(flabby, obese) in Japanese and "fwefwefwe" (springy, elastic) and "saaa" (cool sensation) in Siwu, a language spoken in Ghana (*11*). The ability to coin and appreciate these semantically diverse words is not adequately explained by a discrete set of predisposed cross-modal associations such as the bouba-kiki effect.

Although the bouba-kiki effect cannot account for the wide-ranging manifestations of iconicity found in spoken words, it could have played a role in grounding the first vocal symbols because of the extent to which it is innate and universal across humans. Loconsole *et al.*'s findings in chickens seem to amplify this possibility, which suggests that the effect is deeply ingrained in human phylogeny. Notably, however, the bouba-kiki effect is not an especially strong case of iconicity in speech sounds. For example, a cross-cultural study that included speakers of 28 languages spanning diverse cultures and language families reported a stronger and more consistent association between a trilled /r/ sound with a jagged, "rough" line and a lateral approximant /l/ with a flat, "smooth" line (*12*) than a comparable study of the bouba-kiki effect (*2*).

Beyond these specific associations, the human capacity to communicate with iconic vocalizations—nonlinguistic vocal sounds that depict or resemble their meaning—is expansive. For example, in another cross-cultural study, speakers of different languages were able to interpret new vocalizations created to communicate 30 meanings that included various entities (e.g., child, deer, fire, fruit), actions (e.g., eat, cut, hide), and properties (e.g., big, many, sharp) (*13*). Such findings demonstrate the vast and open-ended human talent for vocal charades, a skill that distinguishes them from all other animals.

As good as they are at vocal communication, humans might be even better at playing charades with their hands (*14*) or through drawing (as in the game of Pictionary). Even if the basis for the bouba-kiki effect is phylogenetically ancient, as shown by Loconsole *et al.*, it is but a single manifestation of the near-boundless human capacity for iconicity that transcends any particular medium of communication. Therefore, although the study of innate cross-modal correspondences such as the bouba-kiki effect may uncover general principles of sensory perception, a broader framework of iconicity is needed to illuminate the origins of language. □

REFERENCES AND NOTES

1. M. Fort *et al.*, *Dev. Sci.* **21**, e12659 (2018).
2. A. Ćwiek *et al.*, *Philos. Trans. R. Soc. London Ser. B* **377**, 20200390 (2022).
3. M. Loconsole, S. Benavides-Varela, L. Regolin, *Science* **391**, 836 (2026).
4. M. Imai, S. Kita, *Philos. Trans. R. Soc. London Ser. B* **369**, 20130298 (2014).
5. V. S. Ramachandran, E. M. Hubbard, *J. Conscious. Stud.* **8**, 3 (2001).
6. P. Perniss, R. L. Thompson, G. Vigliocco, *Front. Psychol.* **1**, 227 (2010).
7. M. Pleyer, M. Perlman, G. Lupyan, K. de Reus, L. Raviv, *Trends Cogn. Sci.* 10.1016/j.tics.2025.10.004 (2025).
8. G. Hodge, L. Ferrara, *Front. Psychol.* **13**, 808896 (2022).
9. W. C. Stokoe Jr., *J. Deaf Stud. Deaf Educ.* **10**, 3 (2005).
10. D. McNeill, *Hand and Mind: What Gestures Reveal about Thought* (Univ. Chicago Press, 1992).
11. M. Dingemanse, D. E. Blasi, G. Lupyan, M. H. Christiansen, P. Monaghan, *Trends Cogn. Sci.* **19**, 603 (2015).
12. A. Ćwiek *et al.*, *J. Acoust. Soc. Am.* **156**, 3468 (2024).
13. A. Ćwiek *et al.*, *Sci. Rep.* **11**, 10108 (2021).
14. N. Fay, M. Arbib, S. Garrod, *Cogn. Sci.* **37**, 1356 (2013).

