Journal of Computer Science and Technology Studies
ISSN: 2709-104X
DOI: 10.32996/jcsts
Journal Homepage: www.al-kindipublisher.com/index.php/jcsts

JCSTS
AL-KINDI CENTER FOR RESEARCH AND DEVELOPMENT

| RESEARCH ARTICLE

The Recognition Gap: How Women's Technical Abilities Remain Invisible in the AI Age

Dr. Kristy Rae Stewart
Program Chair, Technology, Herzing University, Milwaukee, United States
Corresponding Author: Dr. Kristy Rae Stewart, **E-mail**: kristewart@herzing.edu

| ABSTRACT

This paper examines the persistent gender gap in the tech industry, which stems from a lack of appreciation for skills, rather than technical aptitude. Using the principles of Social Cognitive Theory and the concept of the "Second Digital Divide," a meta-analytical review of 54 journal articles published between 2021 and 2025 across 15 countries was conducted. The results show that women score as well as men on technical proficiency tests, yet appreciation of identical tasks is reduced by 12% as soon as the appreciator recognizes the gender of the author. The case analysis of Artificial Intelligence (AI) orchestration for women-led startups in Southeast Asia reveals that, although they achieve 35% higher productivity with AI-integrated workflows, they receive 60% less institutional funding than male-led startups. Here, I propose the "Glass Wall" phenomenon, in which women's technical expertise is labeled "prompt dependency" and relegated to the domain of appreciation for architectural skills. The suggestions for change include implementing blind code review practices, measuring skill delivery progress asynchronously, and leveraging the unseen efforts of documentation and ethical auditing in promotion cycles.

| KEYWORDS

Gender; technology; gatekeeping; women in STEM

| ARTICLE INFORMATION

ACCEPTED: 12 January 2026 **PUBLISHED:** 15 February 2026 **DOI:** 10.32996/jcsts.2026.8.4.6

Introduction

An irreconcilable contradiction marks the state of technology in 2025. It appears, with the democratization of technological knowledge through Artificial Intelligence and open-source platforms, that technology would have achieved a level playing field by now. However, this is not true with respect to representation among engineers and architects in technology. According to current labor market statistics, women contribute only 26% to technical pools overall, and this figure drops significantly in high-stakes sectors such as cybersecurity and Artificial Intelligence infrastructure (Randstad, 2025).

The question this research seeks to answer is not whether women can acquire technological skills, but rather why technological capability and competence do not transform into an equal share of professional recognition and reward. In short, the question is not whether women can solve technology problems; it is rather why these skills and competence do not permit equal recognition and reward in the technology world.

In this article, it is argued that the existence of a technology gap is not necessarily an issue of lacking technical know-how, but rather an issue of lacking recognition of technical skills. I propose the concept of the "Second Digital Divide," which marks a shift in focus from hardware access toward the social construction of technical talent. This framework explores how institutional gatekeeping, through selective code review and exclusionary "bro cultures," actively restricts women from utilizing their skills. Although Generative AI can bypass gatekeeper processes, it also arguably cements current disparities in new language and definitions. AI is at the nucleus of any technical field, and its impact on the technical talent and expertise of women and men is now an essential area for promoting equity and innovation.

The relevance of this study, however, extends beyond the issue of equity in the workplace, which is important in its own right. The underutilization of women's technical capabilities is a significant case of economic inefficiency, as it is estimated that closing the gender gap in technology would add $12 trillion to global GDP by 2030 (McKinsey Global Institute, 2025). Further, the failure to leverage diverse viewpoints in technology development work has been shown to have clear, damaging consequences for product quality, user safety, and algorithmic bias. AI models trained almost exclusively by teams from a similar demographic background demonstrate clear biases against disadvantaged groups, from facial recognition algorithms that fail to accurately identify darker-skinned faces to algorithms that underestimate people of color's experiences of pain in healthcare (Buolamwini & Gebru, 2018; Obermeyer et al., 2019). In other words, closing the gender gap in technology is a matter not just of equity, but a precondition for creating technology in the interest of all humanity.

The skills gap is demonstrated in this research, which utilizes a meta-analysis of technical self-efficacy and objective performance across 15 countries, to illustrate that it is an ideological construct being used as a smokescreen for systemic bias. By analyzing technology hubs in the Global North alongside innovation hubs in the Global South, this study provides empirical proof for "The Proficiency Paradox." This paradox suggests that while technical proficiency is theoretically the primary metric for recognition, advancement, and compensation, its utility is mitigated for women in technology fields by systemic evaluative biases. Furthermore, by advancing the concept of "The Glass Wall," this research identifies a lateral barrier within which women's technical contributions are rendered invisible or categorized as "not technological work at all." This differs fundamentally from "The Glass Ceiling," which describes the vertical obstruction of mobility into executive levels; the Glass Wall instead represents a boundary that devalues and reclassifies technical labor, excluding women from the core of innovation.

This paper first reviews the history of literature regarding women and technology, tracing the transition from access-based initiatives to a modern focus on proficiency and agency. It then establishes a theoretical framework, utilizing Social Cognitive Theory and the Second Digital Divide, to interpret the construction and impact of self-efficacy. The paper then details the methodology, including the specific processes and cases used to conduct this systematic review and analysis. Finally, the results are presented in two parts: the first part comprises the data collected and discussed using a systematic method, and the second part presents the results interpreted from the case example used to focus on the identified issue, the Proficiency Paradox, and the final recommendations to rectify this issue.

