Healthcare Communication Education in the Age of Generative Artificial Intelligence: A systematic review

Zilin Jia   University College London

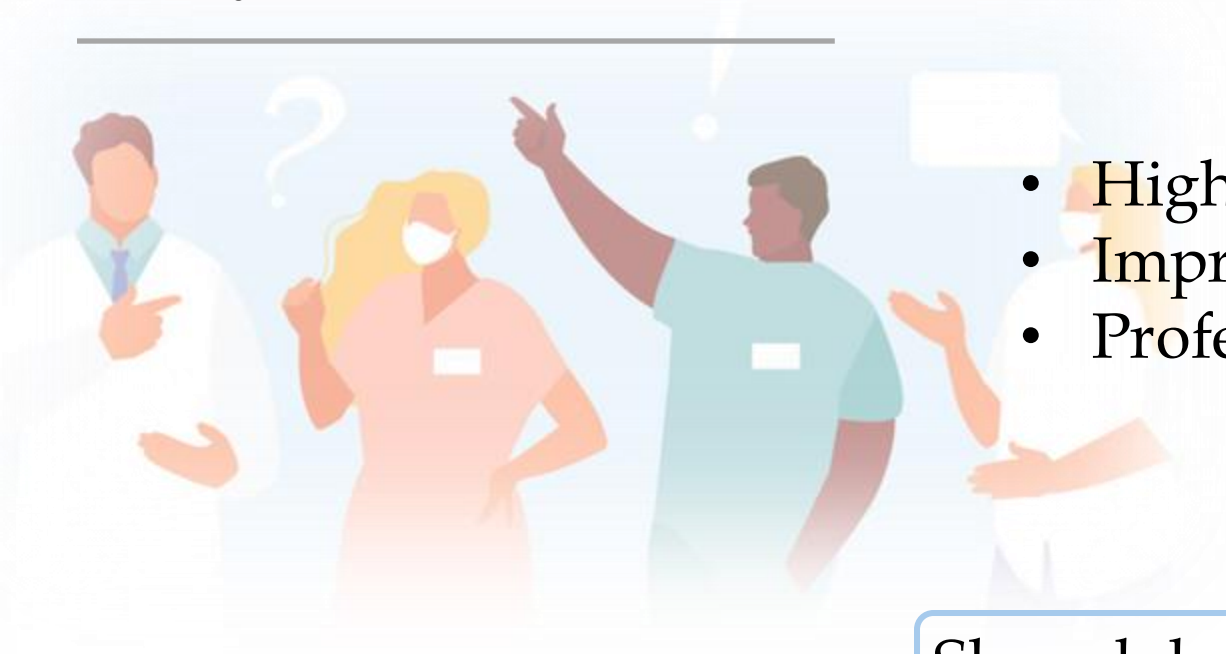Dr David Wei Dai   University College London

Dr Vahid Aryadoust   Nanyang Technological University

Prof Zhu Hua   University College London

Why healthcare communication matters ?

- High-quality patient care
- Improved healthcare outcomes
- Professional development for healthcare providers

(Dai et al., 2024; Choudhary & Gupta, 2015; Kurtz et al., 2017; Teutsch, 2003)

Shared decision-making

Paternalistic consultations → Patient-centred care

Professional communication in the intercultural contexts

Global mobility → Increasing intercultural complexity

(D'Agostino et al., 2017; Fleckman et al., 2015; Sandman & Munthe, 2010; Taylor, 2009)

Challenges in healthcare communication education

Traditional training

- E.g., seminars, workshops
- Passive engagement
- Limited scenarios
- Weak transferability

Patient-centred training

- E.g., role-plays, real/virtual patient simulations
- Active learning
- Financial & Staff resources
- Limited scenarios
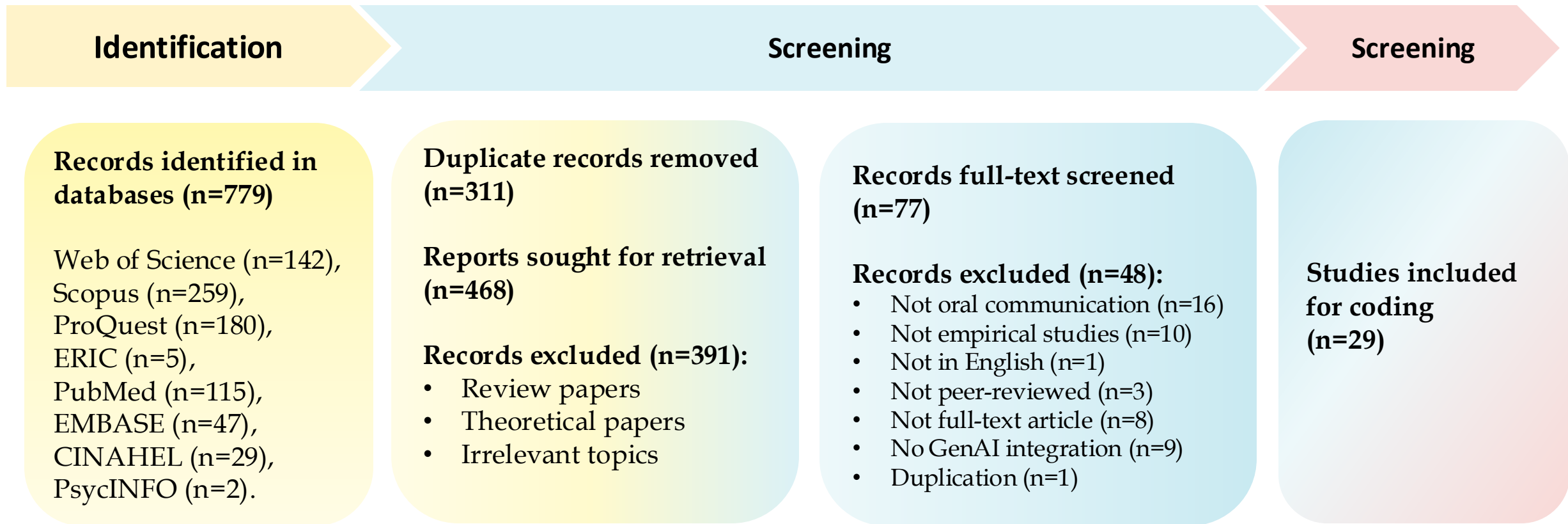
GenAI-mediated training

