Computers and Education: Artificial Intelligence 8 (2025) 100340

Contents lists available at ScienceDirect

Computers and Education: Artificial Intelligence

journal homepage: www.sciencedirect.com/journal/computers-and-education-artificial-intelligence

AI advocates and cautious critics: How AI attitudes, AI interest, use of AI, and AI literacy build university students' AI self-efficacy

Arne Bewersdorff[a,*], Marie Hornberger[a], Claudia Nerdel[a], Daniel S. Schiff[b]

[a] *Technical University of Munich, Germany*
[b] *Purdue University, USA*

ARTICLE INFO

Keywords:
Artificial intelligence
AI literacy
AI self-efficacy
AI education
Higher education
Structural equation model
Gaussian mixture model

ABSTRACT

This study investigates how cognitive, affective, and behavioral variables related to artificial intelligence (AI) build AI self-efficacy among university students. Based on these variables, we identify three meaningful student groups, which can guide educational initiatives. We recruited 1465 undergraduate and graduate students from the United States, the United Kingdom, and Germany and measured their AI self-efficacy, AI literacy, interest in AI, attitudes towards AI, and AI use. Using a path model, we examine the correlations and paths among these variables. Results reveal that AI usage and positive AI attitudes significantly predict interest in AI, which in turn and together with AI literacy, enhance AI self-efficacy. Moreover, using Gaussian Mixture Models, we identify three groups of students: 'AI Advocates,' 'Cautious Critics,' and 'Pragmatic Observers,' each exhibiting unique patterns of AI-related cognitive, affective, and behavioral traits. Our findings demonstrate the necessity of educational strategies that not only focus on AI literacy but also aim to foster students' AI attitudes, usage, and interest to effectively promote AI self-efficacy. Furthermore, we argue that educators who aim to design inclusive AI educational programs should take into account the distinct needs of different student groups identified in this study.

1. Introduction

As artificial intelligence (AI) rapidly permeates our daily lives and workplaces, university students must develop *AI self-efficacy* – the confidence in their capabilities to effectively interact with, understand, learn about, and use (when desired) AI technologies and applications. Self-efficacy, in general, is central for students as it boosts their confidence in their ability to successfully accomplish specific tasks (Ayllón et al., 2019), influencing their aspirations, persistence, and achievements (Schunk & DiBenedetto, 2021). Domain-specific self-efficacy likewise fosters engagement and aspirations within that domain (regarding science, see: Lau & Roeser, 2002; regarding AI, see: Chen et al., 2024), ultimately empowering students to navigate and participate in an AI-driven world.

However, self-efficacy does not develop in isolation; it is rather influenced by variables like literacy, interest, attitudes, and use, as described in many psychological theories across different educational levels and throughout the learning process (e.g., see Eccles & Wigfield, 2002). Studies have shown, for instance, the influence of ability (Gainor & Lent, 1998), anxiety, and prior experience (Johnson, 2005) on self-efficacy. Research has further revealed that the effects of these variables and their specific dynamics are often domain-specific (Pajares, 1997; Usher & Pajares, 2008).

For AI educators, it is thus crucial to understand the interplay between cognitive (AI literacy), affective (positive and negative attitudes towards AI, AI interest), and behavioral (use of AI) variables and how they affect AI self-efficacy. While recent studies have begun to reveal *correlations* between several of these key constructs (Hornberger et al., 2023; Carolus et al., 2023), there remains a lack of clarity on which variables most significantly *shape* AI self-efficacy. Researchers are also working to understand causal effects (Markus et al., 2024); however, to date, we lack a comprehensive understanding of the specific variables that influence self-efficacy in the domain of AI.

To address this gap, our study aims to derive and validate a path model by drawing from existing general educational theories and extending them to understand AI self-efficacy. These unveiled paths between AI literacy, AI interest, attitudes towards AI, and use of AI, ultimately fostering AI self-efficacy, can lead to the design of more effective – and inclusive – AI educational programs.

In line with the concerns raised above, this study therefore aims to

* Corresponding author. Arcisstraße 31, 80333, Munich, Germany.
E-mail address: arne.bewersdorff@tum.de (A. Bewersdorff).

https://doi.org/10.1016/j.caeai.2024.100340
Received 18 October 2024; Received in revised form 5 December 2024; Accepted 5 December 2024
Available online 6 December 2024

answer the following descriptive and exploratory research questions.

- RQ 1: How do key dimensions of AI education among students – AI self-efficacy, AI literacy, AI interest, attitudes towards AI, and AI use – correlate with one another?
- RQ 2: How do AI literacy, AI interest, attitudes towards AI, and use of AI influence AI self-efficacy?

Regarding goals of inclusive education, stakeholders have argued that AI self-efficacy is essential for *all* students, e.g., regardless of discipline or gender to help students effectively integrate AI into their future professional life and to navigate or even challenge the impact of AI on their lives as citizens. As such, inclusive education must ensure that AI learning programs offer equitable opportunities for all students (Ng et al., 2021), sensitive to their individual challenges, attitudes, and needs. Because cognitive (AI literacy), affective (attitudes towards AI, AI interest), and behavioral (use of AI) variables might collectively influence students' relationship with AI and ultimately contribute to AI self-efficacy, we additionally investigate how AI self-efficacy and these variables surrounding AI self-efficacy vary across different groups of students. To better characterize potential groups, we examine demographic and experiential variation, focusing on gender, discipline, and access to AI education programs at students' home institutions. We aim to answer the following research questions.

- RQ 3: What student groups can be identified based on their AI self-efficacy, AI literacy, AI interest, AI attitudes, and use of AI?
- RQ 4: How do factors like gender, discipline, and perceived access to AI courses differ among these student groups, and what are the implications for educational practice?

2. Conceptual foundation

2.1. Cognitive, behavioral, and affective variables related to AI self-efficacy

Understanding the factors that influence students' AI self-efficacy is essential for designing effective AI programs. However, self-efficacy does not develop in isolation. The literature in educational psychology suggests that self-efficacy is influenced by cognitive, affective, and behavioral variables (Eccles & Wigfield, 2002; Pajares, 1997). To understand and find effective ways to boost AI self-efficacy among students, it is crucial to examine variables from each dimension. Table 1 provides definitions of key constructs related to variables in each dimension.

Table 1
Definitions of key constructs related to AI self-efficacy, AI literacy, AI interest, AI attitudes, and AI use.

Key Construct	Dimension	Definition
AI self-efficacy	Cognitive-Affective	Self-efficacy is the belief in one's ability to successfully perform tasks and achieve goals in a specific domain (Bandura, 1994). Wang and Chuang (2024) define AI self-efficacy as individuals' general belief in their ability to use and interact with AI. In the context of education we define AI self-efficacy as students' confidence in their capabilities to use and interact but also to understand and learn about AI technologies and applications.
AI literacy	Cognitive	Literacy is the ability to understand and interpret information in a particular domain. AI literacy can be defined as "a set of competencies that enables individuals to critically evaluate AI technologies; communicate and collaborate effectively with AI; and use AI as a tool online, at home, and in the workplace" (Long & Magerko, 2020).
Interest in AI	Affective	Interest is both a psychological state of attention and affect towards a particular topic and an enduring predisposition to reengage over time (Harackiewicz et al., 2016). In the context of AI, interest refers to students' sustained affect towards and engagement with AI systems and more general AI-related topics.
Attitudes towards AI	Affective	Attitudes are evaluative judgments towards a particular object, person, or concept (Vogel & Wanke, 2016). In the context of AI, attitudes refer to students' positive or negative evaluations of AI technologies and their applications.
Use of AI	Behavioral	We understand use as the engagement with an object, tool, or substance. In the context of our study, the use of AI refers to how often students interact with AI systems for learning as well as in their daily lives.

2.2. Cognitive, affective, and behavioral variables affecting AI self-efficacy

The interplay of variables like attitudes, use, interest, literacy, and self-efficacy is described by many psychological theories on different levels of the learner and across the learning process (e.g., see Eccles & Wigfield, 2002). Many theories focus on some of the mentioned variables and indicate that these variables do not operate in isolation but form a dynamic system.

The *Interest Development Model* (Hidi and Renninger, 2006) provides a framework for understanding how interest develops and evolves. In turn, *Social Cognitive Theory* (Bandura, 2004) highlights self-efficacy as influenced by mastery experiences. In the following section, we will describe how initial AI use (a mastery experience) and a positive attitude might develop into sustained individual interest. This sustained individual interest eventually leads to higher self-efficacy.

According to the Interest Development Model, interest is triggered by the initial attention facilitated by novel or appealing stimuli. In the field of AI, these stimuli could be the students' first encounters with AI systems like ChatGPT. Indeed, only after the release of ChatGPT in November 2022 did we witness the highest user growth rates ever measured in the tech industry (Reuters, 2023), indicating huge triggered situational interest in this AI system. Building on this first situational trigger, personal relevance of the content, novelty, and involvement tend to be sources of maintained or longer-lasting individual interest (Mitchell, 1993; Palmer, 2004; Renninger & Hidi, 2011). Through the use of AI by students then, their perceived personal relevance is boosted, novelty is experienced, and hands-on involvement is increased. Similar to how technology toys have been shown to foster an interest in technology (Haddock et al., 2022), the use of AI systems like ChatGPT could be a source of sustained individual interest. We state as a hypothesis: *H1: The use of AI leads to interest in AI.*

Next, positive attitudes can facilitate the transition from situational to individual interest by promoting engagement and persistence (Renninger & Shumar, 2002). Negative attitudes, on the other hand, can act as barriers, preventing situational interest from developing into a deeper, more sustained individual interest. This leads to the second hypothesis *H2: Positive (negative) attitudes towards AI foster (inhibit) interest in AI.*

When learners have an established individual interest in a domain (e.g., AI), they tend to have relatively higher levels of self-efficacy (Renninger & Hidi, 2002; Nuutila et al., 2020; Renninger, 2010). This effect is described in hypothesis *H3: Higher interest in AI leads to increased self-efficacy.* AI interest, therefore, should serve as a mediator between both AI attitudes and AI self-efficacy, as well as between AI use and AI self-efficacy.

