 188 control).

Codebook derivation. The codebook was derived through a hybrid deductive–inductive process. A deductive scaffold of four families was drawn from three literatures: global-citizenship-education typologies distinguishing local, national, and planetary loyalties [5,10]; astrosociology's Cosmic Perspective and Overview Effect [6,7,11]; and the ENA convention of designing codes so that connections between them carry the substantive claim [8,9]. Within this scaffold, individual codes and definitions were refined bottom-up through three passes of close reading [8]: anchors and non-examples were drawn only from the transcripts, and code boundaries were adjusted until every anchor could be assigned to exactly one code without residual ambiguity. Table 1 lists the twelve codes.

Reliability via automated second rater. Because a fully independent bilingual human second rater was not available for the full corpus, we deployed an nCoder-style regex classifier as Rater 2 [1,4] in three iterative rounds, audited against Cohen's κ and Shaffer's $\rho(.65)$ [2]. Round 1 seeded regex patterns for each code from the anchor excerpts. Round 2 inspected every human-vs-regex disagreement on a stratified subsample and either tightened the pattern or logged a genuine miss. Round 3 re-applied the tightened patterns to the full corpus and recomputed reliability. All twelve codes met $\kappa \geq .65$ (range .740–1.000, Mdn $\approx .90$) and passed $\rho(.65)$ at $p < .05$ (max $\rho = .023$). Full statistics will appear on the poster.

ENA modelling. Following [8,9], ENA was applied with the unit set to student nested in condition, the conversation set to the group, and whole-conversation stanza windows. Edge weights were unit-normalised; SVD rotation compared experimental and control mean networks in the ENA Web Tool.

Table 1: Twelve-code codebook with working definitions and translated anchor examples.

Family	Code	Working definition	Anchor example (translated)
Identity-at-scale	Cosmic Perspective	Humanity at the scale of the universe or life-elsewhere.	"We are only one kind of life in a huge universe."
	Global Identity	Humans, humankind, or Earth as one community.	"Tell them we were all humans, on one Earth."
	Local/National Identity	A country, region, or ethnic group as the primary "we".	"Say we were Chinese and remember our country."
Survival & stewardship	Biological Survival	Food, water, shelter, safety from immediate threats.	"Where to find clean water and food to survive."
	Reproduction	Continuation of species or family line.	"How to have children and raise the next generation."
	Ecological Stewardship	Care for environment, other species, planetary limits.	"Protect forests and the sea; don't kill animals."
	Modern Materialism	Comfort, technology, or wealth as primary goods.	"Bring back phones and cars; life was easier then."
Meaning-making	Humanistic Values	Kindness, love, peace, dignity, family bonds.	"The most important thing was love and not hurting others."
	Cultural Heritage	Language, history, art, tradition as inheritance.	"Keep our language and stories so they know who we were."
	Reflective Communication	Explaining mistakes or lessons to a future reader.	"Tell them what we did wrong so they won't repeat it."
Integration	Socio-Scientific Integration	Links social choices to scientific or ecological reasoning.	"If we use too much fuel the air gets bad; don't do it."
	Technical/Physical Fundamentals	Physics, mathematics, or engineering as rebuild-knowledge.	"Write down that water is H_2O and how electricity works."

3 Results

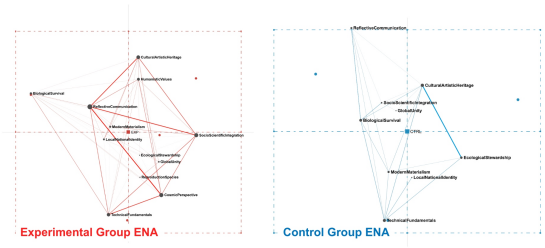


Fig. 1: Mean ENA networks. Node position fixed across conditions; edge thickness encodes normalised co-occurrence. The experimental network foregrounds a triangular structure linking Cosmic Perspective, Humanistic Values, and Global Identity that is absent in the control network.

The experimental and control mean networks differ in *how* codes co-occur rather than in *which* codes appear (Fig. 1). In the experimental condition, the strongest edges connect Cosmic Perspective with Humanistic Values and both with Global Identity; Cultural Heritage and Reflective Communication feed into this humanistic–cosmic triangle. In the control condition, edges cluster instead around Biological Survival, Modern Materialism, and Technical/Physical Fundamentals; Cosmic Perspective and Humanistic Values appear as codes but

are not connected. Individual code frequencies alone do not distinguish the conditions: the shift lives in the connective structure ENA foregrounds.