- E.g., LLM-powered virtual patient simulations
- Active learning
- Diverse scenarios
- *Fragmented empirical evidence*

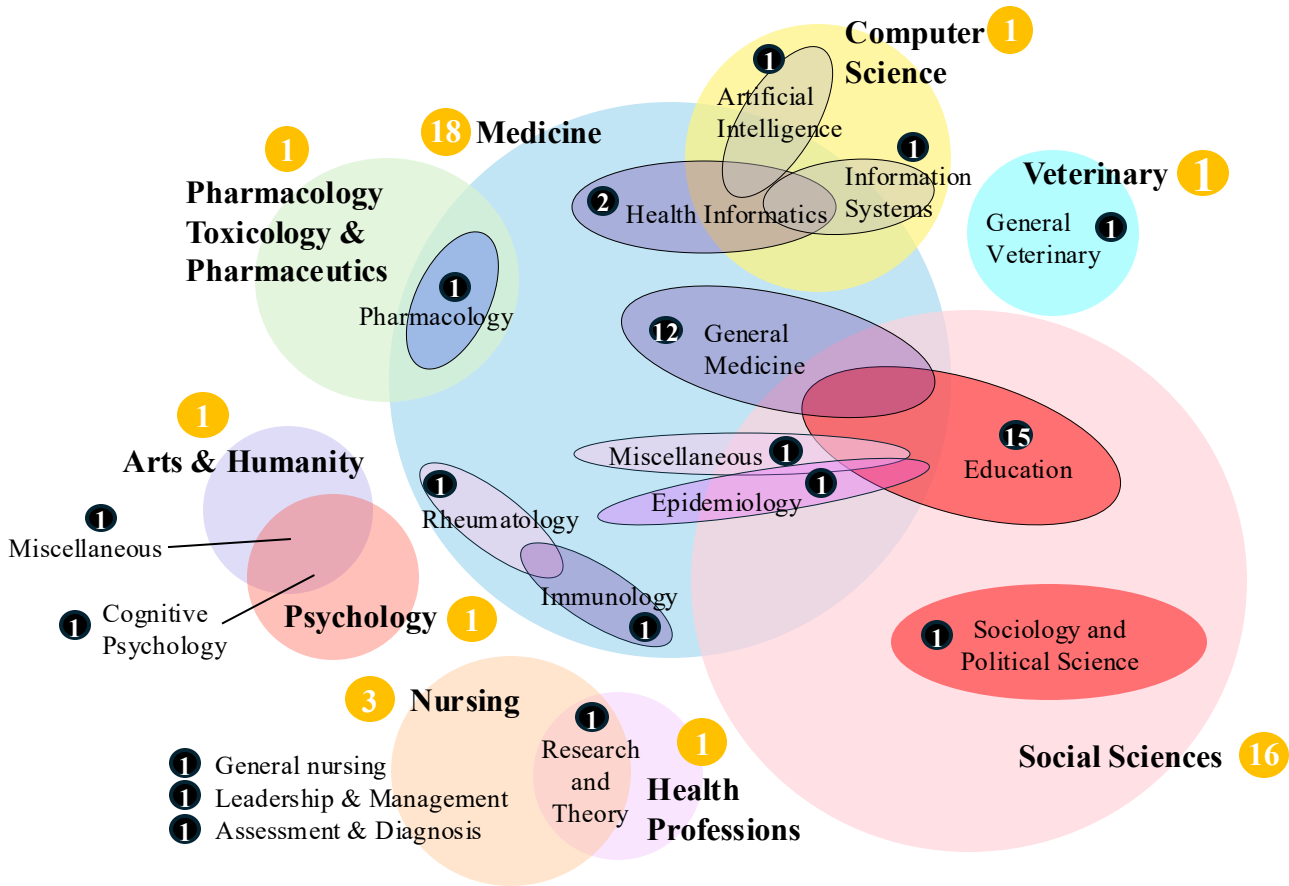(Al-Umran & Adkoli, 2009; Kurtz & Silverman, 1996; Venktaramana et al., 2022; Westmoreland et al., 2019;)

A systematic Review: Research questions

1. What pedagogical practices have been employed in GenAI-supported healthcare communication education ?

2. What are the existing theoretical frameworks and research methodologies in GenAI-supported healthcare communication education research

3. What perceived benefits and challenges have been reported regarding the use of GenAI in healthcare communication education ?