Another aspect to consider in shaping students' AI self-efficacy is the role of AI literacy. Focusing on the personal level, according to Bandura's Social Cognitive Theory (2004), learners experience higher self-efficacy when they achieve success themselves (that is, achieve literacy). This influence on self-efficacy is identified, for example, by Carmichael et al. (2010) regarding mathematical achievement. Findings from Getenet et al. (2024) and Prior et al. (2016) indicate similarly that digital literacy significantly contributes to students' digital self-efficacy. Regarding self-efficacy in learning AI in particular, a study by Chai et al. (2021) indicates that AI literacy significantly predicts self-efficacy in learning about AI. We therefore add AI literacy as a second potential influence on AI self-efficacy to our theoretical model and propose hypothesis

H4 Higher AI literacy leads to increased AI self-efficacy.

Rooted in the Interest Development Model and based on our derived hypotheses, we propose a theoretical model describing the potential paths between the use of AI, AI interest, attitudes towards AI, and AI literacy, collectively leading to AI self-efficacy. The derived path model is illustrated in Fig. 1.

2.3. Groups regarding their relationship with AI and technology

Building on prior research that identified groups associated with variation in technology use and attitudes, we likewise aim to identify student groups based on cognitive, affective, and behavioral variables related to AI self-efficacy. Relatedly, research has already identified various groupings based on affective and behavioral variables like technology use and attitudes, which are closely linked to AI self-efficacy. For instance, technology readiness and attitudes can be conceptualized by two dimensions or parsed into as many as three or four major groups. Technology readiness as understood by Blut & Wang (2020) is described by a two-dimensional construct distinguishing between motivators such as innovativeness and optimism and inhibitors like insecurity and discomfort. Kerschner and Ehlers (2016) proposed a framework for attitudes toward technology, categorizing them into four main types: Enthusiasm, Determinism, Romanticism, and Skepticism. Regarding self-driving cars, Nielsen and Haustein (2018) identified three groups of people according to their attitudes: skeptics (38%), indifferent stressed drivers (37%), and enthusiasts (25%). Enthusiasts are typically young, male, highly educated, and reside in large urban areas, while skeptics tend to be older, more reliant on cars, and often live in less densely populated areas. Indifferent individuals frequently have less access to a car, highlighting the diminished relevance of self-driving technology in their lives.

Views specifically towards AI among learners vary, from some individuals being optimistic to others holding negative views. Negative perceptions of AI's impact on humans and society are well documented (Eagle et al., 2021; Ghotbi et al., 2022; Mertala et al., 2022; Oh et al., 2017). Some have general or vague fears about AI (Antonenko and Abramowitz, 2023; Lindner & Berges, 2020); yet there is also a sense of optimism and promise for addressing problems or benefiting society at large (Antonenko and Abramowitz, 2023; Teng et al., 2022). Meanwhile, some perceive AI as beneficial and a facilitator of personal ease (Mertala et al., 2022). Regarding the development of these conceptions over time, cross-country surveys report that global awareness of AI's potential impact is growing, with sentiment in Western nations slowly improving despite overall pessimism about its economic effects (cf. AI Index 2024 Annual Report: Maslej et al., 2024). In combination, these conceptions indicate that there is a range from people who are generally more skeptical towards AI to people who are more optimistic. Finally, in the field of so-called security-related predictive machines, a subdomain of AI, in a conceptual thought piece, Edwards (2015) described three groups: enthusiasts, critics, and skeptics. By considering underlying profiles that integrate cognitive, affective, and behavioral variables, we might similarly identify distinct groups among students. These analyses may unveil more comprehensive insights than focusing on just one variable and can help identify groups with different characteristics - and needs - based on cognitive, affective, and behavioral dimensions related to AI self-efficacy.

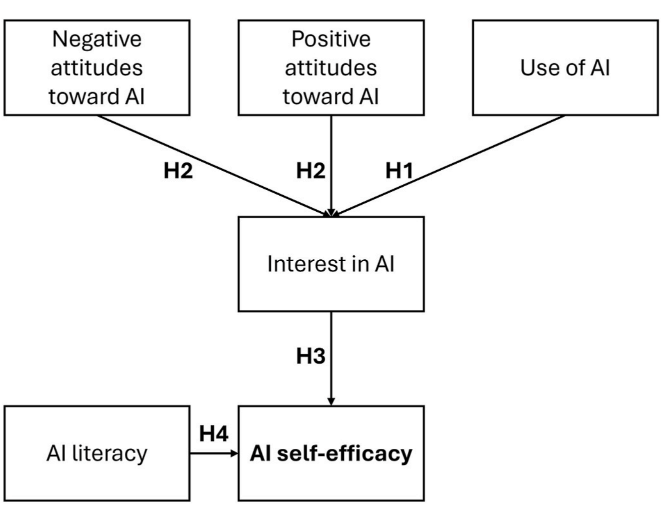

Fig. 1. Derived path model linking AI self-efficacy, AI literacy, AI interest, AI attitudes, and AI use.

Uncovering different groups regarding AI may indicate a need for a broader variety of learning materials that cater to diverse backgrounds, thus promoting inclusivity and accessibility in AI education. By identifying groups such as the critics mentioned above by Edwards (2015) among university students, there would be an indication that educational programs need to be designed to account for potential skepticism or fear of AI (Bewersdorff et al., 2023), perhaps transforming it into more informed caution and constructive criticism, thereby enhancing students' AI self-efficacy. In this context, it seems crucial not only to identify potential groups regarding AI, but also to investigate basic characteristics like the distribution of gender or disciplines among these groups to better tailor educational programs to students' needs.

3. Method

3.1. Sample

During December 2023 we recruited university students from the United States (US), the United Kingdom (UK), and Germany (GER). We focus on these three countries because they are global leaders in AI innovation and adoption (Stanford, 2024), and represent diverse educational systems and cultural and regulatory contexts within the Western world. Additionally, the availability of prior research in these countries enables comparative assessment for a substantial number of variables – the instruments tested are available in English as well as in German.[1]

The students were invited to participate in the web-based study via the Prolific platform. Participation in the study was voluntary; as a reward, participants received \$5.90, about £4.5 or 5.30€ respectively. A total of 1558 students from more than 512 universities and colleges participated. Before analyzing the data, we excluded 93 invalid cases. Fourteen participants were excluded because they failed both attention checks. An additional 75 cases were excluded because they did not finish the survey. Lastly, 3 participants were excluded because they completed the survey too quickly (less than 5 min, $Mdn = 19.1$), which made it implausible that they responded thoughtfully.

The final sample size was $N = 1465$ with $N_{US} = 494$, $N_{UK} = 499$, and $N_{GER} = 472$. The mean age of the participants was $M = 28.4$ (SD = 10.3). Among the participants, 747 (51.0%) were male, 680 (46.4%) were female, 33 (2.3%) were non-binary, and 5 (0.3%) preferred not to disclose their gender. The composition of disciplines can be found in Fig. 6 in

[1] The same sample was used to investigate country-specific differences regarding AI-related variables in Hornberger et al. (2024).

section 4.3.3 and Table 2 in the Technical Report. 1010 students are currently pursuing a bachelor's or similar degree (68.9%), 393 (26.8%) a master's or similar, 61 (4.2%) of students are enrolled in other programs like graduate/Ph.D. programs, and one student did not indicate their student status.

3.2. Instruments

To measure **AI self-efficacy**, we used a scale previously published and validated (Hornberger et al., 2023). The scale is based on the competencies described by Long and Magerko (2020) and consists of eight items (e.g., "I have a good understanding of the basic principles of AI."). The response format consisted of a 5-point Likert scale from 1 = strongly disagree to 5 = strongly agree. With α = .83, internal consistency is good (Tavakol and Dennick, 2011).

AI literacy was measured using the AI literacy test instrument developed by Hornberger et al. (2023) and validated with samples in the US, UK, and Germany (Hornberger et al., 2024). The instrument builds on Long and Magerko's (2020) conceptualization of AI literacy and consists of 27 4-choice items and one sorting item (arranging the steps of machine learning in order). Internal consistency was acceptable (Cronbach's α = .76).

To measure **interest in AI**, we used a scale we published in Hornberger et al. (2023). The scale builds on the 2015 PISA study (Mang et al., 2019) which was originally designed to measure interest in science. We modified the scale by exchanging the word "science" with "artificial intelligence." The scale consists of five items (e.g., "I generally have fun when I am learning about artificial intelligence."). The response format is a 5-point Likert scale ranging from 1 = strongly disagree to 5 = strongly agree. Internal consistency was good at α = .91.

We assessed **positive and negative attitudes towards AI** using eight items from the General Attitudes Towards AI scale by Schepman and Rodway (2020). We used four items measuring positive attitudes (e.g., "There are many beneficial applications of Artificial Intelligence.") and four items measuring negative attitudes (e.g., "I think Artificial Intelligence is dangerous."). The response format was again a 5-point Likert scale from 1 = strongly disagree to 5 = strongly agree. Internal consistency was good, with α = .81 and α = .82, respectively.

We measured the **use of AI** with two items asking about the use of AI systems in respondents' personal lives as well as part of their education, again on a 5-point Likert scale. With α = .79, internal consistency was acceptable.