Literature Review of the Evolution of Technical Self-Efficacy Research

From Access to Agency: Reconceptualizing the Digital Divide

The literature surrounding women in technology has undergone a significant transition in the last two decades. The initial work was carried out in the realm of "access," which comprised the availability of computers and internet access, along with education in tech-related areas (Warschauer, 2004). This can also be referred to as the "first-level digital divide." If women are not adequately represented in tech-related areas, the remedy would lie in providing more computers in schools, offering more scholarships for tech-related education, and making greater efforts to persuade girls to pursue tech-related education. However, this level of intervention was based on the tacit premise that access was the prerequisite for inclusion and success.

However, as access to technology increased, the gap has persisted, leading researchers to rethink the issue. More recent work on the second-level digital divide by Hargittai (2002) emphasized the importance of skills rather than access. Rather than individuals simply having access to computers and the internet, it was recognized that there was a huge variation in the extent to which individuals could use the technology to seek information, communicate, produce content, or solve problems. Studies since then have confirmed second-level gender differences in digital skills, specifically in terms of confidence and usage patterns, with higher values noted in men than in women, even within countries where access was controlled for in terms of computers and internet availability (Hargittai & Shafer, 2006).

Current scholarships today focus even more on issues of "proficiency and agency," recognizing that proficiency does not automatically translate into effective high-tech innovation and leadership (van Dijk, 2020). This third wave of scholarships focuses on the social, cultural, and institutional environments that influence skill development and recognition. Agency—the ability to achieve effective action within technology systems—then becomes the major variable. This helps explain why women possessing equal skill can fail to achieve success in their chosen field: the effect or outcome one can achieve matters not, but rather its recognition or reward.

The Second Digital Divide and the AI Age

Studies carried out in 2024 and 2025 identify the pivotal point where the hurdle is not technology hardware, but rather the stratification of usage (UNESCO, 2024). Men are more likely to engage in the "experimental" aspect of technology utilization, including the development of custom LLM wrappers, software contributions towards open-source repositories, or the creation of novel software solutions, while women are usually encouraged to perform functional or administrative tasks (BCG, 2024). This is perceived as an imbalance between innovation and is instead the result of contrasting levels of socialization between the two sexes in terms of taking risks in technological endeavors, rather than a reflection of divergent levels of aptitude. The gender divide in the allocation of technological endeavors is initiated at a young stage, as research indicates boys are shown increased encouragement towards exploratory technological endeavors, while girls are directed toward systematic and task-related software applications (Margolis & Fisher, 2002).

The rise of Generative AI introduces new layers to this inequality, but initial findings suggest that men are more likely to utilize experimental AI for creative and professional purposes, whereas women are more inclined to leverage productivity-related AI for communication purposes (BCG, 2024). Indeed, it has long been noted that a gender gap persists in computer usage, and it now appears that this gap will be replicated in relation to artificial intelligence usage, reflecting broader social patterns in which the Second Digital Divide has contributed to social inequality. Will artificial intelligence literacy be the new gateway to inequality, as it appears that individuals who leverage experimental artificial intelligence usage are poised to gain greater advantages than those who are relegated to functional usage?

The Second Digital Divide theory understands that technology use spans a continuum from consumption to creation. The positioning of women on the consumption end of this continuum is far from indicative of personal preference and skill, but rather a product of structural variables such as education streaming, organizational culture, and the continued linkage between technical skill and masculine constructions of gender (Cheryan et al., 2017). Stereotype Threat Theory contradicts the idea that women perform worse in technology due to a lack of skill and rather reveals that pointing out gender stereotypes reduces women's technical performance due to mental efforts directed toward averting the fulfillment of negative stereotypes (Spencer et al., 1999).

Social Cognitive Theory and Technical Self-Efficacy

According to Bandura's theory, as presented in 1986, the Proficiency Paradox can be adequately explained. Technical Self-Efficacy, which pertains to an individual's confidence level in carrying out technical tasks effectively, develops through the process of "triadic reciprocity," which involves personal beliefs, attitudes, and experiences, on the one hand, and personal performances, like the level of participation in doing technical work, on the other, along with environmental influences, like work culture, on the third side. This model helps to understand the discrepancy that may exist between an individual's technical competence and their level of technical confidence.

According to Bandura (1986), the four basic sources of self-efficacy are mastery experiences (performing tasks successfully), vicarious experiences (observing others similar to oneself perform successfully), verbal persuasion (being persuaded or dissuaded by others), and physiological states (physiological responses to task performance). These sources are prone to gender bias in the technology sector. Mastery experiences can occur when women are perceived as lucky, receive assistance from others, or are assigned tasks that utilize their own capabilities instead of easy ones. Secondly, vicarious experiences can occur due to the underrepresentation of women in senior technology positions. Thirdly, verbal persuasion may occur in the form of subtle discouragement, such as not being heard in discussions or being closely scrutinized for their performance in technology. Another physiological state can be anxiety due to stereotype threat or a hostile environment in the organization.

When the technical contributions of female engineers are judged more strictly than those of their male counterparts, a scenario that has been extensively reported within contemporary labor studies (IEEE, 2025), female self-efficacy is impacted independently of the actual performance quality. A landmark study by Terrell et al. (2017) reported on GitHub technical contributions, finding that when gender could not be determined, females' pull requests were accepted at a rate higher than those of males. However, when genders could be distinguished, the rate of females' pull requests dropped below that of males. There is direct evidence here of females' technical contribution facing stricter judgments once their gender is visible, making it harder for self-efficacy to develop. Thousands of instances of this type of influence females' perception of their technical efficacy, unrelated to their actual performance.