Methodology

Identification	Screening	Screening

Records identified in databases (n=779)

Web of Science (n=142),
Scopus (n=259),
ProQuest (n=180),
ERIC (n=5),
PubMed (n=115),
EMBASE (n=47),
CINAHEL (n=29),
PsycINFO (n=2).

Duplicate records removed (n=311)

Reports sought for retrieval (n=468)

Records excluded (n=391):
- Review papers
- Theoretical papers
- Irrelevant topics

Records full-text screened (n=77)

Records excluded (n=48):
- Not oral communication (n=16)
- Not empirical studies (n=10)
- Not in English (n=1)
- Not peer-reviewed (n=3)
- Not full-text article (n=8)
- No GenAI integration (n=9)
- Duplication (n=1)

Studies included for coding (n=29)

(Page et al., 2021)

Findings overview: Publication distribution

| Publication resource	N
BMC Medical Education	8
JMIR Medical Education	4
Journal of Medical Internet Research	1
Journal of Healthcare Informatics Research	1
Applied Cognitive Psychology	1
Rheumatology International	1
British Journal of Clinical Pharmacology	1
Teaching and Learning in Nursing	1
BMC Nursing	1
Simulation in Healthcare: The Journal of the Society for Simulation in Healthcare	1
Radiography	1
Educational Technology & Society	1
Journal of Education and Health Promotion	1
Frontiers in Veterinary Science	1
Cureus	1
SVR 2024	1
HCII 2024	1
ACL 2024	1
SCIS-ISIS 2024	1

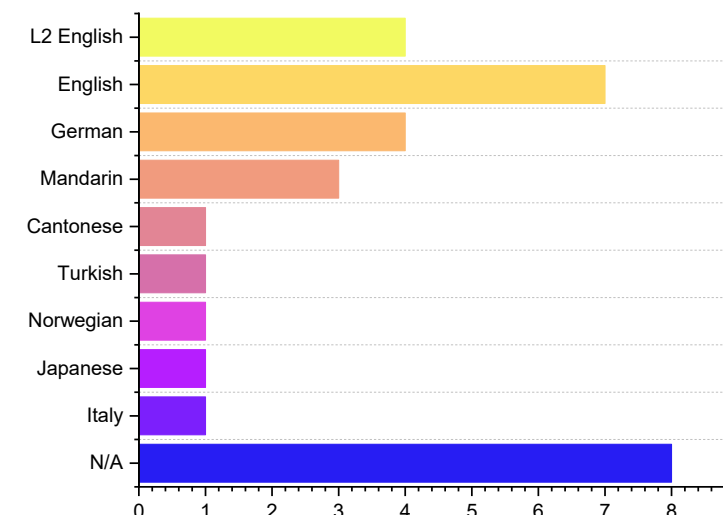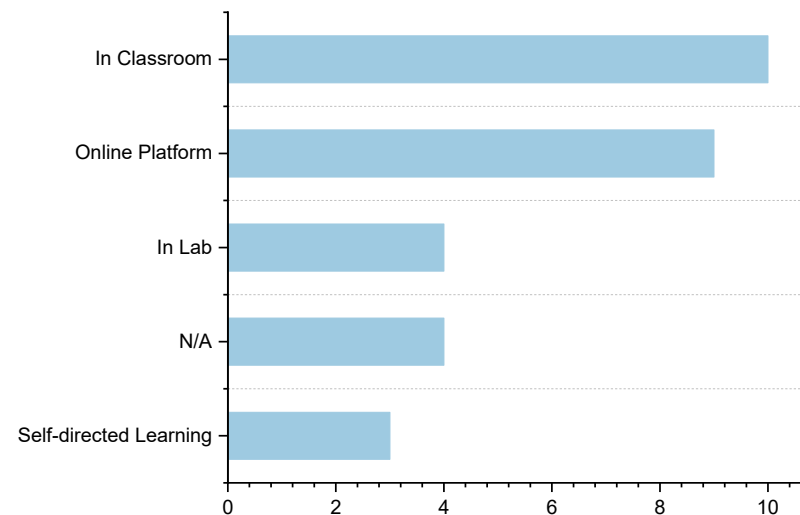# Findings overview: Pedagogical practices

