



Prompt Literacy as a Core Workplace Competency: Development of the Prompt Literacy Index (PLI-20) for Non-Technical Professionals

Shivanand R. Koppalkar, Ajit A. Shenvi

Researcher(s), Walsh College, United States

Abstract

Organizations are adopting generative artificial intelligence faster than their employees can use it well. Training investment targets tools rather than the human skill that determines output quality. This article defines prompt literacy as a distinct, measurable workplace competency: the ability of a worker to instruct, contextualize, verify, refine, and ethically govern the output of a generative AI system during ordinary job tasks. Drawing on digital literacy theory, AI literacy research, and recent evidence on prompting behavior, the article develops a five-dimension conceptual framework covering task framing, context supply, output verification, iterative refinement, and ethical use. Following established scale development procedures, a 40-item pool was generated from peer-reviewed literature published between 2004 and 2025 and reduced through a structured content-validity design to the 20-item Prompt Literacy Index (PLI-20). The full instrument is presented, together with a registered validation protocol specifying the sample, exploratory factor analysis, and reliability criteria for the planned pilot study with non-technical professionals. The article contributes the first profession-agnostic framework and instrument for measuring applied prompting skill outside technical roles, shifting the unit of analysis in workplace AI research from the tool to the worker. Implications for workforce development, training design, and future confirmatory validation are discussed.

Keywords: prompt literacy, AI literacy, generative artificial intelligence, workplace competency, scale development, digital skills

1. Introduction

Generative artificial intelligence has moved from novelty to standard workplace software in only a few years. Large language models now sit inside word processors, email clients, customer service platforms, and search tools. Scholars across disciplines agree that this shift raises both opportunity and risk for organizations, and that the benefits are not automatic (Dwivedi et al., 2023). A central reason is simple. The quality of what a generative model produces depends heavily on the quality of what the user puts in. The same tool, in different hands, returns very different results. Controlled studies confirm that people without technical training struggle systematically to instruct these systems well (Zamfirescu-Pereira et al., 2023). The weak link in workplace AI adoption is often not the model. It is the individual user.

Research has not yet caught up with this problem at the level of the ordinary worker. Two adjacent literatures come close but stop short. The first, AI literacy, defines the broad competencies people need to understand, use, and critically evaluate artificial intelligence (Long & Magerko, 2020; Ng et al., 2021). Its instruments measure general knowledge and attitudes, not the applied skill of instructing a generative model during a work task. The second, prompt engineering, studies how to write effective instructions for large language



International Journal of Science, Technology and Management (IJSTM)

Volume 1 | Issue 1 | 2026 | Article ID: 8622-8235-6697 | <https://doi.org/10.64823/ijstm.2601003>

IORO Publications · Texas, USA · www.ioro.org

models. Its findings are valuable, but the field centers on technical users, system builders, and educational settings (Knoth et al., 2024). Neither literature offers a validated, profession-agnostic instrument that measures how well a human resources officer, a nurse manager, a teacher, or an accountant actually works with generative AI. This article addresses that gap.

The purpose of this study is twofold. First, it defines prompt literacy as a distinct workplace competency and grounds it in a five-dimension conceptual framework: task framing, context supply, output verification, iterative refinement, and ethical use. Second, it develops the Prompt Literacy Index (PLI-20), a 20-item self-report instrument built through established scale development procedures (Boateng et al., 2018; DeVellis & Thorpe, 2022). The full item set is presented here, together with a registered protocol for its empirical validation. Presenting the instrument and the protocol before data collection keeps the validation transparent and invites independent replication.

Three research questions guide the work:

1. **RQ1.** What dimensions constitute prompt literacy as a workplace competency for non-technical professionals?
2. **RQ2.** Can prompt literacy be measured with a short, content-valid self-report instrument?
3. **RQ3.** What protocol will establish the factor structure and reliability of that instrument across occupational groups?

The practical stakes are high. Employers currently buy AI tools and licenses without a way to measure whether their people can use them; a validated index lets training target the weakest skill dimension rather than the tool in general. The article proceeds as follows. Section 2 reviews the lineage from digital literacy to AI literacy, the prompting literature, the productivity evidence, and existing instruments. Section 3 presents the conceptual framework. Section 4 reports the development of the PLI-20. Section 5 specifies the validation protocol. Sections 6 through 8 discuss contributions, limitations, and conclusions.