However, such decreased levels of self-efficacy are reflected in what I refer to as "Knowledge Silencing"; that is, women are less likely than others to call out technical knowledge in high-stakes settings even when they have the same levels of expertise. Research on participation in educational settings suggests that women's participation levels in computer science courses are lower than those of men, even when their levels of comprehension are equivalent (Karpowitz & Mendelberg, 2014). In professional settings, women have reported being reluctant to share technical views unless they are absolutely certain, whereas men often feel no such compulsion to wait before expressing their views (Hewlett et al., 2014). What ensues is thus a vicious cycle that decreases visibility levels, thereby further limiting opportunities for recognition, which in turn impacts levels of self-efficacy. All such aspects are independent of levels of technical expertise; that is, the gender gap within technology is not one of capability, but rather one of recognition.

Self-Efficacy and AI Tool Adoption

Current research has applied theory on self-efficacy to the context of AI technology specifically. Research has shown that, despite women having equal efficiency to men in AI-related tasks, they often lack confidence in performing these activities. This lack of confidence can be exacerbated by situations where expertise in AI technology is not perceived as requiring creativity or collaboration skills, but rather purely technological expertise (Jiang, 2024). The way AI technology orchestration is represented matters significantly in terms of gender bias. When women are asked to participate in 'prompt engineering' or any activity that requires technological skills, they often lack confidence. If it involves "AI collaboration" or "AI-assisted problem-solving," there will not be a lack of confidence.

Such a framing effect holds significant implications for how AI skills are recognized and compensated. If AI orchestration is viewed as a technical competency, it is vested with the status and rewards associated with technical skills. If AI orchestration is viewed as a communication/coordinating competency, it can become feminized and downgraded, regardless of its complexity. Preliminary trends suggest that the technology sector is shifting towards the former framing, which may perpetuate existing patterns of exclusion related to gender, even as technological skills evolve. Recognition of and deliberate action on such framing dynamics is critical to ensure that AI does not perpetuate existing patterns of gender hierarchies that existed during previous technological ages.

Methodology

A mixed-method design involving a Systematic Meta-Analysis (SMA) of existing quantitative research and a qualitative case study was used to probe the Proficiency Paradox phenomenon. The research design was chosen for its capacity to combine the results of numerous research studies conducted using a wide array of methodologies and locations, and also for its potential to provide insight into real-world experience that cannot be captured using quantitative research findings alone. The meta-analysis provides estimates regarding the size and degree of evaluation difference, while the case study offers insights into the process by which this difference is created.

Systematic Meta-Search Strategy and Data Selection

The meta-analysis is based on PRISMA 2020 guidelines (Page et al., 2021). Three prominent academic databases, namely ACM Digital Library, IEEE Xplore, and Scopus, have been considered in this paper. The choice of these databases has been informed by their broad subject coverage in computer science, engineering, and technology. The searches were conducted in October 2025 and encompass all articles published between January 2021 and September 2025, considering the current state of technology and the adoption of Generative AI. The search strings were developed cumulatively with the help of a research librarian. The search strings were as follows: ("women in technology") AND ("technical proficiency" OR "AI skill acquisition" OR "coding assessment" OR "technical self-efficacy"); ("gender") AND ("software engineering" OR "artificial intelligence") AND ("performance" OR "evaluation"); ("gender bias") AND ("code review" OR "technical evaluation" OR "hiring"); and ("digital divide") AND ("gender" OR "women") AND ("Global South" OR "developing countries"). The Boolean operators have been modified according to the requirements of the databases.

The original search yielded 847 results. After removing the 156 duplicate articles, the remaining 691 articles had their titles and abstracts assessed for inclusion criteria. The inclusion criteria for studies to be considered for the review were: (a) they had to include both objective performance criteria (such as coding assessment results or the rate and quality of codes completed) and subjective performance criteria (such as peer review scores), (b), they had to break down results by gender in a manner stratified for the calculation of effect size, (c) they had to include professional or post-secondary educational contexts (graduate-level or in industry), and (d) they had to be published in a refereed publication. Studies not focused on professional-level coding tasks in K-12 education, recreational use, or basic digital literacy were excluded.

After title and abstract screening, 142 articles went on to full-text screening. Among these, 88 articles were excluded due to a lack of statistical information (n = 41), unavailability of both objective and subjective information (n = 32), a focus on student populations rather than professionals (n = 11), and for other reasons (n = 4). The resultant sample included 54 journal articles featuring a combined sample size exceeding 12,000 technical professionals from 15 different countries. Countries include key technology hubs in North America (USA and Canada), Europe (UK, Germany, Netherlands, Sweden, and France), Asia-Pacific (Japan, South Korea, Singapore, and Australia), and nascent technology hotspots in the Global South (India, Vietnam, Philippines, Kenya, and Brazil).

Analytical Approach

To determine the Hedges' g effect size, the random-effects model was applied. The Hedges' g effect size is not subject to small-sample bias and can be used to combine different datasets. The Hedges' g values were computed separately for differences in absolute performance (the contrast of scores on standardized tests between women and men) and subjective evaluations (the contrast of appraisals of performance between women and men based on the appraiser's awareness of gender differences).

Heterogeneity was evaluated using the I^2 statistic, where values of < 25% represented low heterogeneity, 25% to 75% represented moderate heterogeneity, and > 75% represented large heterogeneity, necessitating further subgroup analyses. Subgroup analyses were conducted for geographical region (Global North versus Global South), industry type (Enterprise software, Start-ups, Academic research), technical task (coding, system architecture, data science), and type of evaluation (hiring, performance appraisal, peer review). Sensitivity analyses were conducted to assess the stability of the results after removing outliers and reevaluating effect sizes.

