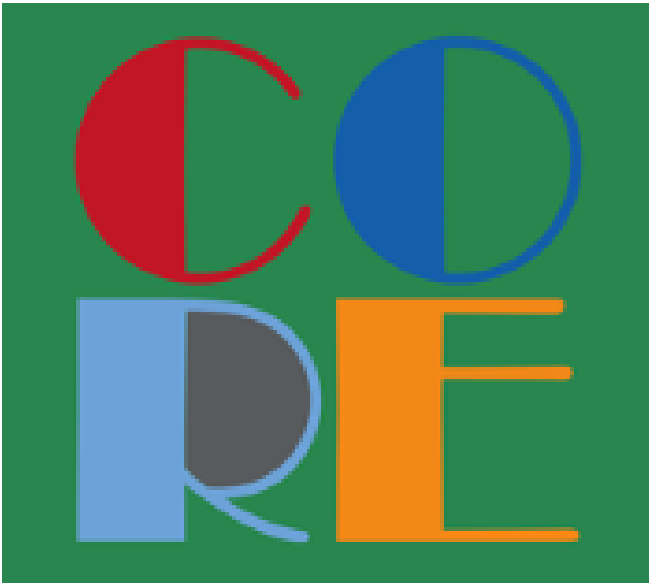


Critical Online Reasoning im Studium der Wirtschaftswissenschaften und Wirtschaftspädagogik: Eine Längsschnittstudie zur Entwicklung und KI-Nutzung



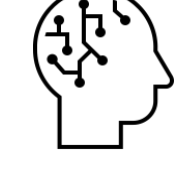
RELEVANZ & FORSCHUNGSFOKUS



Critical Online Reasoning (COR): Fähigkeit zur systematischen Recherche, kritischen Bewertung und evidenzbasierten Argumentation von Online-Informationen^(1; 2; 3)



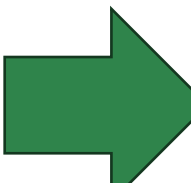
COR gilt als zentrale Kompetenz zur Bewältigung **akademischer und beruflicher Herausforderungen**^(2; 3)



Einsatz von generativen **KI-Tools** wie ChatGPT und Google Gemini verändert Recherche-, Lern- & Wissenserwerbsprozesse im Studium: Interaktion mit dynamisch, algorithmisch generierten Inhalten^(4; 5; 6)



WiWi-/WiPäd-Studierende nutzen zunehmend Online- und generative KI-Ressourcen für akademische Aufgaben^(6; 7), längsschnittliche Befunde zur Kompetenzentwicklung von COR-Fähigkeiten unter generativer KI sind jedoch bislang begrenzt^(7; 8)



Forschungsdesiderat: Wie entwickeln sich die COR-Fähigkeiten und die KI-Nutzung von WiWi/WiPäd-Studierenden im Bachelorstudium?

COR



COR-Facetten⁽²⁾

OIA: Online-Informationsbeschaffung

CIE: Kritische Informationsbewertung

REAS: Schlussfolgern mit Belegen, Argumenten und Synthese

Reguliert durch MCA: Metakognitive Aktivierung und Reflexion

Erfassung an 4 Messzeitpunkten:

t0: Wintersemester 2023/24

t1: Wintersemester 2024/25

t2: Wintersemester 2025/26

KI-Studie: Sommersemester 2025

ASSESSMENT

Virtuelle Testumgebung

Studierende bearbeiten domänenspezifische (DOM) Problemstellungen mit realer Websuche

Problemstellung szenariobasierte Aufgabe

↓

Online-Recherche Webseiten · Suchmaschinen · KI

↓

Produkt Textantwort · Evidenz · Begründung

+ Echtzeit-Aufzeichnung der Recherche-, Evaluations- und Argumentationsprozesse während der Aufgabenbearbeitung

AUFGABENDESIGN

DOM-COR-Aufgaben

- 7 wirtschaftsbezogene Szenarien:** z.B. WiPäd-Studierende beurteilen das Spannungsfeld zwischen Wirtschaftswachstum und ökologischer Nachhaltigkeit.
- in drei **Anwendungskontexten:** fundamental, practical, transdisciplinary
- mit **variierender kognitiver Schwierigkeit** auf drei Kompetenzstufen: basic, advanced, proficient

Integration von KI-Tools im Assessment

- 4 adaptierte Szenarien zur Erfassung und Kontrolle von KI-Nutzung während der Aufgabenbearbeitung über Logdaten, Bildschirmaufzeichnungen und Selbstreflektion
- Evaluation umfasst: Qualität der Prompts, Triangulation von KI-Outputs, kritische Bewertung KI-generierter Inhalte