Location
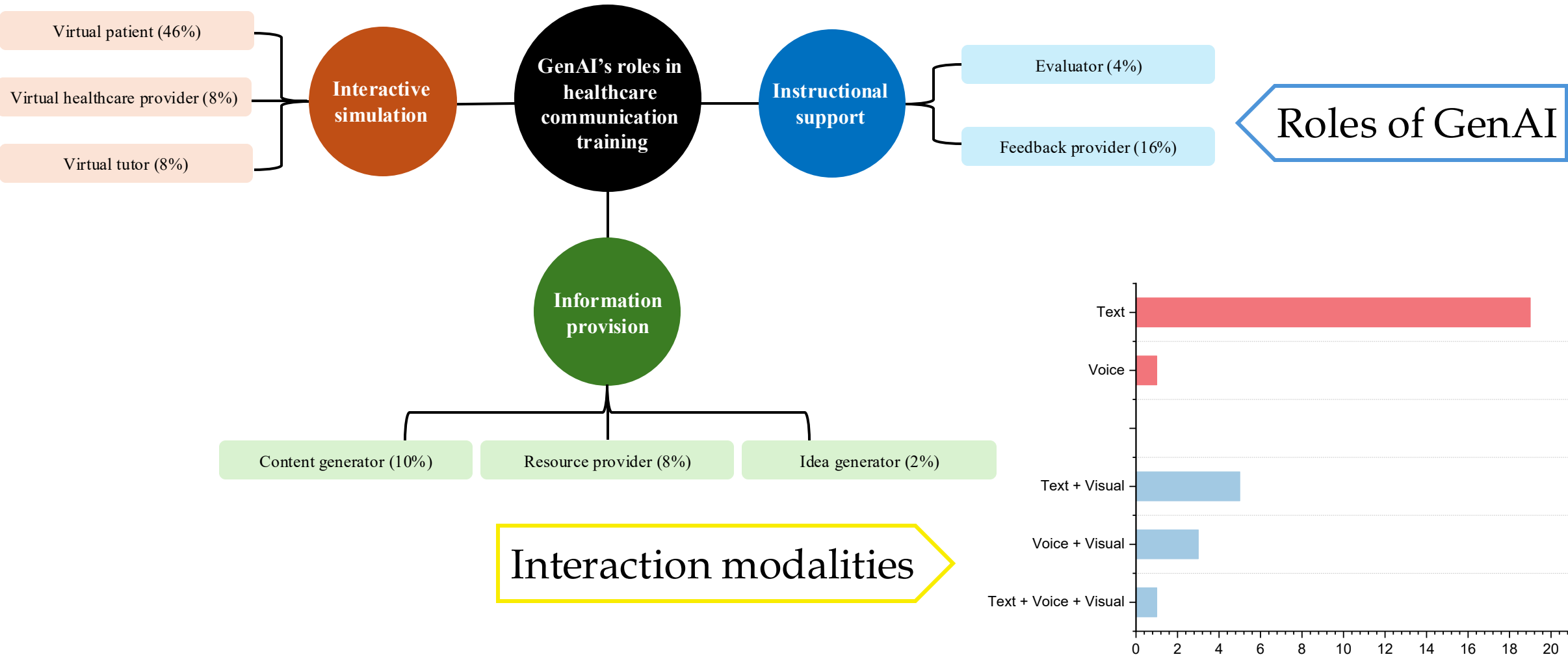
- Geographical distribution
- Interaction settings
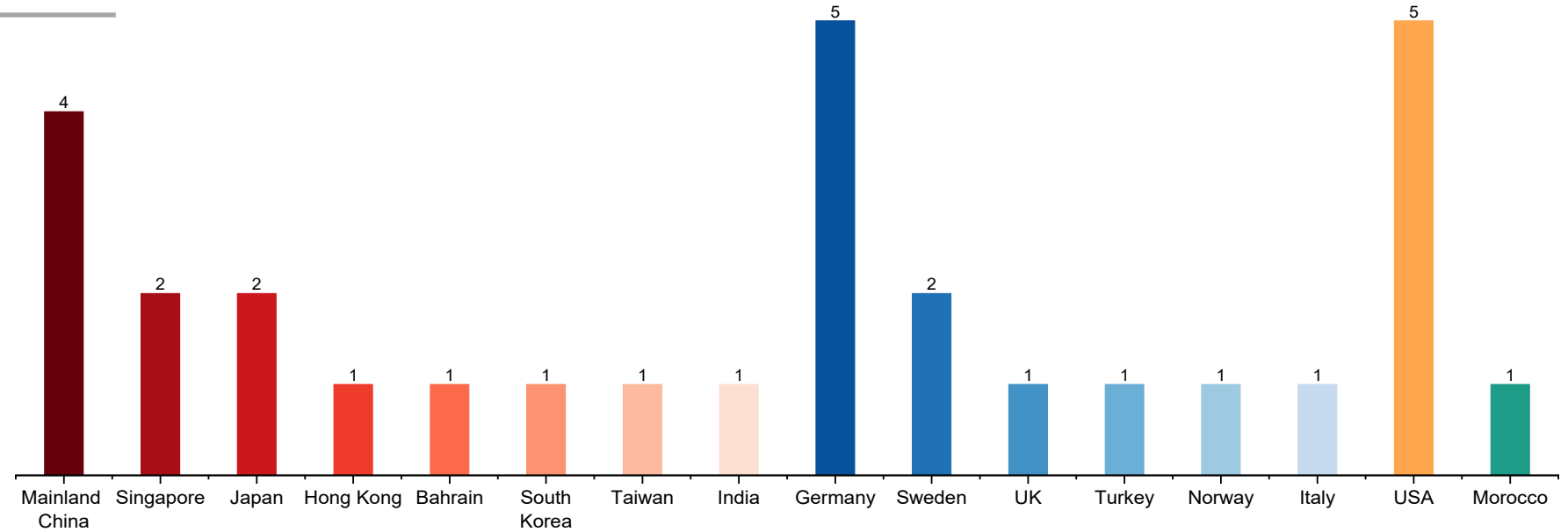
Learner profile