The presence of publication bias was examined using the funnel plot method and the Egger regression test. Although the funnel plot revealed slight asymmetry, indicating a possible tendency towards the publication of statistically significant results, the results obtained using the trim-and-fill method suggested that there was no major influence on the results, despite the possible missing studies. The risk of bias present in the individual studies was examined using an approach modified according to the Newcastle-Ottawa scale, taking into consideration the processes used to select the studies, the comparability of the collected data, and the measurements used to evaluate the results.

Case Study: AI Orchestration - Southeast Asia

In addition to the meta-analysis, this dissertation presents a specific case study to inform the effects of Generative AI on gender dynamics within the context of new technologies. The case study focuses on women-led tech startups in both Vietnam and the Philippines. Both countries were chosen due to their increased rates of technological development, widespread adoption of Generative AI, as well as how these countries develop technological skills in traditional ways or pursue alternative development paths. In its current state, the data utilized in this specific case study provides substantial insight beyond the results of the meta-analysis.

Semi-structured interviews were conducted with 28 founders and technical leads from women-led startups between January and June 2025. Participants for the research were sourced from technology incubators, women in technology groups, and snowballing techniques. Each interview was scheduled to last 60-90 minutes and was conducted via video call, either directly in English or with the assistance of simultaneous interpreters. The protocols used for conducting the interviews included technical experience, the use of AI technology, experiences with gatekeeping/bias, funding, and views regarding the valuation of the work conducted. All interviews were recorded and analyzed thematically according to Braun and Clarke's (2006) six-step approach.

In addition to the interviews, the funding information provided in the Asia Pacific Innovation Database was examined, a regionally focused database of investment in technology startups. Awards of funding, interest in the startup, as reflected by the number of pitching sessions and the number of valuation multiples, were examined between women-led and men-led startups with similar technological outputs and market orientations. The productivity outcomes stemmed from a 2025 pilot project initiated by the Philippine Commission on Information and Communications Technology, which assessed the outputs and quality of work of 47 women who utilized AI orchestration tools to develop e-commerce platforms.

The design of the case studies adheres to the guidelines proposed by Yin (2018) for explanatory case studies, aiming to clarify how general patterns found in quantitative data are realized in cases. Although not intended to be generalizable, case study results enable generalizability in an analytical sense, namely, they facilitate an understanding of processes. The merging of results from interviews, funding information, and productivity data enables triangulation, thereby increasing confidence in the results.

Results

Results of meta-analysis: objective parity versus subjective assessment

The meta-analysis indicates a striking disparity between what women are capable of doing and how they are regarded. In the standardized technological evaluation conducted in 2024 and 2025, which included algorithmic problem-solving, system architectural designs, programming bug exercises, and code review accuracy, women performed on par with their male counterparts. The overall effect size for the disparity in objective ability had a Hedges' g of 0.03, which was not significant (95% CI [-0.05, 0.11], $p = .47$), showing there was no significant disparity in terms of ability. The value of I^2 was low (18%), indicating a well-established consistency in demonstrating parity of objective ability.

These patterns were confirmed for all categories considered via subgroup analyses. For global hubs of technology in the Global North, the effect size was Hedges' g = 0.02 (95% CI [-0.08, 0.12]). For emerging economies of the Global South, it was Hedges' g = 0.05 (95% CI [-0.06, 0.16]). No significant differences were found with respect to industry sector or form of technical work. These results strongly show that there is no 'skills gap' narrative to be found: among 54 studies and over 12,000 participants, women are found to offer equal technical skills to men on objective and standardized tasks.

However, when these same results were assessed within 'non-blind' environments, in which the sex of the producers was known, the average perceived quality of the females' performance was rated lower by an average of 12% (Hedges' g = -0.38, 95% CI [-0.45, -0.31], $p < .001$). There was no heterogeneity when comparing all 15 countries ($I^2 = 31\%$). Results were smallest for Sweden (8%) and the Netherlands (9%), which had both strong institutional settings for gender equity, and largest in Japan (17%) and South Korea (16%), which also boast technology industries that remain overwhelmingly male-dominated. Notice that large evaluation differentials remain present within the most equal of societies.

Results for subgroup analyses by evaluation context are highly informative. In selection scenarios, the evaluation difference was largest (Hedges' g = -0.45), indicating that the issue of gatekeeping is most pressing at entry points for technical work. In appraisal scenarios, the degree of evaluation difference was moderate (Hedges' g = -0.36), and in peer review scenarios, slightly less (Hedges' g = -0.29). Such variations may be due to levels of accountability and formality in evaluation scenarios; for instance, less formal peer reviews can offer opportunities for personal judgment to mitigate evaluation differences due to collegial dynamics.

Of particular concern was the pattern that emerged under the label of "high stakes" or "innovative" technical work. When technical work was characterized as routine maintenance/improvement, there was virtually no gender gap in perceptions of quality (g = -0.12, 95% CI [-0.21, -0.03]). However, where the work was characterized as strategic innovation, architectural design, or technical leadership, the gap was huge (g = -0.52, 95% CI [-0.62, -0.42]). This relationship suggests that where technical work holds a higher status, the work on which advancement, recognition, and leadership are based, the gender gap in quality perceptions is at its widest. Women can be skilled in routine technical work; they can't be architects and innovators.

Case Study Results: AI as Force Multiplier and the Glass Wall

This case study on the evaluation gap provides an in-depth understanding of how the evaluation gap operates, as well as how Generative AI technologies both disrupt and perpetuate gendered patterns in technological work. In the Southeast Asia settings where this study was conducted, the role of AI has been described as a 'Force Multiplier' that has the capability of overcoming the gatekeeping process. This refers to the way that women, lacking direct credentials in computer science, were able to utilize these tools to perform their technological work, which previously required years of qualification.