3.3. Analysis

3.3.1. Calculating IRT scores and correlations

For the analysis, we used Item Response Theory (IRT) scores to adjust for variations in item difficulty and discrimination - the ability of items to differentiate between respondents - allowing us to measure the underlying trait more accurately. IRT is a method that examines how people answer questions, and which links their answers to their level of the measured trait. This approach takes into account both the difficulty of the items and how well they distinguish between individuals, making the measurement more reliable and precise (Embretson and Reise, 2013).

For AI literacy, we used person abilities estimated by a 3-PL model with a 25% guessing rate. For all other scales described in 3.2, we used rating scale models. For detailed information about the validation of the IRT models and the calculation of the IRT scores, see Hornberger et al. (2024). For correlations, we calculated Pearson's *r* (Gupta & Kapoor, 2020) and associated significance levels.

3.3.2. Path model analysis

Following research question 2, we aim to investigate how AI literacy, AI interest, attitudes towards AI, and use of AI influence AI self-efficacy. To investigate basic relationships between these cognitive, affective, and behavioral variables based on the conceptual foundation (see 2.2), we conducted a path analysis, which is a specialized case of Structural Equation Modeling (SEM: Ullman & Bentler, 2012). Path analysis is a statistical technique employed to empirically examine the directional relationships between multiple variables, allowing for the estimation of both direct and indirect effects within a predefined model structure (Lleras, 2005). At its core, path analysis involves specifying a model, based on theoretical considerations and prior empirical evidence, which outlines the expected relationships between observed variables. These relationships are depicted as paths, with each path representing a hypothesized direct influence of one variable on another. The strength and significance of these paths are quantified through path coefficients, which are estimated from the data.

We initiated our analysis by defining a path model structure aligned with our theoretical expectations (see section 3) and previous empirical findings (see Hornberger et al., 2023). This process was facilitated by the use of the Python-based SEM tool semopy (Igolkina & Meshcheryakov, 2020). For all variables, normality of the distribution was checked visually via histograms and Q-Q-plots. Visual inspection showed all variables are roughly normally distributed (see Technical Report, B.1.). To assess the fit of our path model, we investigated multiple fit indices, each offering insights into different aspects of model evaluation. Fit indices investigated were: Degrees of Freedom (DoF) and DoF Baseline, Chi-squared (χ^2) test and its associated p-value, Comparative Fit Index (CFI), Adjusted Goodness of Fit Index (AGFI), the Normed Fit Index (NFI) and the Tucker-Lewis Index (TLI), the Root Mean Square Error of Approximation (RMSEA), the Akaike Information Criterion (AIC), Bayesian Information Criterion (BIC), and Log-Likelihood (LogLik). A detailed description of the fit indices can be found in the Technical Report, section A.1. For an introduction to fit indices and path models, refer to Hu and Bentler (1999), McDonald (1996) and Stage et al. (2004).

To additionally investigate *generalizability* of our path model, we aim to validate the model using the overall sample as well as country-specific subsamples from the United States, the United Kingdom, and Germany. A path model demonstrating a good fit across multiple datasets (cross-validation) is likely to generalize well and would support the model's overall validity and reproducibility.

3.3.3. Component analysis with Gaussian Mixture Models

Following research question 3, we aim to identify distinct student groups based on their AI literacy, AI self-efficacy, AI interest, attitudes towards AI, and AI use. To explore these potential groups of similar cognitive, affective, and behavioral patterns among students, we chose Gaussian Mixture Models (GMM: Bishop & Bishop, 2024). A GMM is a probabilistic model that represents the presence of *components* (also called *groups* or *profiles*) within an overall dataset by modeling the data as a mixture of multiple Gaussian distributions. For all analyses regarding GMM, we used R and the mclust package (Fraley et al., 2024).

To choose and validate a model, we split the dataset into two randomly selected subsets referred to as the training dataset and validation dataset (Fig. 2). As the BIC/AIC scores are sensitive to the sample size, to sustain comparability between training and validation data, we

Table 2
Correlations between AI literacy, AI self-efficacy, AI interest, positive and negative attitude towards AI, and use of AI.

	AI literacy	AI interest	Pos. att. tow. AI	Neg. att. tow. AI	Use of AI
AI self-effic.	.14***	.49***	.33***	−.14***	.29***
AI literacy		.06*	.13***	−.13***	.00
AI interest			.65***	−.32***	.56***
Pos. att. tow. AI				−.44***	.51***
Neg. att. tow. AI					−.25***

choose to split the dataset into equal sizes. On the training data, for 1 to 6 components, we ran all 14 GMMs available with the mclust package with configurations ranging in their variances and covariances for one to six components. For each specified number of components, mclust picks the best-fitting model according to its BIC value. This gives us the best models according to BIC for every specified component structure from 1 to 6, and leaves us with one model for each parsing of components. Out of these six final models, we determined the best-fitting model by evaluating and comparing these models based on their fit indices, including BIC, ABIC, AIC, and CAIC. More information about the fit indices can be found in the Technical Report, section A.2.

The model and identified number of components is then validated with the validation data (see Fig. 2). A model with a specific number of components that shows a good fit for both training and validation datasets suggests that the model is likely generalizing well to new data. Consistency in performance across both subsets would indicate the absence of overfitting and support the validity of the model.

For further insights into the data and components, the total dataset was used. *Z*-scores were calculated for all cognitive, affective, and behavioral variables to standardize the data. This transformation allows for the comparison of variables on a common scale, which makes the results more accessible for interpretation. Next, using MANOVA, we performed an analysis to check if the profiles/groups identified by GMM significantly differ from each other (Tabachnick & Fidell, 2011).

Following research question 4, to investigate differences in socio-demographic variables and perceived AI course availability among members of different groups, we calculated descriptive statistics, including total counts and percentages.

4. Results

4.1. Correlations (RQ1)

Table 2 shows the correlations among cognitive, affective, and behavioral IRT scores regarding AI.

The correlation table reveals strong positive relationships between AI interest and positive attitudes towards AI (0.65) and between AI interest and use of AI (0.56). Additionally, there is a high negative correlation between positive and negative attitudes towards AI (-0.44), suggesting that as positive attitudes increase, negative attitudes decrease. In contrast, AI literacy shows weak correlations with most other variables and is not correlated with the use of AI ($p = .86$).

4.2. Path analysis (RQ2)

4.2.1. Model fit, validation, and generalization of the path model

To confirm the fit and parsimony of our theoretically derived path model (see 2.2), we conducted evaluations using the total sample and country-specific subsamples from the United States, the United Kingdom, and Germany (Table 3). By keeping the model complexity consistent across all groups, evidenced by uniform Degrees of Freedom (DoF) set at 14, our analysis revealed a very good match with the observed data, particularly within the total sample and the US subsample. This finding is strongly supported by a χ^2 value of 3.102 in the total sample and an exceptionally high *p* value of .999, indicating a very good fit of the model. Further validation comes from the Comparative Fit Index (CFI) surpassing the threshold for excellence, along with nearly perfect Goodness-of-Fit Index (GFI) and Adjusted Goodness-of-Fit Index (AGFI) scores, reinforcing this conclusion. An RMSEA of zero for a sample might usually raise overfitting concerns. However, as the model shows an RMSEA of zero for both the total and the US subsample but a very good non-zero RMSEA for the UK and the German sample, it overall seems to highlight the model's fit.

Table 3
Fit indices of the path model for the total sample and country-specific subsamples from the United States, the United Kingdom, and Germany.

Index	Value (Total Sample)	Value (US Sample)	Value (UK Sample)	Value (GER Sample)
DoF	14	14	14	14
DoF Baseline	19	19	19	19
χ^2	3.102	1.502	20.694	14.689
χ^2 p-value	0.999	1.000	0.110	0.400
χ^2 Baseline	2222.967	825.830	856.557	656.358
CFI	1.005	1.015	0.992	0.999
GFI	0.999	0.998	0.976	0.978
AGFI	0.998	0.998	0.967	0.970
NFI	0.999	0.998	0.976	0.978
TLI	1.007	1.021	0.989	0.999
RMSEA	0.000	0.000	0.031	0.010
AIC	13.996	13.994	13.917	13.938
BIC	51.023	43.412	43.405	43.037
LogLik	0.002	0.003	0.041	0.031

Looking at the UK and German subsamples, there is a noticeable change in how well the model fits. Increased χ^2 values for these metrics point to a robust but slightly less ideal fit compared to for the total sample and US subsample. However, the CFI values, though slightly lower, still confirm a good fit of the model. GFI and AGFI scores, although the lowest among the samples, remain within acceptable ranges. Notably, the RMSEA values for these subsamples, which is non-zero for both the UK and Germany, still suggest a close match and offer a nuanced picture of the model's applicability across varied demographic contexts.

The AIC and BIC metrics across different samples illustrate the balance between model simplicity and its ability to fit the data accurately. In particular, the metrics from the US subsample show lower AIC and BIC values, indicating a good mix of simplicity and compatibility with the data.

Additionally, we tested the model for a non-linear relationship between AI literacy and AI self-efficacy, considering the potential influence of the Dunning-Kruger effect (Dunning, 2011). Both linear and non-linear models fit well overall. While the linear model demonstrates slightly better fit statistics, the non-linear model shows marginally improved AIC and BIC values, indicating better parsimony.