- Undergraduate trainees
- Interaction languages

Findings overview: Pedagogical practices – Roles of GenAI

Findings overview: Research design – Theoretical frameworks

Theoretical framework
• Calgary-Cambridge Guide
• ACGME competencies
• Empathic Communication Coding System
• SPIKES framework
• AETCOM competencies

__GATHERING INFORMATION__

__Exploration of Patient's Problems__

8. **Encourages patient to tell the story** of the problem(s) from when first started to the present in own words (clarifying reason for presenting now)
9. **Uses open and closed questioning techniques**, appropriately moving from open to closed
10. **Listens** attentively, allowing patient to complete statements without interruption and leaving space for patient to think before answering or go on after pausing

... **s** patient's responses verbally and non-verbally (e.g. use of ... ement, silence repletion, paraphrasing, interpretation)

... non-verbal **cues** (body language, speech, facial expression, ...

... **and acknowledges** as appropriate

... that are unclear or need amplification (e.g. "Could ... y light headed")

... erify own understanding of what the patient has said; ... rovide further information

... **and comments**, avoids or adequately

Step 1: Setting up the interview

Step 2: Assess the patient's reception

Step 3: Obtain the patient's invitation

Step 4: Give knowledge and information

Step 5: Address the patient's emotion

Step 6: Summary

Findings overview: Research design – Methodology

Data collection instruments	n (%)	Data analysis methods	n (%)
Questionnaire	18 (35.29%)	t-test	6 (11.11%)
Observer-based rating scale	4 (7.84%)	Chi-square test	4 (7.41%)
Self-report rating scale	3 (5.88%)	Kruskal-Wallis test	4 (7.41%)
Automated (GenAI-based) rating scale	2 (3.92%)	Mann–Whitney U test	4 (7.41%)
Generated scripts	2 (3.92%)	Wilcoxon signed-rank test	2 (3.70%)
Statistical table	1 (1.96%)	Dunn's test	1 (1.85%)
Sub-total	**30 (58.82%)**	Friedman test	1 (1.85%)
Semi-structured interview	4 (7.84%)	Fisher's exact test	2 (3.70%)
Focus group interview	2 (3.92%)	Rank-sum test	1 (1.85%)
Open-ended questions	5 (9.80%)	Correlation analysis	5 (9.26%)
Conversational transcription	5 (9.80%)	Regression analysis	1 (1.85%)
Observer-based checklist	1 (1.96%)	Repeated measures ANOVA	1 (1.85%)
Reflective feedback form	1 (1.96%)	ANOVA	1 (1.85%)
Self-regulated learning sheet	1 (1.96%)	ROUGE-L (sequence-based similarity)	1 (1.85%)
Peer assessment	1 (1.96%)	Lexical recall analysis	1 (1.85%)
Human annotation	1 (1.96%)	BLEU-2 (lexical overlap)	1 (1.85%)
Sub-total	**21 (41.18%)**	BERTScore (semantic similarity)	2 (3.70%)
		Sentiment analysis	1 (1.85%)
		Subtotal	**39 (72.22%)**
		Thematic analysis	9 (16.67%)
		Content analysis	6 (11.11%)
		Subtotal	**15 (27.78%)**
Total	**51 (100%)**		**54 (100%)**

Findings overview: Research design – Methodology

Internal reliability	n (%)	Test-retest reliability	n (%)
Previously reliable instruments	1 (3.33%)	Inter-rater agreement percentage	2 (6.67%)
Cronbach's α	3 (10%)	Cohen's κ	2 (6.67%)
		Correlation coefficient	3 (10%)
		Weighted Cohen's κ	2 (6.67%)
Total	4 (13.33%)		9 (30%)

Types of validity	n (%)	Methods	n (%)
Construct validity	8 (26.67%)	• Previously validated instruments	8 (26.67%)
Content validity	9 (30%)	• Previously validated instruments • Expert review • Pilot study	8 (26.67%) 1 (3.33%) 1 (3.33%)
External validity	4 (13.33%)	• Previously validated instruments	4 (13.33%)
Internal validity	6 (20%)	• Previously validated instruments • Control group	3 (10%) 3 (10%)