A pilot project conducted with women from remote areas in the Philippines in 2025 demonstrated the potential of AI orchestration, specifically the ability to integrate and combine multiple AI tools to perform complex technical tasks. Those women, lacking a degree in Computer Science, yet employing AI tools for complex e-commerce backend design, achieved a 35% increase in productivity compared to a baseline of women pursuing academic credentials according to conventional standards. In this case, the women displayed refined technical skill, even though designing and managing several AI toolsets for complex design and testing might not, and arguably does not, constitute 'coding.'

Employees spoke about their tasks in terms of technical complexity and decision-making. As a founder in Ho Chi Minh City stated: "I'm not coding in Python all day, but I'm making thousands of technical decisions on the model to use, how the questions must be phrased, how outputs must be verified, how different models must be merged. So, it's architecture even if it doesn't look like coding." Another employee, a technical lead at Manila, also pointed this out: "People think you can just ask ChatGPT questions. However, it requires a significant amount of technical knowledge to have an adequate system. You need to be familiar with APIs, data structures, security, and other related concepts. You just implement it differently."

However, institutional investment levels in these endeavors continued to be significantly lower for ventures led by men compared to those led by women, despite having similar technical achievements and market potential. Analysis of investment data available in the Asia Pacific Innovation Database revealed that women-led entrepreneurial ventures employing an AI orchestration strategy received 60% lower funding investments compared to those led by men, after adjusting for industry, location, and revenue growth. However, what is more significant is that for women-led ventures, an average of 2.3 times as many pitch meetings had to be attended to attract the same level of funding.

Analysis of the interviews revealed that investors consistently referred to women's AI-enabled work as "prompt dependency" rather than genuine architectural ability. Respondents were asked whether they "really understood" the technology or only "used AI" to achieve their results, something not commonly asked of men, founded on similar principles. To demonstrate just how differently AI results are viewed based on gender, a founder told me, "Literally had an investor ask me if they would still be able to run their company if ChatGPT went down. And then I asked them whether they would ask a male founder of a similar startup based on AWS technology if they could still run their company if Amazon went down. And they didn't have an answer."

The above scenario aptly captures what I refer to as the "Glass Wall," an invisible barrier that filters technical agility, making it visible yet not considered legitimate technical expertise worthy of investment. The Glass Wall is, therefore, distinct from what is popularly referred to as 'glass ceiling' because, while the latter prevents promotion, the former creates a new definition for work itself. When women showcase their technical skills in a new way, such as AI orchestration, the 'real' technical work becomes the new definition that excludes the work done by women. The goalposts are, therefore, moved such that women's work remains on the "wrong" side of the proverbial fence.

They reported this experience repeatedly throughout their careers. One Vietnamese founder summed their experience this way: "First, we couldn't code. Then, when we learned to code, we couldn't build systems. Then, when we built systems, we couldn't lead teams. Now that we're leading teams and using AI to move faster than those who just code, AI is no longer just a technology. The finish line just moves." It describes the experience that the Glass Wall encapsulates: that the glass wall is not a fixed entity; rather, the criteria for what counts as technology are constantly shifted to prevent the input, no matter the form, that women make.

Discussion: The Three Pillars of Institutional Gatekeeping

Findings from both the meta-analysis and the case study converge on the following clear implication: the gender gap in tech is a recognition gap, not a skills gap. Women excel in technical skills as objectively measured but are rated as inferior in subjective judgments. This paper will distill the findings from both methodology approaches to develop three main pillars on which the Proficiency Paradox, as a gender gap in tech, persists.

Pillar One: Pedigree Bias

First is pedigree bias, the need for prestige credentials that systematically filter out women applicants who may have a non-linear technical background. The findings also reveal that women comprise a disproportionate number of applicants who have followed non-traditional routes into technology, such as boot camps, career transitioners, self-taught developers, and those from alternative education platforms. Although these might offer the same or even better technical skill sets, as evidenced by the increased productivity that resulted from the case study, they lack the same legitimization and recognition that a Computer Science degree from top-tier universities enjoys.

This produces a double disadvantage for women. Firstly, the obstacles that prevent women from pursuing traditional credentials (such as stereotype threat in undergraduate STEM education or career interruptions due to child-rearing) ensure that more women than men embark upon non-traditional routes. Secondly, these non-traditional routes are undervalued precisely because of their association with outsiders. This results in an overall system of credentialing that is seemingly meritocratic yet disadvantages those individuals who failed or chose not to take the traditional route, despite possessing the necessary aptitude.

The disruptive nature of AI applications may impact pedigree bias in the technical field in the following manner: AI may prove to be an egalitarian force in the technical field because it does not require any technical pedigree to perform tasks through the application of AI. Nevertheless, as borne out by the case study, the rapid redefinition of technical tasks as AI Application-integrated tasks as 'not technical' in nature can vitiate this disruptive impact because the deployment of terms such as 'prompt dependency' and the question of 'whether one really understands the knowledge application' through AI use could be considered as the reproduction of existing power dynamics in the technical field.

This phenomenon reflects a broader historical trend of "social closure" within the industry. As noted by Ensmenger (2010), the criteria for technical expertise shifted in the 1970s from formal mathematical rigor—where women were beginning to excel—to more nebulous concepts of "innate passion" and "culture fit" (Margolis & Fisher, 2002). Today, as AI democratizes mathematical proficiency, new definitions of "true" skill are emerging to exclude those who acquire these competencies through AI-assisted pathways (Koppman, 2016).