4 Discussion, Limitations, and Next Steps

The humanistic–cosmic pivot sits at three literatures' intersection: global-citizenship-education typologies [5,10] that name but rarely operationalise the local–national–planetary axis; astrosociology's Cosmic Perspective [6,11,7], argued mostly from astronaut testimony rather than classroom data; and ENA's insight that connective structure, not frequency, is where reframings become visible [8,9]. Our contribution locates a humanistic–cosmic connective structure in Grade 7 discourse after a single lesson, using automated reliability on a small translated corpus. The poster invites discussion of the lesson design, the twelve-code codebook, and the nCoder-style automated second rater used when only one human coder and a translated corpus are available. Limitations: N = 508 across 8 groups in one school; translation introduces interpretive distance; durability is untested. Next steps: replication across schools; a delayed post-measure; and a bilingual human second rater on a subsample.

References

1. Cai, Z., Siebert-Evenstone, A., Eagan, B., Shaffer, D.W., Hu, X., Graesser, A.C.: nCoder+: a semantic tool for improving recall of nCoder coding. In: Eagan, B., Misfeldt, M., Siebert-Evenstone, A. (eds.) Advances in Quantitative Ethnography. CCIS, vol. 1112, pp. 41–54. Springer, Cham (2019). https://doi.org/10.1007/978-3-030-33232-7_4
2. Eagan, B., Brohinsky, J., Wang, J., Shaffer, D.W.: Testing the reliability of inter-rater reliability. In: Proc. LAK '20, pp. 454–461. ACM (2020). https://doi.org/10.1145/3375462.3375508
3. Liang, W.: Citizenship project, national identity, and human rights in modern China. In: Human Rights Education in China, pp. 59–72. Springer, Singapore (2022). https://doi.org/10.1007/978-981-19-1304-4_4
4. Marquart, C.L., Swiecki, Z., Eagan, B., Shaffer, D.W.: ncodeR: techniques for automated classifiers (R package). CRAN (2018). https://doi.org/10.32614/CRAN.package.ncodeR
5. Oxley, L., Morris, P.: Global citizenship: a typology for distinguishing its multiple conceptions. Br. J. Educ. Stud. **61**(3), 301–325 (2013). https://doi.org/10.1080/00071005.2013.798393
6. Pass, J.: Astrosociology as the missing perspective. Astropolitics **4**(1), 85–99 (2006). https://doi.org/10.1080/14777620600762865
7. Sagan, C.: Pale Blue Dot: A Vision of the Human Future in Space. Random House, New York (1994)
8. Shaffer, D.W.: Quantitative Ethnography. Cathcart Press, Madison, WI (2017)
9. Shaffer, D.W., Ruis, A.R.: Epistemic network analysis: a worked example of theory-based learning analytics. In: Lang, C., et al. (eds.) Handbook of Learning Analytics, pp. 175–187. SoLAR (2017). https://doi.org/10.18608/hla17.015
10. UNESCO: Global Citizenship Education: Topics and Learning Objectives. UNESCO Publishing, Paris (2015). https://doi.org/10.54675/drhc3544
11. White, F.: The Overview Effect: Space Exploration and Human Evolution, 3rd edn. AIAA, Reston, VA (2014). ISBN 978-1624102622

資格取得学習・専門知学習・推しの学習それぞれの学習中に生起する内言記録のENA比較分析

A Comparative Analysis Using ENA of Inner Speech Records Occurring During Learning for Certification, Disciplinary Knowledge, and Oshi-katsu

藤田ののか[1] 益川弘如[1]
Nonoka Fujita[1] and Hiroyuki Masukawa[1][0000-0001-8840-8905]

[1] 青山学院大学, 東京都渋谷区渋谷4-4-25, 150-8366, 日本
C9126006@aoyama.jp

Abstract. This study explored whether the learning processes characteristic of *Oshi* driven learning could inform disciplinary learning and certification-oriented learning. To capture learners' ongoing thinking processes, inner speech data were collected, and three types of inner-speech records were comparatively analyzed using Epistemic Network Analysis (ENA). The findings suggest that certification-oriented learning is characterized by a relatively constrained and convergent cognitive structure, whereas *Oshi* driven learning and interest-driven disciplinary learning involve more expansive thinking that traverses multiple cognitive tags. These results provide insights into the distinctive cognitive characteristics associated with different forms of learning.