CIE

Auswertung: schriftliche Begründung der Quellenauswahl und -bewertung

- Kodierung durch 2 geschulte Rater:innen anhand 20 binärer Indikatoren⁽⁹⁾ mit mäßig bis moderater IRR ($0.55 < K_{\alpha} < 0.78$)

Message Credibility: Qualität des Inhalts (Konsistenz, Argumentation)

Source Credibility: Glaubwürdigkeit der Quelle (bspw. Expertise)

Surface Credibility: Oberflächenmerkmale

Presumed Credibility: Vertrauen aufgrund von Labels, Ranking, Gewohnheit

REAS

Auswertung: schriftliche Stellungnahme (ca. 200 Wörter)

- Bewertung durch 3 geschulte Rater:innen anhand 8 ordinalen Kriterien mit moderater IRR ($0.70 < K_{\alpha} < 0.74$)
- Kontextunabhängige Kriterien: Problembezug, Kohärenz, Schlussfolgerung, Quellenbelege, Argumente
- Kontextspezifische Kriterien: Argumentation, Ausgewogenheit, Berücksichtigung verschiedener Perspektiven

PANELSTICHPROBE und KI-NUTZUNG

	N	Alter MW (SD)	Weiblich N (%)	HZB-Note MW (SD)	Semester MW (SD)
t0 – t1	56	19.73 (2.03)	29 (51.8)	1.88 (0.60)	1.00 (0.00)
t0 – t2	46	19.96 (2.43)	19 (41.3)	1.85 (0.54)	1.00 (0.00)
t1 – t2	66	20.36 (2.45)	31 (47.0)	1.89 (0.58)	2.03 (1.54)

Anmerkung. Personen mit allen drei Messzeitpunkten erscheinen in allen drei Gruppen.

	N	KI zum Lernen	KI-Fähigkeiten	KI-Relevanz
t0 – t1	56	2.15 (1.38) → 3.85 (1.48) ***	2.49 (1.10) → 2.15 (1.17) *	2.19 (1.08) → 2.04 (0.94)
t0 – t2	46	2.15 (1.30) → 5.00 (1.30) ***	2.69 (1.08) → 2.00 (0.98) ***	2.47 (1.22) → 1.91 (0.76) *
t1 – t2	66	3.83 (1.66) → 4.91 (1.27) ***	2.32 (1.10) → 2.00 (0.98) *	1.98 (0.87) → 1.98 (0.85)

Anmerkung. MW (SD). Personen mit allen drei Messzeitpunkten erscheinen in allen drei Gruppen. Skala zum Lernen: 1=nie bis 6=mehrmals täglich, Skala für Kompetenz und Relevanz (Selbsteinschätzung): 5=sehr schlecht bis 1=sehr gut. * p<.05, ** p<.01, *** p<.001, gepaarte t-Tests

ERGEBNISSE

Paar	1. MZP MW (SD)	2. MZP MW (SD)	p-Wert
REAS ^a			
t0 – t1	0.52 (0.17)	0.49 (0.14)	0.208
t0 – t2	0.52 (0.16)	0.58 (0.14)	0.023
t1 – t2	0.50 (0.15)	0.55 (0.13)	0.014
CIE: Message Credibility ^b			
t0 – t1	0.32 (0.34)	0.46 (0.45)	0.042
t0 – t2	0.34 (0.36)	0.47 (0.45)	0.088
t1 – t2	0.45 (0.40)	0.42 (0.42)	0.654
CIE: Source Credibility ^b			
t0 – t1	0.37 (0.40)	0.49 (0.52)	0.120
t0 – t2	0.37 (0.37)	0.35 (0.40)	0.799
t1 – t2	0.39 (0.39)	0.37 (0.43)	0.708
CIE: Surface Credibility ^b			
t0 – t1	1.35 (0.38)	1.36 (0.51)	0.953
t0 – t2	1.40 (0.37)	1.46 (0.37)	0.391
t1 – t2	1.33 (0.46)	1.49 (0.40)	0.019
CIE: Presumed Credibility ^b			
t0 – t1	1.53 (0.41)	1.45 (0.46)	0.290
t0 – t2	1.49 (0.42)	1.80 (0.33)	< 0.001
t1 – t2	1.53 (0.43)	1.79 (0.32)	< 0.001

Anmerkung. ^a Wilcoxon-Vorzeichen-Rang-Test, Skala: 0=sehr schlecht bis 1=sehr gut. ^b gepaarte t-Tests, Skala: 0=sehr schlecht bis 2=sehr gut