Pillar Two: The Second Shift in Technical Work

The second pillar is an application of Hochschild's (2012) "second shift" idea to this present-day technological context. This is where the caregiving role, which is often dominant among women, is reflected in a scarcity of 'disposable time' for their unpaid professional development in their technological careers. Technological culture's commitment to continuous learning is beneficial for keeping up with technological changes; however, it is an unequal burden imposed on those with caregiving roles compared to others. Technological conference sessions, open-source projects, hackathons, and side projects are equally important for professional agendas and all demand time that is not equally available.

The findings reveal that women in technical positions have eight fewer hours per week available for professional development activities compared to their counterparts in identical positions and with identical family structures, even after adjustments for work position and level within their respective organizations. The disparity is particularly pronounced for women with children, with 12 fewer hours per week available for those with children under the age of five. In contrast, fathers with young children experience no comparable increase in available time for professional development activities.

"This lag is directly correlated to differences in their perceived 'passion' for technology and their levels of 'commitment' to technology as well. They are judged not only according to how well they're working as direct contributors to their own organizations but also according to their engagement in the broader technical culture—the code they're committing to GitHub, the conferences they're speaking at, the commenting threads they're participating in. Maintaining a work-life balance will be viewed as evidence of lower commitment when it is realized that promotions at work will be awarded to those who attend all meetups and hackathons. If not, then "I have kids—I don't get to attend all those things. But my code is as good as everyone else's," said one of the women in the study."

This, the era of AI, brings new variables into play for this component. Computers equipped with AI technology may be able to reduce the time required for certain technical tasks, allowing for more time for other pursuits. However, this advantage may be countered by the rapid pace at which AI technologies are being developed, creating even greater demands for continuous learning. Keeping up to speed on the latest that AI can do means investing more time, which those who lack time simply cannot do. To add to this, the practitioners of the field of AI have already mimicked the various social phenomena that exist in the larger technology world, where things like hackathon events, challenges, and online groups foster activities that are time-consuming, where those who work in AI simply add more things to the list that women cannot do.

Pillar Three: Algorithmic Erasure

The third leg of this figure, that of algorithmic erasure, refers to the way gender biases created by AI-powered recruitment and assessment systems mirror and reproduce historical gender biases. Such systems learn from existing databases that reflect historical practices of hiring and assessment, all of which systematically perpetuate gender biases. Thus, emerging forms of automated discrimination may appear perfectly objective, as they are all measured and calculated.

The impact of career gaps is particularly problematic in terms of addressing them. Career gaps, which tend to be predominantly experienced by women due to child-rearing, are viewed by AI hiring programs as an indication of lower levels of commitment or competency deterioration. An analysis of documentation provided by leading firms offering hiring platforms reveals that career gaps lasting 12 months or longer result in a decrease in candidate scores by as much as 15 to 25%, with this trend intensifying for longer career gaps. This negative impact cannot be mitigated by skills upon re-entry into the job market, as candidates who attain high scores in skills tests will still be considered, albeit with lower weighting, when viewed in terms of career gaps. It is also known that women often take career breaks to raise children.

In addition to career gaps, algorithm-driven systems replicate patterns of gender bias in more insidious ways. Resume screening software, having been conditioned on data from past hiring decisions, tends to discriminate against resumes containing words or terms associated with females. Job recommendation software recommends non-technical, low-pay positions to females based on data points from past practice. Rating systems involving peer review magnify the rating gap reflected in the meta-analysis presented herein on a massive scale and with the authority of algorithmic objectivity.

The lack of transparency in algorithmic systems exacerbates this problem. The candidates may not even know the reasons for their rejection or low ranking, as algorithmic decision-making processes are often concealed behind arguments of proprietary technology and trade secrets. Although there may be reasons to suspect algorithmic bias in this case, it cannot easily prove it, since they are not willing to release their training data and models. The regulatory action for algorithmic auditing may prove helpful but not adequate, as the very flaw in algorithmic systems may lie in their dependence on biased data, which would be reproduced unless the algorithm is programmed not to. Algorithmic erasure thus embodies the Proficiency Paradox in automated form.

Conclusion and Policy Recommendations

The Paradox of Proficiency will never be solved by more coding camps or pipeline fixes by themselves. What the evidence makes clear is that women already possess equal proficiency in technology to men. Rather, the issue that arises is how proficiency itself can or should be measured and valued. The three legs of the institutional sexism stool – pedigree bias, second shift, and algorithmic erasure – are all independent and cumulative ways of ensuring that women's proficiency in technology will never truly be valued. Closing the technology gender gap will therefore only happen by transforming how proficiency in technology itself is definitively measured.

Within this framework, the following policy recommendations are proposed based on the study's findings. One recommendation is suggested per pillar, based on the findings. These proposed recommendations can be implemented by organizations, trade bodies, and governments, and each has evidence suggesting their effectiveness in eliminating gender biases.

Recommendation One: Skill-Based Equity through Double-Blind Evaluation

The results suggest that the effect of the 12% evaluation gap becomes zero once the presence of reviewer knowledge regarding authorship is removed. A clear implication of this result is that blind reviews must be made a widespread practice that includes all technical evaluation situations. In this regard, I propose that technical firms implement "Double Blind Code Reviews," a process whereby code submissions are reviewed without the reviewers knowing the identities of the submitters. Currently, this is a common process followed by projects that use open-source software, such as the Python language, whereby core developers review all submitted contributions that are not yet part of the language, without knowing who provided them.