2. Literature Review

2.1.1. *From Digital Literacy to AI Literacy*

Every major wave of workplace technology has eventually produced a named, teachable, and measurable human skill. Early internet adoption produced digital literacy. Eshet-Alkalai (2004) argued that surviving in a digital environment requires more than operating software; it requires distinct cognitive skills, including photo-visual thinking, information evaluation, and branching navigation. His framework established a principle that still holds: a new technology creates a new literacy, and that literacy can be defined and assessed. A decade later, van Laar et al. (2017) systematically reviewed the relation between 21st-century skills and digital skills. They showed that the two overlap but are not identical, and that core components such as information management, critical evaluation, and problem solving must be anchored to the specific technology in use. The lesson from this lineage is that literacies are technology-contingent. When the technology changes, the skill set must be redefined and re-measured.

Artificial intelligence triggered exactly this redefinition. Long and Magerko (2020) defined AI literacy as a set of competencies that enable individuals to critically evaluate AI technologies, communicate and collaborate effectively with AI, and use AI as a tool at home and in the workplace. Their design-oriented framework identified seventeen competencies, from recognizing AI to understanding its ethical limits. Ng et al. (2021), reviewing the emerging field, organized AI literacy into four aspects: knowing and understanding AI, using

and applying AI, evaluating and creating AI, and addressing AI ethics. Both frameworks predate the mass adoption of conversational generative models. They treat AI mainly as something to understand and judge, not as something the worker instructs in natural language dozens of times a day. Generative AI changed the basic interaction primitive. The user no longer only operates an interface; the user writes instructions, supplies context, and negotiates with the system over successive turns. The literacy cycle that Eshet-Alkalai began for the internet, and that Long and Magerko extended to AI in general, has not yet been completed for this new mode of work. That incomplete step is where prompt literacy sits.

2.1.2. Prompt Engineering and Its Limits as a Lens

A second literature examines the instructions themselves. Prompt engineering emerged as the practice of writing precise, well-structured inputs to elicit useful outputs from large language models. Evidence shows the practice is genuinely difficult for ordinary users. Zamfirescu-Pereira et al. (2023) observed non-AI experts attempting to design prompts for a chatbot task. Participants worked opportunistically rather than systematically, overgeneralized from single successes or failures, and borrowed expectations from human-to-human instruction that the model did not share. The authors concluded that non-experts fail in patterned, predictable ways. Those patterns matter for the present study because they are behavioral: they describe what unskilled prompting looks like, which in turn suggests what skilled prompting consists of.

Knoth et al. (2024) connected prompting behavior to the AI literacy construct. Studying students interacting with a large language model, they introduced a skill-based view of prompt engineering and found that higher-quality prompts predicted higher-quality model output, and that certain facets of AI literacy supported better prompting. Their conclusion, that prompt engineering is a required skill for goal-directed use of generative tools, directly motivates measurement. Yet the prompting literature as a whole remains tool-centric and setting-specific. Its studies concentrate on technical users, students, and laboratory tasks. It has produced techniques and taxonomies, but not a validated instrument that a mid-sized employer could administer to a non-technical workforce. The lens is valuable; the measurement is missing.

2.1.3. Generative AI and Worker Productivity

If prompting skill did not matter economically, measuring it would be an academic exercise. The evidence says otherwise. Noy and Zhang (2023) ran a preregistered experiment in which college-educated professionals completed occupation-specific writing tasks, half with access to ChatGPT. Access raised productivity substantially: time taken fell and graded quality rose. Notably, the gains were largest for participants whose initial performance was weakest, compressing the gap between workers. Brynjolfsson et al. (2025) extended this evidence to the field. Studying thousands of customer support agents given access to a generative AI assistant, they found average productivity gains of about fifteen percent, with the largest improvements among novice and lower-skilled agents, who appeared to absorb the embedded knowledge of more experienced colleagues.

Two implications follow for this study. First, generative AI delivers real returns, which justifies organizational attention to the human side of adoption. Second, the returns are conditional. In both studies, workers had to elect to use the tool and to use it competently; the technology amplified skill rather than replacing it. The field evidence measures outcomes, not the user competence that produces them. A worker-level measure of that competence is the missing variable between tool access and realized productivity. That is the variable the PLI-20 is designed to capture.