DISKUSSION

- ✓ Nutzung von KI zum Lernen verzeichnet unter den Studierenden der Wirtschaftswissenschaften und Wirtschaftspädagogik in allen 3 Kohorten einen signifikanten Anstieg; ebenso bei der Einschätzung der eigenen Fähigkeiten im Umgang mit KI
- ✓ Nur geringe Kompetenzentwicklung bei der REAS-Facette im Bachelorstudium
- ✓ Defizite bei der kritischen Quellenbewertung: Oberflächliche Kriterien wie Surface Credibility und Presumed Credibility verbessern sich leicht, anspruchsvollere, inhaltliche Kriterien wie Source Credibility entwickeln sich im Bachelorstudium nicht weiter
- Analysen zum Zusammenhang zwischen der KI-Nutzung im Studium und der Entwicklung von COR-Kompetenzen erforderlich⁽¹⁰⁾
- Ergebnisse erlauben erste Implikationen für gezielte Förderung von COR-Kompetenzen und KI-Nutzungspraktiken in der WiWi/WiPäd-Hochschulbildung:

Stärkung der traditionellen Fähigkeiten zur Informationsbewertung und zum kritischen Denken in Online-Lernumgebungen mit gleichzeitiger Förderung kritischer KI-Kompetenzen, die es Studierenden ermöglichen, KI-generierte Informationen zu interpretieren, zu hinterfragen und strategisch zu nutzen

PROJEKT

DFG FOR CORE (5404)

www.core.uni-mainz.de

QUELLEN

(1) Osborne, J., Pimentel, D., Alberts, B., Allchin, D., Barzilai, S., Bergstrom, C., Coffey, J., Donovan, B., Kivinen, K., Kozyreva, A., & Wineburg, S. (2022). *Science Education in an Age of Misinformation*. Report. Stanford. <https://sciedandmisinfo.stanford.edu/>

(2) Molerov, D., Zlatkin-Troitschanskaia, O., Nagel, M.-T., Brückner, S., Schmidt, S., & Shavelson, R. (2020). Assessing university students' critical online reasoning ability: A conceptual and assessment framework with preliminary evidence. *Frontiers in Education*, 5. <https://doi.org/10.3389/feduc.2020.577843>

(3) Zlatkin-Troitschanskaia, O., Beck, K., Fischer, J., Braunheim, D., Schmidt, S., & Shavelson, R. J. (2020). The Role of Students' Beliefs When Critically Reasoning From Multiple Contradictory Sources of Information in Performance Assessments. *Frontiers in Psychology*, 11. <https://doi.org/10.3389/fpsyg.2020.02192>

(4) Al-Zahrani, A.M. & Alasmari, T.M. (2024). Exploring the impact of artificial intelligence on higher education: The dynamics of ethical, social, and educational implications. *Humanit Soc Sci Commun* 11, 912 (2024). <https://doi.org/10.1057/s41599-024-03432-4>

(5) Labadze, L., Grigolia, M., & Machaidze, L. (2023). Role of AI chatbots in education: systematic literature review. *International Journal of Educational Technology in Higher Education*, 20(1), 56. <https://doi.org/10.1186/s41239-023-00426-1>

(6) Braunheim, D. (2024). Effects of Regular Media Use of Young Professionals on Their Online Source Selection During a Situated Digital Performance Assessment. In: Zlatkin-Troitschanskaia, O., Nagel, M.T., Klose, V., Mehler, A. (eds) *Students', Graduates' and Young Professionals' Critical Use of Online Information*. Springer. https://doi.org/10.1007/978-3-031-69510-0_3

(7) Federiakin, D., Molerov, D., Zlatkin-Troitschanskaia, O., & Maur, A. (2024). Prompt engineering as a new 21st century skill. *Frontiers in Education*, 9. <https://doi.org/10.3389/feduc.2024.1366434>

(8) Zlatkin-Troitschanskaia, O., Hartig, J., Goldhammer, F., & Krstev, J. (2021). Students' online information use and learning progress in higher education – A critical literature review. *Studies in Higher Education*. <https://doi.org/10.1080/03075079.2021.1953336>

(9) Wathen, C. N., & Burkell, J. (2002). Believe it or not: Factors influencing credibility on the Web. *Journal of the American Society for Information Science and Technology*, 53(2), 134–144. <https://doi.org/10.1002/asi.10016>

(10) Molerov, D., Federiakin, D., Zlatkin-Troitschanskaia, O., Shenavai, K., Trierweiler, L. & Nagel, M. (2026). The relationship between AI-chatbots use, student assessment performance and learning outcomes in higher education. *Unterrichtswissenschaft*, 1-32. <https://doi.org/10.1007/s42010-026-00242-2>