Keywords: Inner speech, First-person research, Epistemic Network Analysis

1 背景と目的

頭の中で展開される内面化された言葉である「内言」[1]に着目し，学習中に生起した思考を意識的に捉え，リアルタイムに記録する「内言記録システム」を構築，第一著者自身を研究対象とし，利用，分析をおこなった一人称研究である．一人称研究は，自身の身体知における獲得過程の言語化を対象とした諏訪[2]の研究や，自身の過去を振り返り価値付ける近藤[3]の回顧的エスノグラフィなど広がりを見せている．本研究では，いつでもデータ取得可能な一人称研究の強みをいかし，学習中に生起する内言を長期的に記録した．記録システムのツールとしてはGoogle スプレッドシートを採用し，テキストデータでの内言記録と，その内言が意味する認知状態を表す8つの認知タグの選択を学習と同時並行でリアルタイムに行えるようになっている．

学習対象には様々な種類が存在する．最終的に定められた知識技能を測定されるために学ぶ学習や専門的な知識技能を得たくて能動的に学ぶ学習，さ

らには，自身の趣味を追求していく学習もあるだろう．しかし，それら学びに対して内的に湧き上がってくる動機や感情は異なるものであり，趣味の追求ではあっという間に対象について深く理解し，主体的にもっと学習を進めようとする自分に気づく．どうすれば様々な学習において，趣味の追求のような学習が実現できるのだろうか.

本報告では，その第一歩として，３種類の異なる学習に関する内言記録データをEpistemic Net-work Analysis（ENA）で分析することで，タグの出現頻度ではなく，学習中にどのタグ同士が関連して生起するかという共起構造を比較し，それぞれの認知状態の特徴を明らかにする．

2　研究の方法

3種類の異なる学習における記録期間と記録件数は以下の通りである．

・資格取得学習（旅行業務取扱管理者資格の取得を目的として行った，資格試験のための参考書学習）：2024年5月2日～10月16日（全502件）
・専門知学習（自身の研究関心である学習科学についてより深く理解することを目的として行った専門書学習）：2025年6月11日～10月8日（全471件）
・推しの学習（SNSや公式コンテンツを通して推しのアイドルについて知ること自体が楽しいため行った学習）：2025年6月1日～9月30日（全571件）

Table 1の通り，コーディングは，内言記録と同時にタグ付けした８つのコードとした．これは内言記録と同時にその都度タグを選択しており，学習中の新鮮な認知状態を表すのに有力であると考えたためである．Horizon単位は学習中に扱われているトピックごとに設定し，学習内容のまとまりを基準として区切った．また，Windowサイズは4に設定し，unitは各学習とした．そして生成されたネットワーク図をもとに特徴的な共起関係が見られた箇所について，実際の内言記録データを参照しながら内容的検討を行った．

Table 1. 分析コードブック

コード名	定義	例
納得	自分の中で納得・了承したこと	AはBだ
理解のための疑問	得た情報自体について理解のための疑問	Aとは何だ？
派生した疑問	得た情報から派生して生まれた疑問	AならBはどうか？
理解のための言換	言い換えや説明のしなおし	つまりA‘だ
思い出し	すでに持っている情報の思い出し	前回学んだな
ひらめき（推測）	得た情報から生まれた考えや推測	恐らくAならBだ
リアクション	感想	知らなかった！
話変わるけど	学習とはまったく関係ない逸脱	お腹すいたな

3　結果と考察

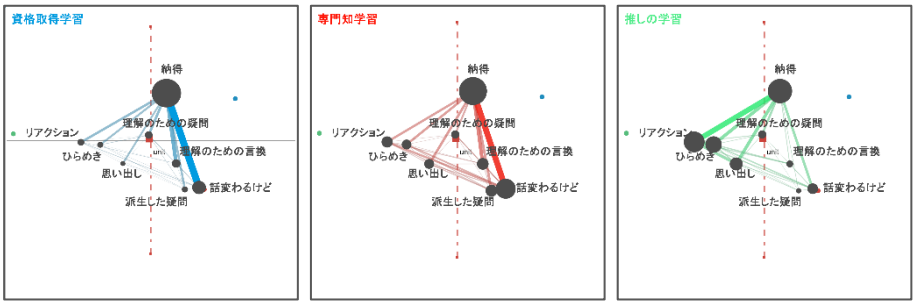


Fig. 1. 各学習のENAネットワーク図

Fig. 1は各学習におけるENAネットワーク図である．強い共起がみられた傾向を見ると，資格取得学習と専門知学習は似ており，推しの学習とは異なるが，派生した疑問と思い出し，ひらめきの共起を見ると，資格取得学習は共起が弱く，専門知学習と推しの学習は比較的共起が強い．