2.1.4. Existing AI Literacy Instruments and the Gap

Several validated AI literacy scales now exist, and each is a genuine contribution. None, however, measures applied prompting skill in non-technical work. Carolus et al. (2023) developed the Meta AI Literacy Scale (MAILS), a 34-item modular instrument covering the facets of using and applying AI, understanding AI, detecting AI, and AI ethics, plus the ability to create AI and psychological meta-competencies such as AI self-efficacy. Wang et al. (2023) validated an artificial intelligence literacy scale organized around awareness, usage, evaluation, and ethics, designed for general users. Laupichler et al. (2023) developed an item set specifically for non-experts, focused on conceptual understanding of how AI systems work. Table 1 compares these instruments against the construct targeted here.

Table 1. Published AI Literacy Instruments and Their Coverage of Applied Prompting Skill

Instrument	Source	Items	What it measures	Why it does not capture prompt literacy
MAILS	Carolus et al. (2023)	34	AI literacy facets plus self-efficacy and self-management	General competencies and attitudes; no items on instructing a generative model during work tasks
AI Literacy Scale	Wang et al. (2023)	12	Awareness, usage, evaluation, and ethics of AI	Measures broad usage competence, not the behaviors of framing, contextualizing, and refining prompts
Non-experts' AI literacy scale	Laupichler et al. (2023)	31	Conceptual understanding of AI by non-experts	Targets knowledge about AI systems, not applied interaction skill

The gap can now be stated precisely. Existing instruments measure what people know and feel about artificial intelligence. The productivity literature shows that what matters economically is what people do with generative systems during real tasks. The prompting literature describes those behaviors but has not operationalized them into a validated, profession-agnostic measure. No published instrument measures the day-to-day skill of a non-technical professional instructing, verifying, and governing generative AI output at work. The present study constructs that instrument.

2.2. Conceptual Framework

This study defines prompt literacy as *the ability of a worker to instruct, contextualize, verify, refine, and ethically govern the output of a generative AI system in the course of ordinary job tasks*. The definition is deliberately behavioral. It describes what skilled users do, not what they know about artificial intelligence in the abstract. It is also deliberately interface-agnostic. The five abilities apply whether the system is a chat window, an assistant embedded in office software, or a future voice or agent interface, because every such interaction still requires a stated goal, supplied context, checked output, improvement over turns, and responsible handling.

The framework places prompt literacy in a clear relation to its parent constructs. It is a subset of AI literacy in the sense of Ng et al. (2021): it operationalizes the “using and applying” and “evaluating” aspects for generative systems specifically. It is narrower than prompt engineering as practiced by system builders, because it excludes model configuration, tooling, and programmatic techniques, and it is broader than any single prompting technique, because it includes verification and ethics, which technique taxonomies omit. Each dimension is grounded in the reviewed literature.

Task framing is the ability to state the goal, audience, format, and constraints of a task before or while requesting output. Non-experts commonly issue underspecified requests and then attribute the poor result

to the tool (Zamfirescu-Pereira et al., 2023). Prompt quality, of which specification is a core component, predicts output quality (Knoth et al., 2024).

Context supply is the ability to give the system the facts, source material, and boundaries the task requires. Generative models produce fluent text with or without adequate grounding; the user decides whether the model works from the right material. Supplying the policy excerpt an answer must follow, or the data a summary must reflect, is a distinct and observable behavior.

Output verification is the ability to check generated content against trusted sources before it is used. Generative systems can produce confident but false statements, a risk raised consistently across disciplines (Dwivedi et al., 2023). Verification corresponds to the evaluative aspect of AI literacy (Long & Magerko, 2020; Ng et al., 2021) applied at the level of a single work product.

Iterative refinement is the ability to treat the first output as a draft and improve it through follow-up instruction. Observed novice behavior is the opposite: abandoning the attempt, or starting over, after one weak answer (Zamfirescu-Pereira et al., 2023). Skilled users diagnose what was wrong and direct the correction.

Ethical use is the ability to respect confidentiality, disclosure, and organizational rules while using generative tools. Ethics is a recognized facet of AI literacy (Carolus et al., 2023; Ng et al., 2021). At the workplace level it becomes concrete: not pasting client data into external tools, disclosing substantial AI assistance where disclosure is expected, and following organizational AI policy.

Together the five dimensions span the life cycle of a single assisted task, from the instruction, through the material, to the checking, the improvement, and the governance of the result. Each dimension is behavioral, teachable, and itemizable, which makes the framework suitable for instrument construction.