10.1126/science.aee8641

Department of Linguistics and Communication, University of Birmingham, Birmingham, UK. Email: m.perlman@bham.ac.uk

SOCIAL SCIENCES

AI raises the productivity bar

Understanding how AI affects productivity has implications for labor markets

Lingfei Wu[1] and **Bogdan Vasilescu**[2]

Artificial intelligence (AI) is often assumed to democratize productivity by lowering the cost of production, especially for less-experienced workers (*1*, *2*). An opposing view is that AI concentrates productivity among the experienced. These competing expectations arise because they rest on different definitions of productivity. If it means generating more text, figures, or lines of computer code, then AI should benefit novices more by making the production of such output cheaper. But if productivity depends on whether the output has market value and keeps people employed, then AI should benefit experienced workers more by amplifying the value that they bring from their accumulated knowledge, skills, and judgment. On page 831 of this issue, Daniotti *et al.* (*3*) report that entry-level software developers show little productivity gain when using AI compared with their experienced counterparts.

Software development is one domain where empirical evidence on the impact of generative AI is accumulating quickly, thanks in large part to the wealth of trace (time-stamped) data available from GitHub, a public platform where developers store and manage code. Although measuring productivity from public repository trace data depends on many assumptions, the scale of such data is orders of magnitude larger than self-reports or controlled experiments, thus offering considerable statistical power. Daniotti *et al.* analyzed millions of open-source changes to Python files on GitHub using a machine learning model (classifier) trained to recognize AI-generated code. The authors show that adoption of generative AI has grown rapidly, accounting for roughly one-third of Python functions written by late 2024. They also found that only experienced developers—those with 6 or more years of GitHub activity—show a measurable rise in commit rates (a measure of software development velocity that can be interpreted as a proxy for productivity), whereas junior developers show little gain. If this finding were to hold broadly for production-grade software beyond the public Python files of the study, it would mean that even as AI use spreads quickly, the capacity to convert generated code into working software does not grow with it.

The findings point to a potential threshold model of human-AI productivity in which the quality of human input translates into productivity gains only after a minimum amount of expertise is reached. When content can be generated cheaply, productivity depends less on producing output and more on integrating AI-generated material into complex systems. Below this threshold, it may be that more generated content simply absorbs working time—output must still be read, edited, and maintained at substantial cost (*4*). Once the threshold is crossed, however, AI could accelerate productive work by reducing execution time. In this sense, AI does not lower the productivity bar but raises it.

In software development, human input takes two kinds of judgment. One is evaluation. Junior developers often treat AI-generated code as tentative solutions—they insert outputs into projects, keep them if they run, and revisit problems only when something breaks (*5*). This workflow generates false positives—code that appears functional in isolation but fails under real-world complexity or accumulates technical debt (*6*). Experienced developers, by contrast, draw on accumulated knowledge of abstractions, design patterns, and failure modes to predict how code will behave, filtering out fragile AI outputs (*4*). As a result, evaluation—not generation—becomes the primary bottleneck in turning AI assistance into productivity.

The other form of judgment is delegation. Senior developers spend much time deciding what work to assign, how much autonomy to grant, and how closely to monitor results, skills developed through years of system design and team coordination. This experience carries over to their interactions with AI, helping them write effective prompts and get useful solutions. Junior developers are often asked to carry out assigned tasks rather than directing others (*5*). As AI systems shift from code completion toward more autonomous execution, productivity may depend less on evaluating outputs than on deciding what to delegate.

This threshold model has implications for labor markets in software development. If AI rewards workers who already have strong evaluation and delegation skills, then adoption will widen productivity differences rather than compress them. Indeed, recent labor market analyses suggest that demand for senior developers has expanded, whereas hiring for junior roles has slowed (*7*, *8*). This reflects reduced entry-level hiring rather than increased separations or faster promotion (*9*), which suggests that AI alters who is getting hired. By raising the minimum amount of expertise required to achieve net productivity gains, AI changes the structure of opportunity, making experience more valuable, entry harder, and early-career pathways more fragile.

The same threshold logic may extend to other forms of knowledge work. In creative writing, AI generates fluent prose, but producing a compelling narrative requires sustained judgment—deciding which characters matter, which details to keep, and which ideas to discard (*10*). In science, AI increases research output while weakening traditional quality signals, such as writing style and presentation, thereby raising the bar for identifying high-quality work (*11*). Across domains, AI is better at amplifying human judgment than at replacing it, rewarding those who can evaluate outputs and decide how to direct them toward useful ends.