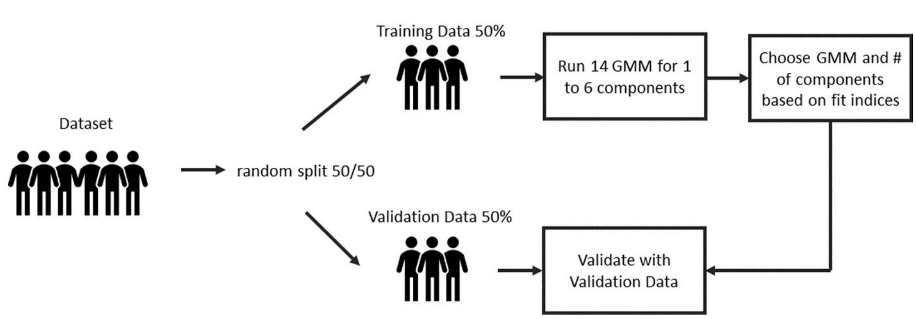

Fig. 2. Process of selecting and validating the GMM and the number of components.

However, given that the path coefficient in the non-linear model is non-significant ($p = .99$), we opted for the linear model, which provides an excellent fit and avoids introducing a non-significant term.

4.2.2. Unveiled relationships between cognitive, affective and behavioral variables

Based on the theoretically-derived hypotheses, we fitted our model using the entire sample to describe the potential pathways leading to AI self-efficacy. In line with our hypothesis H2, a positive attitude towards AI is significantly and positively linked to AI interest, with a sizable path coefficient of 0.618. On the other hand, the influence of negative attitudes towards AI seems to negatively (−0.057) impact AI interest, but only to a very marginal degree. However, the lack of statistical significance ($p = .12$) for this relationship questions the viability of negative attitudes in reliably forecasting interest in AI. This is a deviation from the hypothesized model.

Next, use of AI is a strong (0.434) and significant ($p < .01$) predictor of interest in AI, confirming H1. This finding, with important practical implications, emphasizes the role of AI interaction in cultivating AI interest.

As predicted in H3 and H4, AI self-efficacy is indeed positively affected by both AI literacy and AI interest. AI literacy shows a modest yet positive (0.144) and significant ($p < .01$) impact on AI self-efficacy, suggesting the presence of other influential factors not accounted for in the current model. The significant ($p < .01$) path coefficient of 0.245 for the impact of AI interest on AI self-efficacy highlights another robust influence on AI self-efficacy. All paths and their coefficients are displayed in Fig. 3.

The results unveil robust paths from positive attitudes towards AI and from AI use to AI interest, as well as paths from AI interest and AI literacy to AI self-efficacy. Conversely, there is only a small and statistically insignificant influence of negative attitudes toward AI on interest.

We additionally tested the direct effects of both positive and negative attitudes towards AI, as well as use of AI, on AI self-efficacy. All direct effects were non-significant, confirming that AI interest serves not just as a medaitor, but as a *full* mediator between both AI attitudes and AI self-efficacy, as well as between AI use and AI self-efficacy.

4.3. Student groups regarding AI (RQ3 & RQ4)

4.3.1. Model fit, validation, and generalization of the GMM

To identify the preferred number of distinct student groups, we fitted multiple GMMs for one to six components (Fig. 2). A variety of fit indices were evaluated for the selection of the model and appropriate number of components (groups). Generally, the model showing the best fit across all fit indices (see 3.3.3) should be chosen.

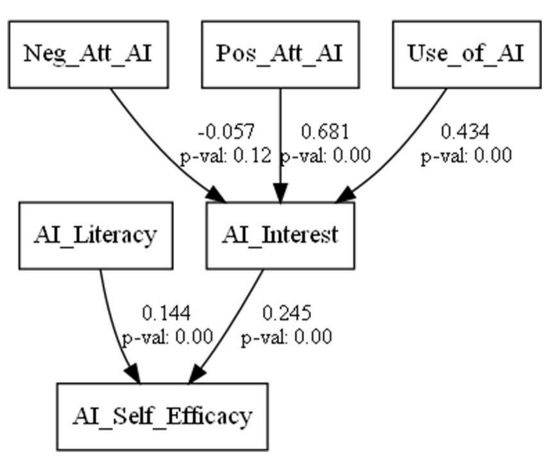

Fig. 3. Path model of the interplay of attitudes, use, interest, literacy, and self-efficacy regarding AI ($N = 1465$).

The BIC/ABIC scores of our training dataset suggest that the best fitting model for our data is the "VVE" model with three components. The "VVE" model is a flexible yet constrained approach that allows the covariance ellipsoids of each component to have different volumes and shapes while sharing the same orientation (see the *mclust* package, Fraley et al., 2024). The fit of this model is confirmed by the AIC/CAIC of our training dataset as well as by the AIC/CAIC of our validation data (Fig. 5). Only the BIC/ABIC (Fig. 4) for the validation dataset prefers a model with only two components. This can be explained by the BIC generally having an emphasis on parsimony but a tendency towards underfitting (Dziak et al., 2020). Further, the AIC and CAIC are more likely to favor models with more components for larger datasets because the improved fit often outweighs the penalty for additional parameters. This explains why the AIC and the CAIC of the total dataset indicate five components over three components. Overall, the BIC/ABIC of the full dataset (test and training dataset combined) supports the choice of three components or student groups.

The fit indices BIC, ABIC (Fig. 4), as well as the AIC and CAIC (Fig. 5) of the training as well as the validation dataset favor the three-component "VVE" model.

4.3.2. Major components revealed by the GMM

In the previous section, 4.3.1, we identified distinct groups within our data based on clusterings with respect to AI literacy, AI self-efficacy, AI interest, attitudes towards AI, and AI use. We now aim to further investigate differences among these groups regarding these variables. In particular, the analysis with GMM revealed three different components (groups) of students (Fig. 6). 703 of 1465 students (47.99%) can be labeled as 'AI Advocates.' This group shows above-average AI literacy, AI self-efficacy, AI interest, and positive attitudes toward AI, coupled with below-average negative attitudes and high usage of AI. The group of 'Cautious Critics' (305 of 1465, 20.82%) shows below-average AI literacy, AI self-efficacy, AI interest, and positive attitudes toward AI, coupled with high negative attitudes and low usage of AI. Finally, around three of ten students can be described as 'Pragmatic Observers' (457 of 1465, 31.19%). This group has scores close to the mean in most variables, suggesting a balanced and moderate relationship related to AI. A table of the group labels, sizes, and mean z-scores across variables,

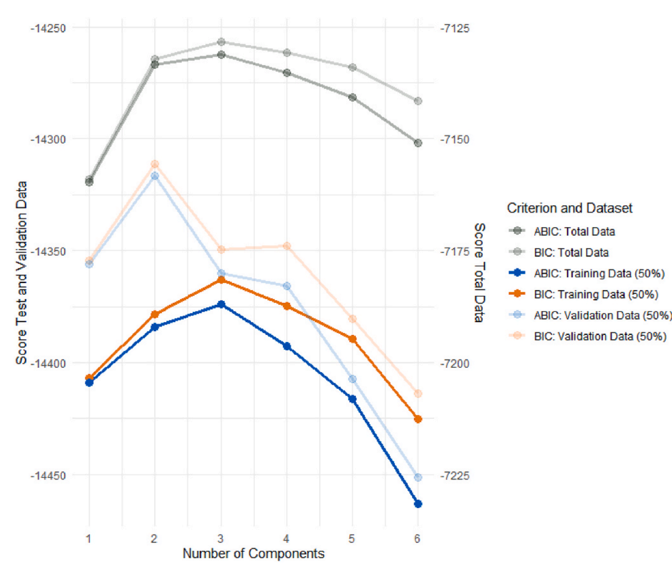

Fig. 4. BIC and ABIC for Gaussian Mixture Models (GMM) with 1–6 components.
Note. The y-axis for the training and validation data differ from the y-axis for the full dataset. However, for better visualization and comparison, all three datasets are plotted on a single graph. With the mclust package, values closer to zero for all indices indicate better fit (Wardenaar, 2024).

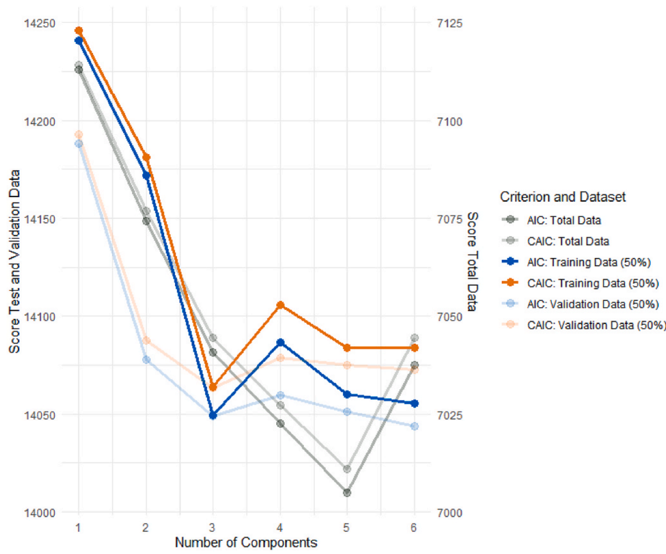

Fig. 5. AIC and CAIC for Gaussian Mixture Models (GMM) with 1–6 components.
Note. The y-axis for the training and validation data differ from the y-axis for the full dataset. However, for better visualization and comparison, all three datasets are plotted on a single graph. With the mclust package, values closer to zero for all indices indicate better fit (Wardenaar, 2024).