2.3. Development of the Prompt Literacy Index

Instrument development followed the staged best practices synthesized by Boateng et al. (2018) and the scale development procedures of DeVellis and Thorpe (2022): domain definition, item generation, content validity assessment, and preparation for psychometric testing.

2.3.1. Item Generation

The domain was defined by the five-dimension framework in Section 3. Candidate items were generated deductively from the reviewed literature on prompting behavior, AI literacy, and digital skills published between 2004 and 2025, with particular weight on the behavioral failure patterns documented for non-experts (Zamfirescu-Pereira et al., 2023) and the skill components linked to output quality (Knoth et al., 2024). Forty candidate items were drafted, eight per dimension. Four wording rules governed drafting. Each item expresses one idea. Each item uses plain English with no technical vocabulary. Each item stays under twenty words. Each item describes a behavior the respondent can observe in their own work, not an attitude toward AI in general. Items are answered on a 5-point frequency scale (1 = *never*, 5 = *always*), because the construct concerns habitual behavior rather than agreement with propositions.

2.3.2. Content Validity Design

Content validity is assessed by an expert panel of six to eight judges spanning three perspectives: researchers in human-AI interaction or AI literacy, practitioners in corporate learning and development or human resources, and at least one specialist in technology ethics. Each judge rates every candidate item for relevance to its intended dimension on a 4-point scale. Item-level content validity indices (I-CVI) and the average scale-level index (S-CVI/Ave) are computed, with retention thresholds of I-CVI $\geq$.78 and S-CVI/Ave

$\geq .90$, consistent with the standards reported by Boateng et al. (2018). Judges also flag ambiguity, double-barreled wording, and dimension overlap. The two strongest surviving items per dimension are retained outright; the next strongest are retained where needed to reach four items per dimension, producing the 20-item index presented in Table 2.

2.3.3. The PLI-20 Instrument

Table 2. *The Prompt Literacy Index (PLI-20). Response scale: 1 = never, 2 = rarely, 3 = sometimes, 4 = often, 5 = always. Stem: "When I use an AI tool for a work task..."*

Code	Dimension	Item
TF1	Task framing	I state what the output will be used for.
TF2	Task framing	I tell the AI tool who the audience for the output is.
TF3	Task framing	I specify the format I need, such as a table, a list, or an email.
TF4	Task framing	I state limits up front, such as length, tone, or scope.
CS1	Context supply	I give the AI tool the background information the task needs.
CS2	Context supply	I paste or attach the source material the answer must be based on.
CS3	Context supply	I tell the AI tool which rules or policies the output must follow.
CS4	Context supply	I say what the AI tool should leave out as well as what it should include.
OV1	Output verification	I check AI-generated facts against another source before using them.
OV2	Output verification	I read AI output fully before passing it on to anyone else.
OV3	Output verification	When the AI tool cites a source, I confirm the source exists and says what is claimed.
OV4	Output verification	I treat AI output as a draft that I remain responsible for.
IR1	Iterative refinement	When an answer is weak, I refine my request instead of giving up.
IR2	Iterative refinement	I ask follow-up questions to improve the first answer.
IR3	Iterative refinement	I point out what was wrong in the output and ask for a corrected version.
IR4	Iterative refinement	I compare more than one response before choosing what to use.
EU1	Ethical use	I avoid entering confidential information into AI tools.
EU2	Ethical use	I remove names and personal details from prompts where they are not needed.
EU3	Ethical use	I disclose substantial AI assistance in my work where disclosure is expected.
EU4	Ethical use	I follow my organization's rules on which AI tools may be used for which tasks.

Scoring is the mean of the items within each dimension, yielding a five-dimension profile, plus an overall mean. The profile form is deliberate: for training purposes, knowing that a team scores high on refinement but low on verification is more useful than a single total.

2.4. Proposed Validation Protocol

This section registers the design of the planned pilot study so that the subsequent empirical report can be judged against a protocol fixed in advance.

Participants. The target sample is 180 working professionals in non-technical roles, recruited through professional associations and a commercial research panel. Five occupational groups are targeted in roughly equal numbers: human resources, education, healthcare administration, finance and accounting, and public administration. A sample of 180 provides approximately nine respondents per item, within the commonly recommended range for exploratory factor analysis of a 20-item instrument (Boateng et al., 2018). Screening

questions exclude respondents whose primary role is in information technology, software development, or data science, and respondents who have not used a generative AI tool for work in the past month.