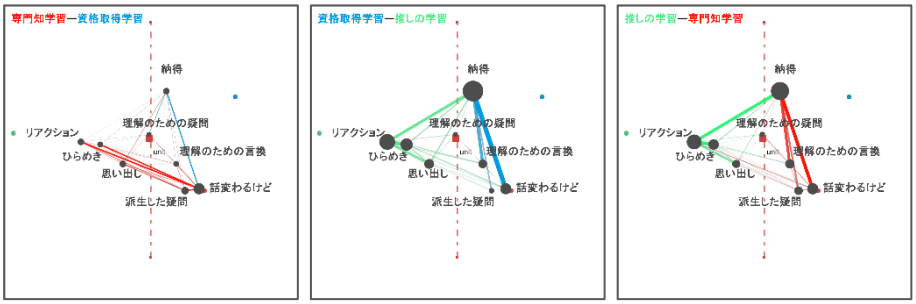


Fig. 2. ENAネットワーク図の比較

また，3種類のネットワーク図をペアごとに差分比較したFig2でも，やはり専門知学習と資格取得学習の差分比較は共起関係が似ているため細い線で結ばれた図となった一方で，派生した疑問と思い出し，ひらめきの結びつきは目立つ図となっており，この特徴は推しの学習でも見られた．このことから興味関心を契機とする学習は生まれた疑問に対して過去に得た情報を思い出したり積極的に推測しているのに対し，資格取得学習は比較的収束的なタグにとどまったことがわかった．この結果は，過去ENAを用いずに認知タグ間の遷移率を分析した研究 [4][5]でも同様であった．

ただし専門知学習においては，疑問に対して過去の知識や経験を参照しながら思考を展開する場面が事例数としては他の２学習と比べて多く確認できた．しかし，残りの約半数の共起は意味的連続性が見られず，専門知学習の認知状態の遷移では必ずしも連続的・線形的に起きないという特徴も発見された．

次に，推しの学習と専門知学習，資格取得学習の差分比較では，納得と強く共起したものの違いが似た構造で可視化され，推しの学習に対して専門知学習と資格取得学習が類似した構造を取ることが示された．

この結果は，推しの学習は学習内容が体系的に構造化されていたわけではなく，偶発的に得られる個々の情報を，学習者が主体的に結び付けて学習を進めることで理解を深めていたためだと考えられる．

そこで，推しの学習のみに見られた特徴的な共起について分析した．納得とひらめき，リアクション，思い出しタグとの強い共起の事例を追ったのがFig3である．その結果，積極的に過去に得ていた情報を思い出して結びつけたり，その際，それらから推測しながら納得を得つつ学習を進める様子が確認された．

納得	えええルウク退所してんの
思い出し	おはよう朝日のこじけんからのメッセージ
ひらめき（推測）	もしかして関係あるのでは？？？
納得	こじけんのインスタ、たろちゃんのみならずもう1人はルウクなんだ
話変わるけど	てかbgmにしてる音楽、こんなんが好きなんやな
リアクション	まじか…びっくり

Fig. 3. 推しの学習における「納得」タグとの共起していた内言の事例

これらの結果から，多様な認知状態を展開するのは，興味関心を契機とする学習であることがわかった一方，推しの学習では学習対象そのものへのより強い愛着や熱狂があるからこそ，より強く自発的に過去の知識と結び付け，さらなる情報へと展開させていく学習を成立させたのではないだろうかという考察を得た．

今後，これらの分析成果を参考に，推しの学習のような学び方を専門知学習や資格取得学習に広げる方策を考えていきたい．

利益相反の開示 本論文の内容に関連する利益相反は，著者にはありません．

参考文献

1. 中村和夫.: ヴィゴツキー心理学.新読書社,49 (2004)
2. 諏訪正樹.: 身体知獲得のツールとしてのメタ認知的言語化.人工知能学会誌 20(5) , 523–532 (2005)
3. 近藤百玲.: 主体的問題解決のための思考過程の解明の試み-自己エスノグラフィを通した探索的検討-.教育論叢 59, 19–34 (2016)
4. 藤田ののか, 齋藤ひとみ, 益川弘如.: 内言記録システムを用いた学習の検証.日本教育工学会第 46 回全国大会講演論文集, 441–442 (2025)
5. 藤田ののか, 齋藤ひとみ, 益川弘如.: 学習内容の違いによる内言記録の特徴分析.日本教育工学会第 47回全国大会講演論文集, 811–812 (2026)