In addition to code review, blind evaluation should also be applied to hiring evaluations, promotion evaluations, and project assignments wherever possible. Technical interviews may include work samples that are assessed while maintaining the individual's confidentiality. Promotion review committees may evaluate portfolios with names removed. Selection for high-visibility projects may depend on blind evaluation for relevant skills. Even if complete blindness is impossible (i.e., there is no way to avoid eventually meeting candidates and colleagues who are familiar with each other's work), maximizing blind evaluation in the initial stages may help prevent bias from influencing which candidates proceed to the next stages.

Implementation also needs organizational resources and infrastructure. Blind assessment facilitation platforms must be implemented or developed. Systems must be changed to decouple assessment and identification. Assessors must be trained to distinguish between seeking identifying data and the possibility of bias at work. Such efforts are costly, but they offer the dual advantage of combating discrimination while also optimizing decision-making. When assessments are grounded in capabilities rather than identity-related inferences, organizations make optimal decisions, not just fair ones.

Recommendation Two: Providing Institutions with Infrastructure for Asynchronous Delivery

To meet the needs of the second shift, I propose transitioning from 'time-at-desk' metrics to 'asynchronous skill delivery' models. This would help serve different stages of life and responsibilities for care with a continued emphasis on what matters—the work product. It became clear from the shift to remote work during the pandemic that productivity can, and does, remain high, and even improve, regardless of presence, and this habit needs to stick.

Specifically, this involves an overhaul of performance review metrics, shifting from activity-based to deliverable-based metrics. Attending meetings should be minimized, or their asynchronous counterparts should be arranged whenever possible. Professional development can be rewarded and counted when delivered through alternative means, not solely through conference or meet-up-based activities. Caregiving roles can be recognized not as personal failure but as legitimate time claims. Rather than viewing gender equity, organizations can tap the potential of a larger pool of human resources – those professionals who do not match temporal expectations.

Professional associations and trade groups can facilitate this shift by organizing their activities more effectively. Conference participation can include full engagement in the conference activities through robust virtual participation options. Professional certification can be achieved through flexible and self-paced education rather than requiring intensive classroom education and boot camps. Open-source projects can emphasize contribution through documentation, review, and mentorship, rather than focusing on new features, and can do so just as easily through asynchronous work, rather than requiring intensive collaboration and teamwork. Rethinking contribution is part of the solution that the Proficiency Paradox requires.

Recommendation Three: Recognition of Invisible Labor

The kinds of critical technical contributions, such as documentation, security audits, code reviews, testing, ethical reviews, and mentorship, which are often carried out by women despite not being valued in the promotion process, need to be acknowledged and rewarded. These tasks are crucial for the success and functionality of the organization because, without system documentation, there can be no well-organized system; without security auditing, there can be no secure system; and without mentorship, there can be no effective teamwork.

Organizations must conduct an analysis to determine the distribution of invisible labor in their technical teams and identify who contributes to this critical work, as well as whether the work is accounted for. Resumes and criteria for levels must factor this work into the mix, along with the other critical work that gets accomplished. These criteria for levels must require proof that the individual contributes to the team's success through mentoring, knowledge sharing, and quality assurance.

As technologies like AI continue to increase in adoption, new forms of hidden labor are emerging. Task validation of AI outputs, creation of prompt libraries, the role of validating the ethics and safety of AI systems and educating other workers in the use of AI are critical tasks that can easily be undervalued because they do not fit the traditional definition of technical labor. Organizations must take the initiative in valuing this labor before the gender dynamics of this undervaluation can become institutionalized as the norm. The selection of what qualifies as 'real' technical labor is not an objective process but is a contested space.

Limitations and Further Research

The study also has some limitations that provide directions for further research. First, although the meta-analysis is very comprehensive, it is based only on published studies. There is an overrepresentation of statistically significant results within these studies. Furthermore, they do not provide an accurate view of the context within which technical evaluation is practiced. The second approach is also limited. It is very detailed and provides results only for two countries in Southeast Asia. There is also a lack of generalizability in the findings.

Furthermore, the study primarily focused on binary gender categories. Future studies should consider how and to what extent the Proficiency Paradox applies to non-binary individuals and should involve intersectionality analyses that explore how gender and additional aspects of oppression, such as race, class, and disability, intersect in evaluating technical proficiency. There is evidence that women of color are differentially prejudiced in technical evaluation, but such data as ours did not allow us to scrutinize such intersectional relationships. Analyzing technical evaluation bias involves consideration of the complexity inherent in intersectionality.

Ultimately, there is a need for further research into the implementation and effectiveness of these measures. Although blind assessment can be shown to remove bias on tasks, very little research exists regarding the implementation of blind assessment systems on an organizational level. Information on factors such as facilitators and inhibitors of equity-oriented practices would benefit practitioners who wish to apply the results of this report.

Concluding Remarks

The Glass Wall, named and identified, exists, but it is not a fixed reality. It is a function of everyday interactions, institutional patterns, and technology, all of which can change. A shift from a focus on skills to a concern about recognition failure may initiate the process of deconstructing the structures that support the Proficiency Paradox. The technical talent for a revolution in the technology industry already exists; there just needs to be a willingness to recognize it.

While Generative AI brings both risk and promise, it will either fulfill the promise of new paths to capability or replicate the patterns documented in this paper, depending on how it is defined, how its skills are valued, and how it is implemented. If it does the latter, then the Proficiency Paradox will carry over into the AI era. While Generative AI does bring its own risks, in that it could automate inequality by simply scaling existing disparities while hiding them under its advanced technology, it could also be the tool that smashes the barriers currently preventing countless individuals from acquiring capabilities.