Procedure. Participants complete the PLI-20 online, followed by demographic items (occupation, tenure, frequency of AI use) and two single-item criterion measures: self-rated confidence with generative AI tools and self-reported usefulness of AI output in their work. Institutional ethics approval is obtained before data collection. Participation is voluntary, consent is informed, responses are anonymous, and data are stored in compliance with the UK GDPR.

Analysis plan. Analysis proceeds in four fixed steps.

1. *Item screening.* Item means, standard deviations, inter-item correlations, and floor and ceiling effects are examined; items with near-zero variance are flagged.
2. *Exploratory factor analysis.* Principal axis factoring with oblique rotation is applied, given expected correlations between dimensions. Sampling adequacy requires $KMO \geq .80$ and a significant Bartlett's test. The number of factors is decided by parallel analysis rather than eigenvalue cutoffs alone. The hypothesis is a five-factor structure matching Section 3; a different recovered structure will be reported, not forced.
3. *Reliability.* Cronbach's alpha and McDonald's omega are computed per subscale, with .70 as the minimum acceptable value.
4. *Refinement.* Items with loadings below .40, or substantial cross-loadings, are candidates for removal, with the constraint that each dimension retains at least three items.

Planned extensions. A second, independent sample of at least 300 respondents is planned for confirmatory factor analysis and tests of measurement invariance across the five occupational groups, which addresses RQ3 fully. Convergent and discriminant validity will be examined against an established AI literacy scale (Carolus et al., 2023; Wang et al., 2023), with the expectation of positive but non-redundant correlations.

3. Discussion

This article set out to name, define, and operationalize the human skill on which workplace returns from generative AI depend. Its contributions are theoretical, methodological, and practical.

Theoretically, the article completes a literacy cycle that previous technology waves completed before it. Digital literacy named the skills the internet demanded (Eshet-Alkalai, 2004; van Laar et al., 2017). AI literacy named the competencies required to understand and evaluate artificial intelligence (Long & Magerko, 2020; Ng et al., 2021). Prompt literacy, as defined here, names the applied interaction skill that conversational generative systems demand of every worker who uses them. Positioning the construct as the behavioral, generative-AI-specific subset of AI literacy keeps the field coherent: prompt literacy does not compete with AI literacy, it operationalizes its application layer for the dominant current interface.

Methodologically, the article contributes a complete, transparent instrument rather than a described one. The full PLI-20, its wording rules, its content-validity thresholds, and its analysis plan are all published in advance of data, which allows independent researchers to critique the items, administer the index in their own settings, and compare factor solutions. Validated instruments, once available, tend to structure a research area; the comparison in Table 1 shows that generative-AI-specific workplace measurement currently has no such structure to build on.

Practically, the index gives organizations a measure they currently lack. Employers can establish an AI-readiness baseline before a tool rollout and target training at the weakest dimension rather than delivering generic tool demonstrations. The dimension profile matters here. The productivity evidence shows that gains flow to workers who use the tools competently (Brynjolfsson et al., 2025; Noy & Zhang, 2023), while the behavioral evidence predicts that verification and disciplined framing are where non-experts fail first (Zamfirescu-Pereira et al., 2023). An organization whose staff score high on refinement but low on verification faces a quality and compliance risk that a single total score would hide. For educators, the five dimensions supply a rubric for curriculum design in professional and higher education, each dimension mapping to teachable, assessable behaviors.

A final point concerns durability. Interfaces for generative systems are changing quickly, toward agents, voice, and automatic prompt improvement. This might seem to threaten any prompting-related construct. The framework was built to survive that change. Whatever the interface, a worker must still convey the goal, judge whether the system has the right material, decide whether to trust the output, improve an unsatisfactory result, and handle the interaction within ethical and organizational bounds. The surface behaviors in individual items may need periodic re-anchoring; the five dimensions themselves are interface-agnostic.

3.1. Limitations and Future Research

Four limitations bound the present work. First, the PLI-20 is a self-report instrument; it measures reported habitual behavior, which can diverge from demonstrated performance. Pairing the index with task-based assessments, in which respondents complete standardized AI-assisted tasks that are independently scored, is a priority for future work. Second, the empirical validation reported here is a protocol, not a result; until the pilot and confirmatory studies are complete, the factor structure and reliability remain hypotheses. Third, the initial validation targets English-speaking professionals in five occupational groups; translation and cross-cultural validation are required before wider use. Fourth, the construct is anchored to the current generation of generative tools, and a scheduled re-validation cycle should test whether item wording still reflects practice as interfaces evolve.