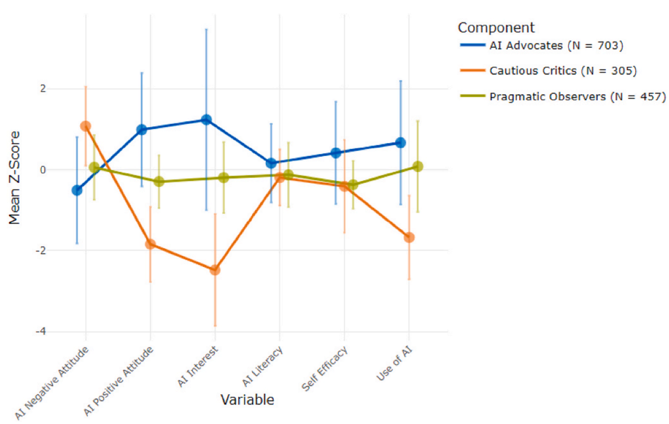

Fig. 6. Component means and standard derivations across variables.

as well as a visualization of the groups by their first principal component, can be found in the Technical Report (B.2., Table 1).

The MANOVA results, summarized using Pillai's trace, confirm significant differences between the three groups ($F = 53.44$; $p < .01$). However, the Pillai's trace value (0.180) suggests that the groupings explain a relatively modest proportion of the overall variance in the dependent variables.

4.3.3. Distribution of disciplines among student groups

In the full sample, the most prevalent discipline is Engineering and Technology, constituting 29.42% of individuals, followed by students of Social Sciences at 25.94%, and Art and Humanities at 20.00%. In comparison with the total sample, the three subgroups exhibit distinct disciplinary profiles. 'Cautious Critics' primarily consists of students enrolled in Arts and Humanities (30.16%) and Social Sciences (29.51%), diverging from the overall trend with less presence in technical fields. Conversely, 'AI Advocates' stand out with a strong majority in Engineering and Technology (38.69%), highlighting a significant deviation towards more technical disciplines in contrast to the total sample. Finally, 'Pragmatic Observers' display a mostly balanced disciplinary distribution. Results are displayed in Fig. 7 and Table 2 in the Technical Report.

4.3.4. Distribution of gender among student groups

In the total sample, gender distribution leans slightly towards males at 50.99%, with females representing 46.42% of the respondents, indicating a fairly balanced but slightly male-skewed demographic. Among the groups, 'Cautious Critics' have a majority female representation at 61.97%, contrasting with 'AI Advocates' where males constitute 62.78%, indicating a notable gender disparity in these groups. 'Pragmatic Observers' show a balanced gender distribution, with females at 54.05% and males at 44.64%. Results are displayed in Fig. 8 and Table 3 in the Technical Report.

4.3.5. Perceived course availability among student groups

Within the total sample, there is notable uncertainty about AI course availability, with 32.63% of students indicating they're "not sure" if course availability is sufficient. The next largest group, 27.85%, say there are not enough courses, suggesting a general perception of either uncertainty or insufficient AI course offerings. More than half of the 'Cautious Critics' report unawareness of potential AI courses (55.41%). They are likewise most skeptical about the sufficiency of suitable courses (15.08%). On the other hand, 'AI Advocates' are the group most satisfied with the course offerings (34.00%), though a significant portion still perceive there could be more courses (30.01%). Results are displayed in Fig. 9 and Table 4 in the Technical Report.

5. Discussion

5.1. Discussion of the derived path model: fit, accuracy, and generalizability

Based on prior theoretical findings, we derived and validated a potential path model - though we acknowledge there might be alternative models which lead to better theoretical as well as empirical fit. This is especially plausible in this domain as many observed relationships are sensitive to tasks and content characteristics (Nuutila et al., 2020). Indeed, it's important to keep in mind that studies often report mixed and inconclusive results, depending on the models used and the variables and their definitions.

Yet, despite slight variations in fit indices, the strong fit indices in all four (sub)samples suggest the model's validity and its ability to effectively capture the intended constructs across diverse settings in Western countries. This analysis reinforces the model's theoretical and empirical utility, enhancing our confidence in its applicability and generalizability

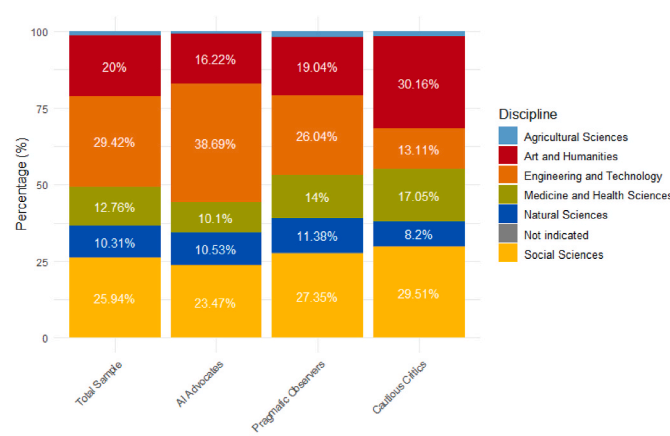

Fig. 7. Distribution of disciplines among the total sample and identified student groups.
Note. Percentages <5 are not labeled.

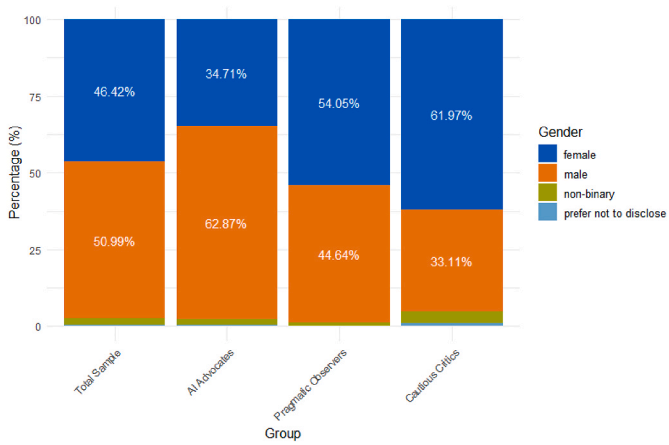

Fig. 8. Distribution of gender among the total sample and identified student groups.
Note. Percentages <5 are not labeled.

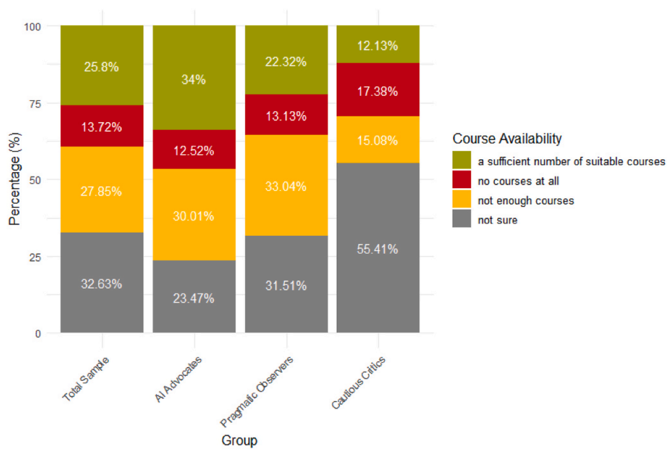

Fig. 9. Perceived course availability by the total sample and the three different groups.

across varied contexts in Western nations. Nevertheless, while our model has shown validity and good fit across three prominent Western countries, we recognize that it may not be directly applicable to countries worldwide. More research is needed to explore model fit in other educational systems as well as other potential models, as this model, though theoretically grounded, represents only one possible approach that could be considered.

5.2. Fostering AI self-efficacy: the interplay of AI literacy, AI interest, use of AI, and attitudes towards AI

The core variable studied here - AI self-efficacy, AI literacy, AI interest, attitudes towards AI, and use of AI - show significant correlations, unveiling complex interconnected dynamics across these variables. As expected, negative attitudes towards AI correlate negatively with the other more 'positive' metrics. Our findings also indicate that AI literacy has weaker correlations with variables like AI interest compared to what has been reported in other studies (e.g., Hornberger et al., 2023). Notably, AI literacy is not at all correlated with the use of AI, which strongly cautions that there may be no meaningful direct relationship between AI literacy and use of AI. This could mean that students may use AI tools without developing understanding of the underlying concepts and suggests that casual use of AI doesn't contribute much to a deeper understanding of AI. This finding aligns with the debunked myth of 'digital natives,' who were often imagined as having deep knowledge of technology when, in reality, they are often limited to basic skills and uses of technology (Kirschner & De Bruyckere, 2017). This finding is somewhat contradictory to arguments made by Koch et al. (2024), who suggest that individuals who use AI might become more competent regarding AI. One possible reason for this contradiction is the difference in the operationalization of 'use of AI.' While Koch et al. (2024) focused on structured educational settings, our study incorporated casual use, which may not contribute significantly to AI literacy, but nonetheless may constitute a major portion of AI use.

With our path model, we demonstrate that fostering AI self-efficacy seems to be a multi-step process involving cognitive, affective, and behavioral variables. Confirming our hypothesis H1, use of AI has a strong and significant influence on AI interest. This aligns with the Interest Development Model (Hidi and Renninger, 2006; see also 2.2) stating that the initial capture of attention due to novel or appealing stimuli triggers interest. The finding here emphasizes the role of (first-time) interaction with AI specifically in fostering an interest in AI. Hands-on experiences seem to make AI more tangible and accessible, thereby sparking interest. One example is the introduction of AI systems like ChatGPT and Gemini in late 2022 and early 2023, which sparked a significant increase in AI interest. Google searches for the term 'artificial intelligence' increased about six to tenfold from September 2022 to mid-2023 (Google Trends, 2024). AI education should thus might (e.g., at the beginning of a course) incorporate interaction with AI systems to kindle student interest - especially among students who might not otherwise use AI in their daily lives (see 4.3).