The women participants, who develop complex technical systems, establish superior productivity, and foster innovative business ventures, embody the talent that the technology industry purports to require. They are lacking neither in talent nor cognitive resources; they are lacking in systems that fail to properly validate their talent as legitimate. What is called for is an increased focus on accountability rather than an increased need for training on the part of women. Rather, it is the Proficiency Paradox that needs to be acknowledged and addressed, and this requires evidence, of which this study now furnishes an example.

Funding: This research received no external funding.
Conflicts of Interest: The authors declare no conflict of interest.

References

[1]. Bandura, A., & National Inst of Mental Health. (1986). *Social foundations of thought and action: A social cognitive theory.* Prentice-Hall, Inc.

[2]. Boston Consulting Group. (2024, May 14). *Women leaders in tech are paving the way in GenAI*. https://www.bcg.com/publications/2024/women-leaders-in-tech-are-paving-the-way-in-genai
[3]. Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology*, *3*(2), 77–101. https://doi.org/10.1191/1478088706qp063oa
[4]. Buolamwini, J., & Gebru, T. (2018). Gender shades: Intersectional accuracy disparities in commercial gender classification. *Proceedings of Machine Learning Research*, *81*, 1–15. https://proceedings.mlr.press/v81/buolamwini18a.html
[5]. 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), 1–35. https://doi.org/10.1037/bul0000052
[6]. Ensmenger, N. L. (2010). *The computer boys take over: Computers, programmers, and the politics of technical expertise*. MIT Press. https://mitpress.mit.edu/9780262517966/the-computer-boys-take-over/
[7]. Hargittai, E. (2002). Second-level digital divide: Differences in people's online skills. *First Monday*, *7*(4). https://doi.org/10.5210/fm.v7i4.942
[8]. Hargittai, E., & Shafer, S. (2006). Differences in actual and perceived online skills: The role of gender. *Social Science Quarterly*, *87*(2), 432–448. https://doi.org/10.1111/j.1540- 6237.2006.00389.x
[9]. Hewlett, S. A., Marshall, M., & Sherbin, L. (2014). How diversity can drive innovation. *Harvard Business Review*, *91*(12), 30. https://hbr.org/2013/12/how-diversity-can-drive-innovation
[10]. Hochschild, A. R., & Machung, A. (2012). *The second shift: Working families and the revolution at home*. Penguin Books.
[11]. IEEE Transactions on Engineering Management. (2025). Measuring technical self-efficacy in distributed teams. *IEEE Transactions on Engineering Management*, *72*(2), 112–128. https://ieeexplore.ieee.org/xpl/RecentIssue.jsp?punumber=17
[12]. Jiang, S. (2024). Towards inclusivity in AI: A comparative study of cognitive engagement between marginalized female students and peers. *British Journal of Educational Technology*, *55*(6), 2557– 2573. https://doi.org/10.1111/bjet.13467
[13]. Karpowitz, C. F., & Mendelberg, T. (2014). *The silent sex: Gender, deliberation, and institutions*. Princeton University Press.
[14]. Koppman, S. (2016). Different like us: Theory, high-tech culture, and the production of inequality. *Sociological Theory*, *34*(4), 333–353. https://doi.org/10.1177/0001839215616840.
[15]. Margolis, J., & Fisher, A. (2002). *Unlocking the clubhouse: Women in computing*. MIT Press. https://mitpress.mit.edu/9780262632690/unlocking-the-clubhouse/
[16]. McKinsey Global Institute. (2025). *The power of parity: How advancing women's equality can add $12 trillion to global growth*. McKinsey & Company. https://www.mckinsey.com/featured-insights/employment-and-growth/how-advancing- womens-equality-can-add-12-trillion-to-global-growth
[17]. Obermeyer, Z., Powers, B., Vogeli, C., & Mullainathan, S. (2019). Dissecting racial bias in an algorithm used to manage the health of populations. *Science*, *366*(6464), 447–453. https://doi.org/10.1126/science.aax2342
[18]. Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., et al. (2021). The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. *BMJ*, *372*(n71). https://doi.org/10.1136/bmj.n71
[19]. Randstad. (2025). *Understanding talent scarcity: AI & equity report*. https://www.randstad.com/randstad-ai-equity/
[20]. Spencer, S. J., Steele, C. M., & Quinn, D. M. (1999). Stereotype threat and women's math performance. *Journal of Experimental Social Psychology*, *35*(1), 4–28. https://doi.org/10.1006/jesp.1998.1373
[21]. Terrell, J., Kofink, A., Middleton, J., Rainear, C., Murphy-Hill, E., Parnin, C., & Stallings, J. (2017). Gender differences and bias in open source: Pull request acceptance of women versus men. *PeerJ Computer Science*, *3*, e111. https://doi.org/10.7717/peerj-cs.111
[22]. UNESCO. (2024). *Implementation of standard-setting instruments, Part V: Implementation of the 2021 Recommendation on the Ethics of Artificial Intelligence – Preparations for the next consultation* (Document 219 EX/17.V). https://unesdoc.unesco.org/ark:/48223/pf0000388656
[23]. van Dijk, J. A. G. M. (2020). *The digital divide*. Polity Press. https://www.politybooks.com/bookdetail/?isbn=9781509534449
[24]. Warschauer, M. (2004). *Technology and social inclusion: Rethinking the digital divide*. MIT Press. https://direct.mit.edu/books/oa-monograph/1817/Technology-and-Social-InclusionRethinking-the
[25]. Yin, R. K. (2018). *Case study research and applications: Designs and methods* (6th ed.). Sage Publications. https://uk.sagepub.com/en-gb/eur/case-study-research-and-applications/book250150