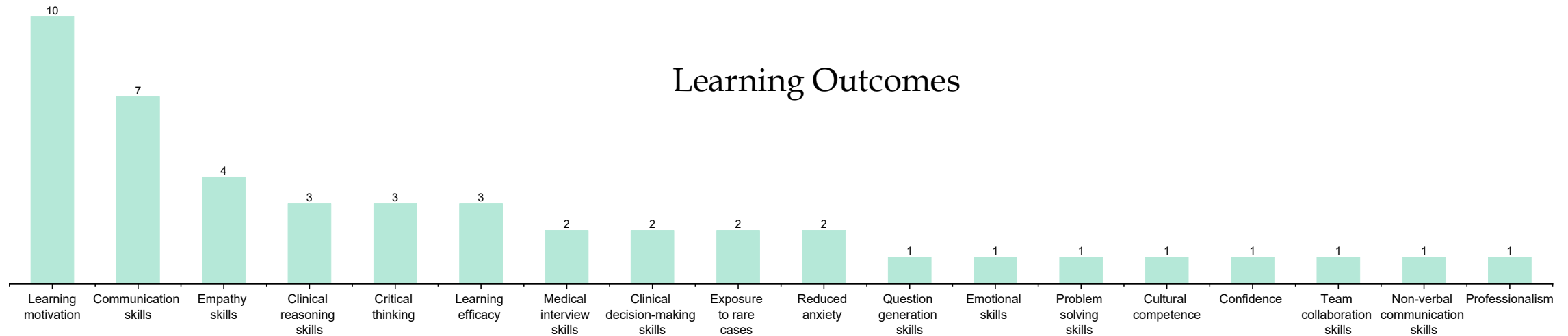# Findings overview: Perceptions – Benefits
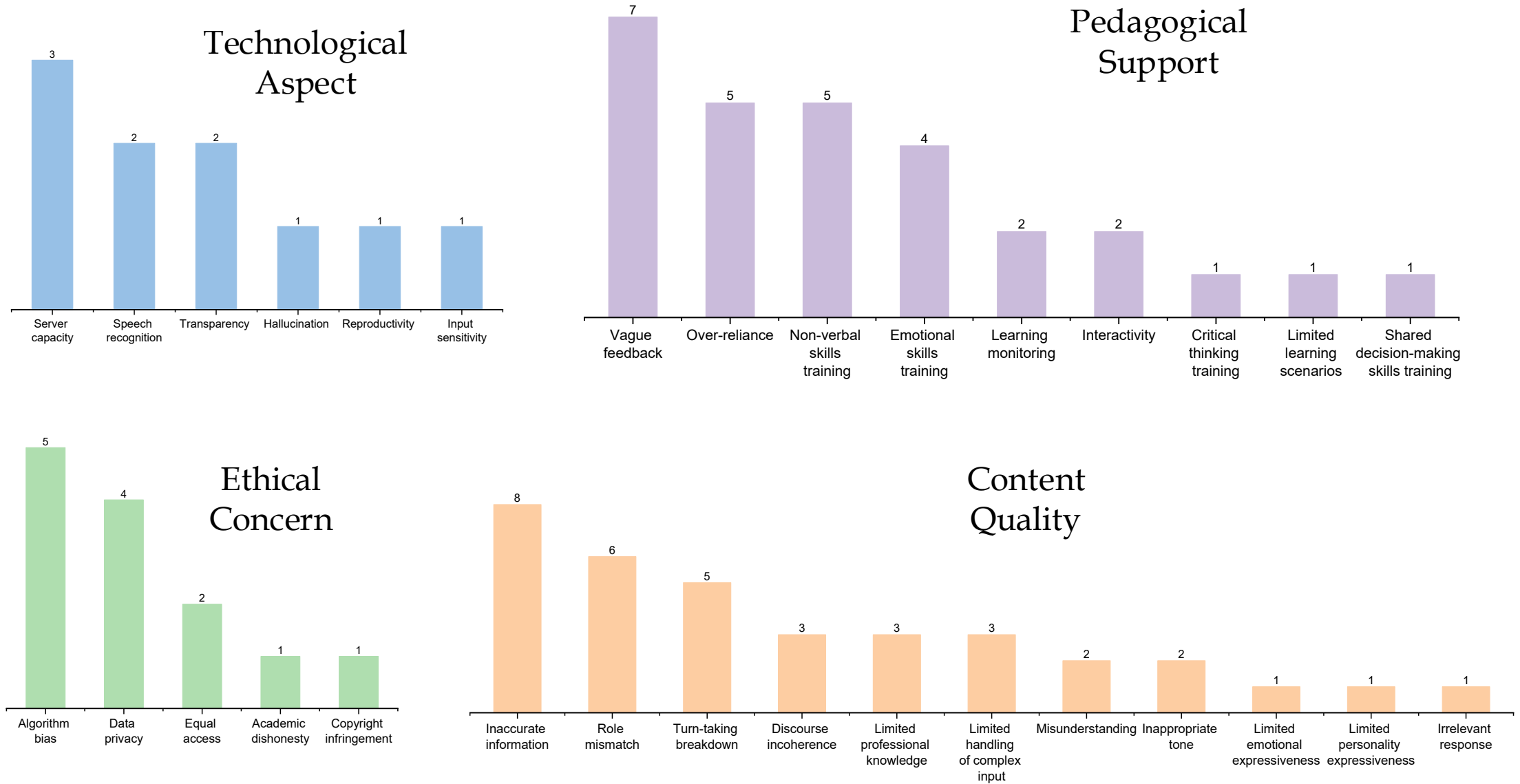