Next, in line with our hypothesis H2, positive attitudes regarding AI *are* a significant predictor of interest in AI. Attitudes can foster individual interest by promoting engagement and persistence (Renninger & Shumar, 2002). However, contrary to our expectations, the influence of negative attitudes on interest in AI was found to be non-significant. One possible explanation is that the positive aspects of AI might overshadow negative attitudes, thereby diminishing the impact of negative attitudes on overall AI interest. There might also be a potential methodological issue of collinearity between positive and negative attitudes, which could have obscured the unique contribution of negative attitudes in our model, resulting in non-significant findings. However, Schepman and Rodway (2020) demonstrated that their positive and negative attitude scales load onto two distinct factors, thereby mitigating concerns regarding collinearity. Thus, AI education programs which focus on building positive attitudes toward AI could promote higher AI self-efficacy and greater AI interest. However, educators should also grapple with how to present balanced information about AI's benefits and risks so that students are empowered to make their own informed decisions about when to use – or not use – AI.

Along these lines, as per Renninger and Hidi (2002) and hypothesis H3, students with interest in AI have relatively higher levels of AI self-efficacy. Education strategies should, therefore, aim to maintain and grow AI interest.

Finally, the relationship between AI literacy and AI self-efficacy modestly supports hypothesis H4 that understanding AI concepts can increase AI self-efficacy. That is, the modest size of this coefficient (0.144) implies that AI literacy alone does not account for the majority of variance in AI self-efficacy. Focusing solely on AI literacy may not be sufficient. Instead, educators should combine AI literacy efforts with fostering attitudes and interest as well as interaction with AI systems.

Summarized, the validated path model suggests that providing more opportunities for direct interaction with AI, coupled with fostering a positive attitude towards it, can significantly enhance interest in AI. In turn, interest, in combination with fostering AI literacy, can boost AI self-efficacy.

5.3. Three groups regarding AI: AI Advocates, Cautious Critics, and Pragmatic Observers

Based on the fit indices BIC and ABIC (Fig. 4) as well as AIC and CAIC

(Fig. 5), from both the training and validation datasets, empirical data prefer a three-component model. This three-component spectrum of groups aligns both with theoretical frameworks (Edwards, 2015) and prior research regarding people's attitudes toward self-driving cars (Nielsen & Haustein, 2018). Given the empirical evidence and its alignment with prior findings in the literature, we selected a model with three components: AI Advocates, Cautious Critics, and Pragmatic Observers. This decision to prefer a three-component model instead of a five-component model is further supported by the principle of parsimony (often referred to as 'Occam's Razor'; Braithwaite, 2007), which advises choosing the simplest model when multiple models demonstrate adequate fit. However, despite our preference for the three-component model, we acknowledge the possibility that a greater number of components might reveal more nuanced insights into our dataset.

Although the variance in AI self-efficacy explained is modest - suggesting few *fundamental* differences among many members - the groups still represent a meaningful division along a continuum of socio-cognitive relationships with AI. This finding reflects the nuanced and dynamic nature of students' relationships with AI. From a statistical perspective, the modest proportion in how much of AI self-efficacy is explained could be due to the shared variance among the often correlated variables, which can reduce the distinct contribution of each variable to the overall effect. Nevertheless, the three groups provide a useful framework for understanding differences in student engagement with AI education.

For instance, one of the key findings is that AI literacy manifests through only minor differences across the three groups, indicating that it plays a less significant role in shaping students' relationship with AI. This suggests that focusing exclusively on improving AI literacy might not be enough to fully engage students, particularly those with low self-efficacy or negative attitudes toward AI. The weak correlations between AI literacy and self-efficacy (see section 4.1) instead highlight the need for a broader approach to AI education that considers not only knowledge acquisition but also students' emotional and cognitive relationships with AI.

Caveats asid, the three groups - AI Advocates, Cautious Critics, and Pragmatic Observers - do represent distinct relationships with AI. AI Advocates, nearly half of the students (47.78%), are characterized by high AI literacy, self-efficacy, interest, and positive attitudes toward AI, alongside high usage of AI. This group's enthusiasm suggests they are likely to engage deeply with AI education, though there is a risk that their optimism could lead to underestimating potential risks or ethical concerns associated with AI technologies. Educational strategies for this group should maintain their enthusiasm while introducing critical perspectives on the ethical, societal, and security-related aspects of AI. Cautious Critics (20.82%) exhibit low AI literacy, AI self-efficacy, and AI interest, coupled with negative attitudes and low engagement with AI. This group aligns with the 'Critics' in security-related AI contexts identified by Edwards (2015), who may hold apprehensions about the implications of AI technologies. For this group, educational initiatives should prioritize building AI interest and addressing their concerns directly, potentially through more accessible and relatable AI content that demystifies the technology and highlights its practical, safe applications and the prospects for the students themselves to responsibly develop and use of AI in light of concerns. Pragmatic Observers (30.17%) fall near the mean across most variables, showing a balanced, moderate attitude toward AI. Their lack of strong opinions or engagement resembles the 'Indifferent' group noted by Nielsen and Haustein (2018), indicating a passive, observant approach to AI. This group could benefit from targeted educational programs that encourage deeper exploration and engagement without overwhelming them with either overly technical content or with accentuated promises or fears about AI's benefits or risks.

The identification of distinct student groups thus highlights the need for differentiated educational strategies in AI curricula. Current AI programs, which often emphasize technical skills and AI literacy alone, may fall short in addressing the diverse cognitive, affective, and behavioral aspects of students' relationships with AI. Identifying these student groups enables educators to better allocate resources and develop support systems that directly target the barriers each group faces. For instance, AI Advocates, with their high AI literacy, AI self-efficacy, and positive attitudes, might thrive in traditional AI courses that focus on technical knowledge. For AI Advocates, the challenge might be in maintaining their engagement while encouraging critical reflection on AI's societal and ethical implications. Cautious Critics might need support in building AI self-efficacy and AI interest, potentially through more accessible, relatable content and active participation with AI. Pragmatic Observers, who exhibit moderate engagement and balanced attitudes, may require additional motivation and tailored content to deepen their involvement and interest with AI learning. By applying such tailored approaches, educators can foster more inclusive AI programs that focus on the diverse needs of all students, ensuring no group is left behind.

5.4. *Socio-demographic characteristics of the student groups regarding AI*

5.4.1. *Distribution of disciplines among student groups*

Building on the identification of these groups, we further explored the socio-demographic characteristics to understand the underlying factors influencing AI relationships. The distribution of disciplines among students reveals insights into how different academic backgrounds shape relationships to AI. The finding that Cautious Critics are predominantly from Arts and Humanities (30.16%) and Social Sciences (29.51%) is in line with prior research (Mercader & Gairín, 2020; Stöhr et al., 2024). It might point to a potential disconnect between AI education efforts and non-technical fields. This group's skepticism towards AI might be rooted in the perception that AI is not relevant to their disciplines, and beyond their ability to exercise agency. However, demonstrating the applicability of AI in their disciplines could address this gap. Projects that integrate AI into real-world applications within these disciplines may encourage more interest and participation from non-technical students. On the other hand, AI Advocates are primarily from Engineering and Technology (38.69%), which reflects their increased comfort and familiarity with AI. However, their significant representation in Social Sciences (23.47%) shows growing interest in non-technical fields. Promisingly, this provides an opportunity for universities to foster interdisciplinary collaborations between technical and non-technical students in AI education. Lastly, pragmatic Observers are distributed more evenly across various disciplines, suggesting that they may take a broader, more practical 'wait-and-see' approach to AI. Their heterogeneous academic backgrounds call for AI education that is relevant to a wide range of fields, including individuals who are not likely to specialize in AI.

5.4.2. *Distribution of gender among student groups*

The gender distribution within the groups also highlights noteworthy differences in how men and women perceive AI. The male dominance among AI Advocates (62.87%) reflects broader trends seen in technology-related fields, where men are traditionally overrepresented (Verdugo-Castro et al., 2022). In contrast, the female majority among Cautious Critics (61.97%) suggests that women are more likely to be skeptical of AI and experience lower AI self-efficacy and AI interest. Our findings align with Stöhr et al. (2024), who report in a study on the perception of students of AI chatbots that males have higher usage, more positive attitudes, and less concern than females. This gender gap additionally aligns with societal and cultural factors often discouraging women from engaging with technology (Cheryan et al., 2017). To address this gap, universities should work to break down these barriers by making AI education more inclusive. For example, they could add courses to the existing curriculum that highlight the relevance of AI in fields where women are traditionally more represented, such as healthcare, education, and social services. Alternatively, existing courses could be transformed to incorporate more gender-relevant content.

5.4.3. Perceived course availability among the student groups

The findings on perceived AI course availability reveal significant challenges in ensuring that students have access to AI education. The fact that 32.63% of students are unsure about course availability suggests a communication gap between universities and students. This lack of awareness could be due to insufficient promotion of AI courses or inadequate integration of these courses into broader curricula. To address this, universities need to enhance their outreach strategies to ensure that students across all disciplines are well-informed about the AI courses available to them. Improved course advertising, collaboration between departments, and better visibility of AI-related opportunities are essential steps in bridging this gap.

Cautious Critics are the most uncertain, with 55.41% unsure about course availability and 17.38% stating that no AI courses are offered. This group's lack of awareness and lackluster feelings towards AI coursework highlights the need for more targeted outreach and AI courses that show the relevance of AI in their respective domains. On the other hand, AI Advocates are generally more satisfied with course availability, though a notable 30.01% still report a lack of courses. This suggests that even students who are already inclined toward AI see room for improvement in the course offerings.