4. Conclusion

The workplace value of generative artificial intelligence depends on a human skill that has, until now, lacked a precise name and a measure. This article defined that skill as prompt literacy, grounded it in five behavioral dimensions, and delivered the PLI-20 together with a registered protocol for its validation. Measuring the skill is the precondition for teaching it, and teaching it is how the documented productivity gains of generative AI reach the workers and organizations that adopt it.

References

- [1] Boateng, G. O., Neilands, T. B., Frongillo, E. A., Melgar-Quinonez, H. R., & Young, S. L. (2018). Best practices for developing and validating scales for health, social, and behavioral research: A primer. *Frontiers in Public Health*, 6, Article 149. <https://doi.org/10.3389/fpubh.2018.00149>
- [2] Brynjolfsson, E., Li, D., & Raymond, L. (2025). Generative AI at work. *The Quarterly Journal of Economics*, 140(2), 889–942. <https://doi.org/10.1093/qje/qjae044>
- [3] 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. <https://doi.org/10.1016/j.chbah.2023.100014>
- [4] DeVellis, R. F., & Thorpe, C. T. (2022). *Scale development: Theory and applications* (5th ed.). Sage.
- [5] Dwivedi, Y. K., Kshetri, N., Hughes, L., Slade, E. L., Jeyaraj, A., Kar, A. K., Baabdullah, A. M., Koohang, A., Raghavan, V., Ahuja, M., Albanna, H., Albashrawi, M. A., Al-Busaidi, A. S., Balakrishnan, J., Barlette, Y., Basu, S., Bose, I., Brooks, L., Buhalis, D., ... Wright, R. (2023). "So what if ChatGPT wrote it?" Multidisciplinary perspectives on opportunities, challenges and implications of generative conversational AI for research, practice and policy. *International Journal of Information Management*, 71, Article 102642. <https://doi.org/10.1016/j.ijinfomgt.2023.102642>
- [6] Eshet-Alkalai, Y. (2004). Digital literacy: A conceptual framework for survival skills in the digital era. *Journal of Educational Multimedia and Hypermedia*, 13(1), 93–106.
- [7] Knoth, N., Tolzin, A., Janson, A., & Leimeister, J. M. (2024). AI literacy and its implications for prompt engineering strategies. *Computers and Education: Artificial Intelligence*, 6, Article 100225. <https://doi.org/10.1016/j.caeai.2024.100225>
- [8] Laupichler, M. C., Aster, A., Haverkamp, N., & Raupach, T. (2023). Development of the "Scale for the assessment of non-experts' AI literacy" — An exploratory study. *Computers in Human Behavior Reports*, 12, Article 100338. <https://doi.org/10.1016/j.chbr.2023.100338>
- [9] Long, D., & Magerko, B. (2020). What is AI literacy? Competencies and design considerations. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems* (pp. 1–16). ACM. <https://doi.org/10.1145/3313831.3376727>
- [10] Ng, D. T. K., Leung, J. K. L., Chu, S. K. W., & Qiao, M. S. (2021). Conceptualizing AI literacy: An exploratory review. *Computers and Education: Artificial Intelligence*, 2, Article 100041. <https://doi.org/10.1016/j.caeai.2021.100041>
- [11] Noy, S., & Zhang, W. (2023). Experimental evidence on the productivity effects of generative artificial intelligence. *Science*, 381(6654), 187–192. <https://doi.org/10.1126/science.adh2586>
- [12] van Laar, E., van Deursen, A. J. A. M., van Dijk, J. A. G. M., & de Haan, J. (2017). The relation between 21st-century skills and digital skills: A systematic literature review. *Computers in Human Behavior*, 72, 577–588. <https://doi.org/10.1016/j.chb.2017.03.010>
- [13] Wang, B., Rau, P.-L. P., & Yuan, T. (2023). Measuring user competence in using artificial intelligence: Validity and reliability of the artificial intelligence literacy scale. *Behaviour & Information Technology*, 42(9), 1324–1337. <https://doi.org/10.1080/0144929X.2022.2072768>
- [14] Zamfirescu-Pereira, J. D., Wong, R. Y., Hartmann, B., & Yang, Q. (2023). Why Johnny can't prompt: How non-AI experts try (and fail) to design LLM prompts. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems* (Article 437, pp. 1–21). ACM. <https://doi.org/10.1145/3544548.3581388>