User Experience
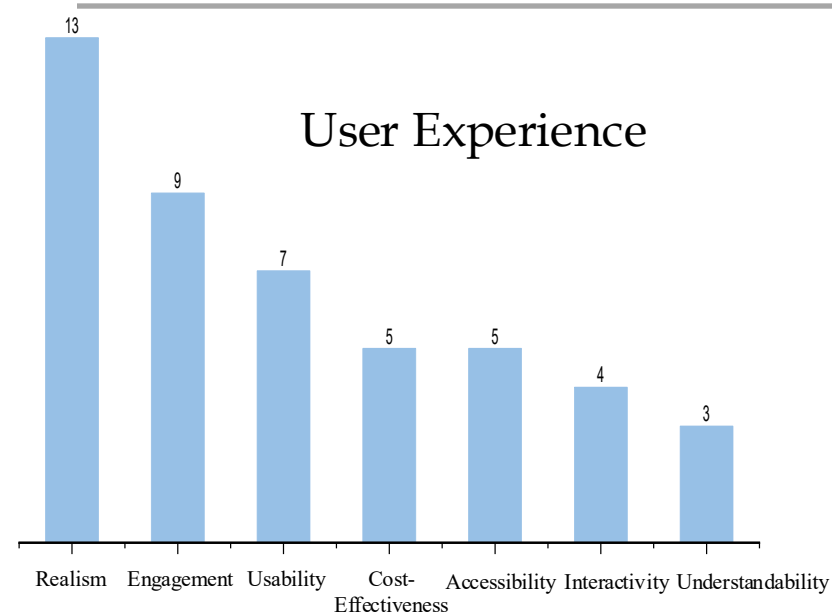
Instructional Support
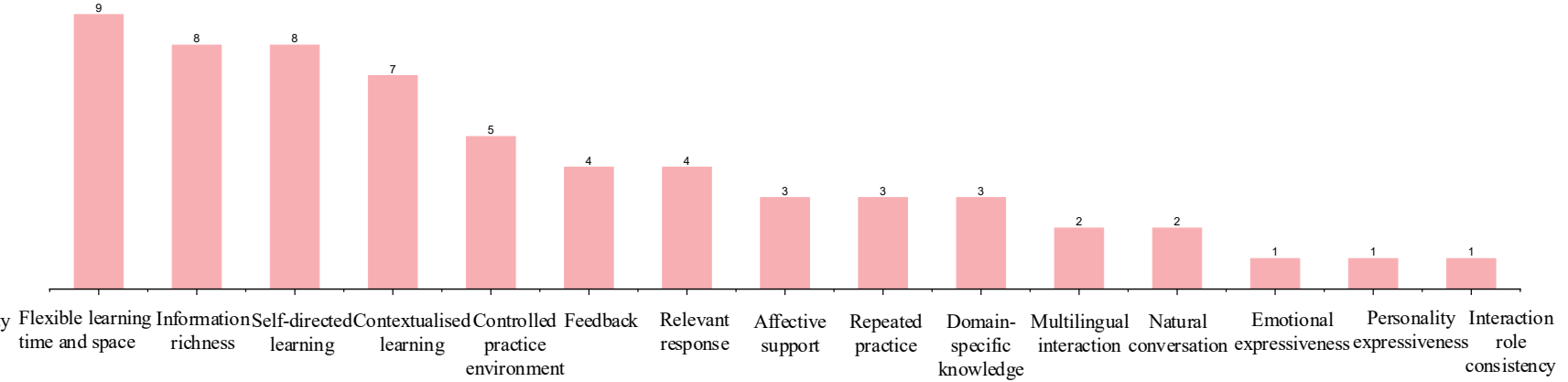
Learning Outcomes

Findings overview: Perceptions – Challenges

A systematic Review: Research gap

Pedagogical practices

- English as the main language
→ Multilingual learning

- Limited patient profiles
→ Self-directed learning

- Text-based interactions
→ Multimodal interactions

Research design

- Conceptual foundations
→ Reframe healthcare communication competence

- Quantitative dominance
→ Qualitative methods

- Limited reliability & validity
→ Suit GenAI contexts

Stakeholder perceptions

- Content quality
e.g., turn-taking breakdown, role mismatch, discourse incoherence

- Ethical concerns
e.g., algorithm bias

- Unexplored long-term influence

Towards an applied linguistics research agenda

- Reframe healthcare communication competence

→social, situated, interactional

- Employ qualitative research methods

→Conversation Analysis, Membership Categorisation Analysis

- Call for long-term research
- Self-directed learning setting
- Multilingual, multimodal learning