Finally, while we analyzed the student groups regarding AI by course availability, discipline, and gender, we note that there may be additional underlying patterns influencing these relationships worth exploring that we did not examine here. For instance, Hong (2022) demonstrated that social status influences AI self-efficacy and intentions to use AI.

6. Conclusion

This study highlights the critical role of cognitive, affective, and behavioral variables in shaping university students' AI self-efficacy. We derived and validated a path model that describes interconnections between these variables.

The findings emphasize the importance of fostering positive attitudes toward AI, creating opportunities for hands-on interaction, and addressing the specific concerns of students with lower AI self-efficacy or skepticism. While AI literacy remains important, the results suggest that AI interest, use of AI and positive attitudes also play a ssubstantial role in fostering AI self-efficacy. Furthermore, we identified three distinct groups of university students regarding AI: 'AI Advocates,' 'Cautious Critics,' and 'Pragmatic Observers,' each representing different levels of interest in and engagement with AI, as well as varying attitudes toward AI. This underscores the need for differentiated AI programs that meet the diverse needs of these students, ensuring that no group is neglected when key educational design choices are made. Overall, this research contributes to the literature by identifying distinct student profiles concerning AI engagement, highlighting the limited role of AI literacy alone in fostering AI self-efficacy, and indicating other viable pathways for promoting AI self-efficacy.

Several limitations should be acknowledged. First, the research was conducted exclusively in Western countries, which might limit the generalizability of the findings to non-Western contexts with different cultural and educational systems. Second, the study employed a cross-sectional design, capturing data at a single point in time. This approach does not account for changes and developments over time, important given the dynamic nature of AI technically and socially. Third, we used the platform Prolific to generate our sample, but did not rely on representative samples given our focus on cross-national comparisons. The sample thus is unlikely to fully represent all demographics, e.g., by socioeconomic status, especially given our focus on university students. This potential lack of diversity could influence the results, as these factors might play a significant role in AI self-efficacy and attitudes toward AI.

Future studies should investigate the impact of tailored educational interventions on the AI self-efficacy of the individual groups identified here, such as the Cautious Critics, over time. Additionally, examining cultural factors in non-Western contexts could reveal whether certain patterns are universal or culture-specific. For a more comprehensive understanding, research should also investigate how socioeconomic status, access to resources and other surrounding variables influence socio-cognitive relationships with AI, further providing insights for the development of effective AI programs accessible to all. Finally, this research should be extended to other populations, like non-students.

CRediT authorship contribution statement

Arne Bewersdorff: Writing – review & editing, Writing – original draft, Visualization, Project administration, Methodology, Investigation, Conceptualization. **Marie Hornberger:** Writing – review & editing, Project administration, Data curation, Conceptualization, Methodology. **Claudia Nerdel:** Writing – review & editing, Supervision, Resources, Project administration, Funding acquisition, Conceptualization. **Daniel S. Schiff:** Writing – review & editing, Validation, Supervision, Project administration, Methodology, Funding acquisition, Conceptualization.

Declaration of AI and AI-assisted technologies in the writing process

During the preparation of this work, the authors used ChatGPT (www.chat.openai.com) as well as Grammarly (www.grammarly.com) in order to improve the readability and language of single sentences. After using these tools, the authors reviewed and edited the content as needed and take full responsibility for the content of the publication.

Funding

This work was supported in part by funding from Google Research. It was funded in part by the German Federal Ministry of Education and Research (16DHBKI051). The responsibility for the content of this publication lies with the authors.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.caeai.2024.100340.

References

Antonenko, P., & Abramowitz, B. (2023). In-service teachers'(mis) conceptions of artificial intelligence in K-12 science education. *Journal of Research on Technology in Education, 55*(1), 64–78.

Ayllón, S., Alsina, Á., & Colomer, J. (2019). Teachers' involvement and students' self-efficacy: Keys to achievement in higher education. *PLoS One, 14*(5), Article e0216865.

Bandura, A. (1994). Self-efficacy. In V. S. Ramachaudran (Ed.)*, Vol. 4. Encyclopedia of human behavior* (pp. 71–81). New York: Academic Press (Reprinted in H. Friedman [Ed.], Encyclopedia of mental health. San Diego: Academic Press, 1998).

Bandura, A. (2004). Social cognitive theory for personal and social change by enabling media. In A. Singhal, M. J. Cody, E. M. Rogers, & M. Sabido (Eds.), *Entertainment-education and social change: History, research, and practice* (pp. 75–96). Lawrence Erlbaum Associates Publishers.

Bewersdorff, A., Zhai, X., Roberts, J., & Nerdel, C. (2023). Myths, mis-and preconceptions of artificial intelligence: A review of the literature. *Computers & Education: Artificial Intelligence, 4*, Article 100143.

Bishop, C. M., & Bishop, H. (2024). *Deep learning: Foundations and concepts*. Cham: Springer. https://doi.org/10.1007/978-3-031-45468-4

Blut, M., & Wang, C. (2020). Technology readiness: a meta-analysis of conceptualizations of the construct and its impact on technology usage. *Journal of the Academy of Marketing Science, 48*, 649–669.

Braithwaite, J. J. (2007). *Occam's razor: The principle of parsimony*. UK: Behavioural Brain Sciences Centre, School of Psychology, University of Birmingham.

Carmichael, C., Callingham, R., Hay, I., & Watson, J. (2010). Statistical literacy in the middle school: The relationship between interest, self-efficacy, and prior mathematics achievement. *Australian Journal of Educational and Developmental Psychology, 10*, 83–93.

Carolus, A., Koch, M. J., Straka, S., Latoschik, M. E., & Wienrich, C. (2023). MAILS-Meta AI literacy scale: Development and testing of an AI literacy questionnaire based on well-founded competency models and psychological change-and meta-competencies. *Computers in Human Behavior: Artificial Humans, 1*(2), Article 100014.

Chai, C. S., Lin, P.-Y., Jong, M. S.-Y., Dai, Y., Chiu, T. K. F., & Qin, J. (2021). Perceptions of and behavioral intentions towards learning artificial intelligence in primary school students. *Educational Technology & Society, 24*(3), 89–101 (International Forum of Educational Technology & Society).

Chen, S. Y., Su, Y. S., Ku, Y. Y., Lai, C. F., & Hsiao, K. L. (2024). Exploring the factors of students' intention to participate in AI software development. *Library Hi Tech, 42*(2), 392–408.

Cheryan, S., Ziegler, S. A., Montoya, A. K., & Jiang, L. (2017). Why are some STEM fields more gender balanced than others? *Psychological Bulletin, 143*(1), 1.

Dunning, D. (2011). The Dunning–Kruger effect: On being ignorant of one's own ignorance. *Advances in Experimental Social Psychology, 44*, 247–296 (Academic Press).

Dziak, J. J., Coffman, D. L., Lanza, S. T., Li, R., & Jermiin, L. S. (2020). Sensitivity and specificity of information criteria. *Briefings in Bioinformatics, 21*(2), 553–565.

Eagle, R., Lander, R., & Hall, P. D. (2021). Questioning 'what makes us human': How audiences react to an artificial intelligence–driven show. *Cognitive Computation and Systems, 3*(2), 91–99.

Eccles, J. S., & Wigfield, A. (2002). Motivational beliefs, values, and goals. *Annual Review of Psychology, 53*(1), 109–132. https://doi.org/10.1146/annurev.psych.53.100901.135153

Edwards, A. (2015). *Big data, predictive machines and security: Enthusiasts, critics and sceptics'*. Discover Society. Online publication at https://core.ac.uk/reader/444014720.

Embretson, S. E., & Reise, S. P. (2013). *Item response theory*. Psychology Press.

Fraley, C., Raftery, A. E., Scrucca, L., Murphy, T. B., & Fop, M. (2024). mclust: Gaussian mixture modelling for model-based clustering, classification, and density estimation. *The Comprehensive R Archive Network (CRAN)* (Version 6.1.1) [R package] https://mclust-org.github.io/mclust/. h ttps://doi.org/10.32614/CRAN.package.mclust.

Gainor, K. A., & Lent, R. W. (1998). Social cognitive expectations and racial identity attitudes in predicting the math choice intentions of Black college students. *Journal of Counseling Psychology, 45*(4), 403–413.

Getenet, S., Cantle, R., Redmond, P., et al. (2024). Students' digital technology attitude, literacy and self-efficacy and their effect on online learning engagement. *Int J Educ Technol High Educ, 21*, 3. https://doi.org/10.1186/s41239-023-00437-y

Ghotbi, N., Ho, M. T., & Mantello, P. (2022). *Attitude of college students towards ethical issues of artificial intelligence in an international university in Japan*. AI & SOCIETY.

Google Trends. (2024). *Artificial intelligence*. Retrieved from https://trends.google.com/trends/explore?date=today%205-y&q=artificial%20intelligence&hl=en.

Gupta, S. C., & Kapoor, V. K. (2020). *Fundamentals of mathematical statistics*. Sultan Chand & Sons.

Haddock, A., Ward, N., Yu, R., & O'Dea, N. (2022). Positive effects of digital technology use by adolescents: A scoping review of the literature. *International Journal of Environmental Research and Public Health, 19*(21), Article 14009. https://doi.org/10.3390/ijerph192114009

Harackiewicz, J. M., Smith, J. L., & Priniski, S. J. (2016). Interest matters: The importance of promoting interest in education. *Policy insights from the behavioral and brain sciences, 3*(2), 220–227. https://doi.org/10.1177/2372732216655542

Hidi, S., & Renninger, K. A. (2006). The four-phase model of interest development. *Educational Psychologist, 41*(2), 111–127. https://doi.org/10.1207/s15326985ep4102_4

Hong, J.-W. (2022). I was born to love AI: The influence of social status on AI self-efficacy and intentions to use AI. *International Journal of Communication, 16*, 172–191.