Studies included for review

Borg et al.	2025	Virtual patient simulations using social robotics combined with large language models for clinical reasoning training in medical education: Mixed methods study.
Zeng et al.	2025	Application of ChatGPT-assisted problem-based learning teaching method in clinical medical education
Rädel-Ablass et al.	2025	Teaching opportunities for anamnesis interviews through AI based teaching role plays: a survey with online learning students from health study programs
Pears et al.	2025	Non-technical skills for urology trainees: A double-blinded study of ChatGPT4 AI benchmarking against consultant interaction
Öncü et al.	2025	AI-powered standardised patients: evaluating ChatGPT-4o's impact on clinical case management in intern physicians
Chan et al.	2025	Using ChatGPT for medical education: the technical perspective
Ko et al.	2025	Using an AI-Driven child chatbot avatar as a training tool for information gathering skills of dental and medical professionals: A pilot study
Borg et al.	2024	Enhancing clinical reasoning skills for medical students: a qualitative comparison of LLM-powered social robotic versus computer-based virtual patients within rheumatology
Yamamoto et al.	2024	Enhancing medical interview skills through AI-simulated patient interactions: Nonrandomized controlled trial
Holderried et al.	2024	A language model–powered simulated patient with automated feedback for history taking: Prospective study
Sridharan & Sequeira	2024	Evaluation of artificial intelligence-generated drug therapy communication skill competencies in medical education
Benfatah et al.	2024	Assessing the efficacy of ChatGPT as a virtual patient in nursing simulation training: A study on nursing students' experience
Lee et al.	2024	Collaborative project-based learning in global health: Enhancing competencies and skills for undergraduate nursing students
Xu et al.	2024	Designing and evaluating an emotionally responsive virtual patient simulation
Ba et al.	2024	Enhancing clinical skills in pediatric trainees: A comparative study of ChatGPT-assisted and traditional teaching methods
Bonfitto et al.	2024	Harnessing ChatGPT dialogues to address claustrophobia in MRI - A radiographers' education perspective
Huang & Lin	2024	ChatGPT as a life coach for professional identity formation in medical education
Roy et al.	2024	Efficacy of ChatGPT in solving attitude, ethics, and communication case scenario used for competency-based medical education in India: A case study
Yao et al.	2024	Enhancing empathic communication in healthcare education through virtual conversations: Leveraging large language models for real-time feedback
Kapadia et al.	2024	Evaluation of large language model generated dialogues for an AI based VR nurse training simulator
Huang et al.	2024	Implementation and evaluation of an optimized surgical clerkship teaching model utilizing ChatGPT
Huang et al.	2024	Benchmarking large language models on communicative medical coaching: A dataset and a novel system
Holderried et al.	2024	A generative pretrained transformer (GPT)–powered chatbot as a simulated patient to practice history taking: Prospective, mixed methods study
Gray et al.	2024	Increasing realism and variety of virtual patient dialogues for prenatal counselling education through a novel application of ChatGPT: exploratory observational study
Brügge et al.	2024	Large language models improve clinical decision making of medical students through patient simulation and structured feedback: A randomized controlled trial
Aster et al.	2024	Development and evaluation of an emergency department serious game for undergraduate medical students
Artemiou et al.	2024	Introducing AI-generated cases (AI-cases) & standardized clients (AI-SCs) in communication training for veterinary students: perceptions and adoption challenges
Agatsuma et al.	2024	Building a role-play interactive system using LLM for health guidance education
Webb	2023	Proof of concept: Using ChatGPT to teach emergency physicians how to break bad news

References

- Al-Umran, K. U., & Adkoli, B. V. (2009). Experience of a workshop on communication skills in health professional education. *Journal of Family and Community Medicine*, 16(3), 115.
- Choudhary, A., & Gupta, V. (2015). Teaching communications skills to medical students: Introducing the fine art of medical practice. *International Journal of Applied and Basic Medical Research*, 5(1), S41. https://doi.org/10.4103/2229-516X.162273
- D'Agostino, T. A., Atkinson, T. M., Latella, L. E., Rogers, M., Morrissey, D., DeRosa, A. P., & Parker, P. A. (2017). Promoting patient participation in healthcare interactions through communication skills training: A systematic review. *Patient Education and Counselling*, 100(7), 1247–1257. https://doi.org/10.1016/j.pec.2017.02.016
- Dai, D. W. (2024). Interactional Competence for professional communication in intercultural contexts: Epistemology, analytic framework and pedagogy. *Language, Culture and Curriculum*, 37(4), 435–455. https://doi.org/10.1080/07908318.2024.2349781
- Fleckman, J. M., Dal Corso, M., Ramirez, S., Begalieva, M., & Johnson, C. C. (2015). Intercultural Competency in Public Health: A Call for Action to Incorporate Training into Public Health Education. *Frontiers in Public Health*, 3. https://doi.org/10.3389/fpubh.2015.00210
- Kurtz, S. M., & Silverman, J. D. (1996). The Calgary—Cambridge Referenced Observation Guides: An aid to defining the curriculum and organizing the teaching in communication training programmes. *Medical Education*, 30(2), 83–89. https://doi.org/10.1111/j.1365-2923.1996.tb00724.x
- Kurtz, S., Draper, J., & Silverman, J. (2017). *Teaching and Learning Communication Skills in Medicine* (2nd edn). CRC Press. https://doi.org/10.1201/9781315378398
- Sandman, L., & Munthe, C. (2010). Shared decision making, paternalism and patient choice. *Health Care Analysis*, 18(1), 60–84. https://doi.org/10.1007/s10728-008-0108-6
- Taylor, K. (2009). Paternalism, participation and partnership—The evolution of patient centeredness in the consultation. *Patient Education and Counselling*, 74(2), 150–155. https://doi.org/10.1016/j.pec.2008.08.017
- Teutsch, C. (2003). Patient–doctor communication. *Medical Clinics*, 87(5), 1115–1145. https://doi.org/10.1016/S0025-7125(03)00066-X
- Venktaramana, V., … Krishna, L. K. R. (2022). A systematic scoping review of communication skills training in medical schools between 2000 and 2020. *Medical Teacher*, 44(9), 997–1006. https://doi.org/10.1080/0142159X.2022.2054693
- Westmoreland, K. D., Isaksson, E., & Fassl, B. A. (2019). A standardized low-cost peer role-playing training intervention improves medical student competency in communicating bad news to patients in Botswana. *Palliative & Supportive Care*, 17(1), 60–65. https://doi.org/10.1017/S1478951518000627

Thank you!

Zilin Jia

zilin. jia.21@ucl.ac.uk