Hornberger, M., Bewersdorff, A., & Nerdel, C. (2023). What do university students know about artificial intelligence? Development and validation of an AI literacy test. *Computers & Education: Artificial Intelligence, 5*, Article 100165.

Hornberger, M., Bewersdorff, A., Schiff, D., & Nerdel, C. (2024). A multinational assessment of AI literacy among university students in Germany, the UK, and the US. *Preprint*. https://doi.org/10.13140/RG.2.2.29184.65281

Hu, K. (2023). ChatGPT sets record for fastest-growing user base - analyst note. *Reuters*. https://www.reuters.com/technology/chatgpt-sets-record-fastest-growing-user-base-analyst-note-2023-02-01/.

Hu, L. T., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling: A Multidisciplinary Journal, 6*(1), 1–55.

Igolkina, A. A., & Meshcheryakov, G. (2020). semopy: A python package for structural equation modeling. *Structural Equation Modeling: A Multidisciplinary Journal, 27*(6), 952–963.

Johnson, R. D. (2005). An empirical investigation of sources of application-specific computer-self-efficacy and mediators of the efficacy—performance relationship. *International Journal of Human-Computer Studies, 62*(6), 737–758.

Kerschner, C., & Ehlers, M. H. (2016). A framework of attitudes towards technology in theory and practice. *Ecological Economics, 126*, 139–151.

Kirschner, P. A., & De Bruyckere, P. (2017). The myths of the digital native and the multitasker. *Teaching and Teacher Education, 67*, 135–142.

Koch, M. J., Wienrich, C., Straka, S., Latoschik, M. E., & Carolus, A. (2024). Overview and confirmatory and exploratory factor analysis of AI literacy scale. *Computers and Education: Artificial Intelligence, 7*, Article 100310.

Lau, S., & Roeser, R. W. (2002). Cognitive abilities and motivational processes in high school students' situational engagement and achievement in science. *Educational Assessment, 8*(2), 139–162.

Lindner, A., & Berges, M. (2020). Can you explain AI to me? Teachers' pre-concepts about artificial intelligence. In *2020 IEEE frontiers in education conference (FIE)* (pp. 1–9). IEEE. https://doi.org/10.1109/FIE44824.2020.9274136.

Lleras, C. (2005). Path analysis. *Encyclopedia of social measurement, 3*(1), 25–30.

Long, D., & Magerko, B. (2020). *What is AI literacy? Competencies and design considerations*. CHI.

Mang, J., Ustjanzew, N., Leßke, I., Schiepe-Tiska, A., & Reiss, K. (2019). *PISA 2015 Skalenhandbuch: Dokumentation der Erhebungsinstrumente*. Waxmann Verlag.

Markus, A., Pfister, J., Carolus, A., Hotho, A., & Wienrich, C. (2024). Effects of AI understanding-training on AI literacy, usage, self-determined interactions, and anthropomorphization with voice assistants. *Computers and Education Open, 6*, Article 100176.

Maslej, N., Fattorini, L., Perrault, R., Parli, V., Reuel, A., Brynjolfsson, E., … Clark, J. (2024). *The AI Index 2024 Annual Report. AI Index Steering Committee, Institute for Human-centered AI*. Stanford University.

McDonald, R. P. (1996). Path analysis with composite variables. *Multivariate Behavioral Research, 31*(2), 239–270. https://doi.org/10.1207/s15327906mbr3102_5

Mercader, C., & Gairín, J. (2020). University teachers' perception of barriers to the use of digital technologies: The importance of the academic discipline. *International Journal of Educational Technology in Higher Education, 17*(1), 4.

Mertala, P., Fagerlund, J., & Calderon, O. (2022). Finnish 5th and 6th grade students' pre-instructional conceptions of artificial intelligence (AI) and their implications for AI literacy education. *Computers & Education: Artificial Intelligence, 3*, Article 100095.

Mitchell, M. (1993). Situational interest: Its multifaceted structure in the secondary school mathematics classroom. *Journal of Educational Psychology, 85*(3), 424–436. https://doi.org/10.1037/0022-0663.85.3.424

Ng, D. T. K., et al. (2021). Conceptualizing AI literacy: An exploratory review. *Computers & Education: Artificial Intelligence, 2*, Article 100041.

Nielsen, T. A. S., & Haustein, S. (2018). On sceptics and enthusiasts: What are the expectations towards self-driving cars? *Transport Policy, 66*, 49–55.

Nuutila, K., Tapola, A., Tuominen, H., Kupiainen, S., Pásztor, A., & Niemivirta, M. (2020). Reciprocal predictions between interest, self-efficacy, and performance during a task. *Frontiers in Education, 5*. https://doi.org/10.3389/feduc.2020.00036

Oh, C., Lee, T., Kim, Y., Park, S., Kwon, S., & Suh, B. (2017). Us vs. them: Understanding artificial intelligence technophobia over the google deepmind challenge match. In *Proceedings of the 2017 CHI conference on human factors in computing systems* (pp. 2523–2534).

Pajares, F. (1997). Current directions in self-efficacy research. In M. Maehr, & P. R. Pintrich (Eds.), *Advances in motivation and achievement* (Vol. 10, pp. 1–49). Greenwich, CT: JAI Press.

Palmer, D. (2004). Situational interest and the attitudes towards science of primary teacher education students. *International Journal of Science Education, 26*(7), 895–908.

Prior, D. D., Mazanov, J., Meacheam, D., Heaslip, G., & Hanson, J. (2016). Attitude, digital literacy and self efficacy: Flow-on effects for online learning behavior. *The Internet and Higher Education, 29*, 91–97. https://doi.org/10.1016/j.iheduc.2016.01.001

Renninger, K. A. (2010). Working with and cultivating interest, self-efficacy, and self-regulation. In D. Preiss, & R. Sternberg (Eds.), *Innovations in educational psychology: Perspectives on learning, teaching and human development* (pp. 158–195). Springer. https://works.swarthmore.edu/fac-education/49.

Renninger, K. A., & Hidi, S. (2002). Student interest and achievement: Developmental issues raised by a case study. In A. Wigfield, & J. S. Eccles (Eds.), *Development of achievement motivation* (pp. 173–195). Academic Press. https://doi.org/10.1016/B978-012750053-9/50009-7.

Renninger, K. A., & Hidi, S. (2011). Revisiting the conceptualization, measurement, and generation of interest. *Educational Psychologist, 46*(3), 168–184. https://doi.org/10.1080/00461520.2011.587723

Renninger, K. A., & Shumar, W. (2002). Community building with and forteachers: The Math Forum as a resource for teacher professional development. In K. A. Renninger, & W. Shumar (Eds.), *Building virtual communities: Learning and change in cyberspace* (pp. 60–95). New York: Cambridge University Press.

Schepman, A., & Rodway, P. (2020). Initial validation of the general attitudes towards artificial intelligence scale. *Computers in Human Behavior Reports, 1*, Article 100014.

Schunk, D. H., & DiBenedetto, M. K. (2021). Self-efficacy and human motivation. *Advances in Motivation Science, 8*, 153–179.

Stöhr, C., Ou, A. W., & Malmström, H. (2024). Perceptions and usage of AI chatbots among students in higher education across genders, academic levels and fields of study. *Computers & Education: Artificial Intelligence, 7*, Article 100259.

Stage, F. K., Carter, H. C., & Nora, A. (2004). Path analysis: An introduction and analysis of a decade of research. *The Journal of Educational Research, 98*(1), 5–13. https://doi.org/10.3200/JOER.98.1.5-13

Stanford, H. A. I. (2024). *Global AI power rankings: Stanford HAI tool ranks 36 countries in AI*. Stanford HAI. Retrieved from https://hai.stanford.edu/news/global-ai-power-rankings-stanford-hai-tool-ranks-36-countries-ai.

Tabachnick, B. G., & Fidell, L. S. (2011). Multivariate analysis of variance (MANOVA). In M. Lovric (Ed.), *International encyclopedia of statistical science*. Berlin, Heidelberg: Springer. https://doi.org/10.1007/978-3-642-04898-2_394.

Tavakol, M., & Dennick, R. (2011). Making sense of Cronbach's alpha. *International Journal of Medical Education, 2*, 53–55. https://doi.org/10.5116/ijme.4dfb.8dfd

Teng, M., Singla, R., Yau, O., Lamoureux, D., Gupta, A., Hu, Z., … Field, T. S. (2022). Health care students' perspectives on artificial intelligence: Countrywide survey in Canada. *JMIR medical education, 8*(1), Article e33390.

Ullman, J. B., & Bentler, P. M. (2012). Structural equation modeling. *Handbook of psychology* (2nd ed., 2).

Usher, E. L., & Pajares, F. (2008). Sources of self-efficacy in school: Critical review of the literature and future directions. *Review of Educational Research, 78*, 751–796.

Verdugo-Castro, S., García-Holgado, A., & Sánchez-Gómez, M. C. (2022). The gender gap in higher stem studies: A systematic literature review. *Heliyon, 8*(8), Article e10300.

Vogel, T., & Wanke, M. (2016). *Attitudes and attitude change* (2nd ed.). Psychology Press. https://doi.org/10.4324/9781315754185

Wang, Y. Y., & Chuang, Y. W. (2024). Artificial intelligence self-efficacy: Scale development and validation. *Education and Information Technologies, 29*(4), 4785–4808. https://doi.org/10.1007/s10639-023-12015-w

Wardenaar, K. J. (2024). *Latent Profile Analysis in R: A Tutorial and Comparison to Mplus (Version 1.2)*. Groningen: University Medical Center.