By making execution cheap, AI shifts value upstream toward cognitive tasks—judgment, problem framing, and integration—that are unevenly distributed (*12*) and slow to acquire (*13*). As a result, access to AI alone is unlikely to equalize outcomes. What matters is not whether workers can use AI but whether they can turn its output into useful work. Understanding how AI shifts—rather than removes—productivity is therefore essential, not only for hiring and management but also for guiding policy responses to technological change, including experiments with social safety nets (*14*). □

REFERENCES AND NOTES

1. E. Brynjolfsson *et al.*, *Q. J. Econ.* **140**, 889 (2025).
2. S. Noy, W. Zhang, *Science* **381**, 187 (2023).
3. S. Daniotti *et al.*, *Science* **391**, 831 (2026).
4. J. Becker *et al.*, https://doi.org/10.48550/arXiv.2507.09089 (2025).
5. S. Nguyen *et al.*, "How beginning programmers and code LLMs (mis)read each other" in *Proceedings of the 2024 CHI Conference on Human Factors in Computing Systems*, F. F. Mueller *et al.*, Eds. (Association for Computing Machinery, 2024), pp. 1–26.
6. H. He *et al.*, https://arxiv.org/abs/2511.04427v3 (2026).
7. S. M. Hosseini Maasoum, G. Lichtinger, Generative AI as Seniority-Biased Technological Change: Evidence from U.S. Résumé and Job Posting Data (Social Science Research Network, 2025); https://doi.org/10.2139/ssrn.5425555.
8. M. R. Frank *et al.*, https://doi.org/10.48550/arXiv.2601.02554 (2026).
9. E. Brynjolfsson *et al.*, "Canaries in the coal mine? Six facts about the recent employment effects of artificial intelligence" (2025); https://digitaleconomy.stanford.edu/app/uploads/2025/12/CanariesintheCoalMine_Nov25.pdf.
10. T. Chiang, "Why A.I. Isn't Going to Make Art," *The New Yorker*, 31 August 2024; https://www.newyorker.com/culture/the-weekend-essay/why-ai-isnt-going-to-make-art.
11. K. Kusumegi *et al.*, *Science* **390**, 1240 (2025).
12. Y. Lin *et al.*, *Nature* **623**, 987 (2023).
13. F. Xu *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **119**, e2200927119 (2022).
14. OpenResearch, Unconditional Cash Study, (2024); https://www.openresearchlab.org/studies/unconditional-cash-study/study.

ACKNOWLEDGMENTS

L.W. and B.V. acknowledge support from the National Institute of General Medical Sciences of the National Institutes of Health (R01GM164731).

10.1126/science.aef5239

[1]School of Computing and Information, University of Pittsburgh, Pittsburgh, PA, USA. [2]School of Computer Science, Carnegie Mellon University, Pittsburgh, PA, USA. Email: liw105@pitt.edu

DIABETES

Cyborg pancreatic islet organoids

Bio-nanoelectronic islets are new tools for diabetes research and therapy

Jochen Lang and **Matthieu Raoux**

Stem cells can be manipulated to form organoids—three-dimensional structures that mimic the vital functions of organs. Organoids are useful for research into developmental, physiological, and pathophysiological mechanisms, and for improving therapies. The ability to monitor and steer organoid maturation in situ, noninvasively, in real time, and over months has the potential to transform biomedical research and regenerative medicine but poses considerable biotechnological challenges (*1*). On page 786 of this issue, Li *et al.* (*2*) report the differentiation of human pluripotent stem cells into pancreatic islet micro-organs containing microelectrodes. These "cyborg pancreatic organoids" enable cell-specific long-term monitoring of islet cell activities, opening new avenues in diabetes research and cell therapy.

The…platform enabled gene expression profiling and electrical recording at multiple sites at the same time…

Pancreatic islets are complex networks of endocrine cells, including glucagon-secreting α cells and insulin-secreting β cells, that raise and lower blood glucose concentrations, respectively (*3*, *4*). The destruction or dysfunction of β cells, as well as α-cell dysregulation, are present in both type 1 and type 2 diabetes (*3*, *4*), which are chronic diseases that affect 10% of people worldwide. Islet cells use bioelectric signals to transduce nutrient and neuro-hormonal inputs into secretion of hormones. These electric signals, which are difficult to detect with extracellular probes, have different durations and amplitudes compared with those of neurons and cardiomyocytes (*5*). Ever since extracellular electrodes were used to record electrical activity in islets isolated from rodents and humans (primary islets) (*6*), substantial progress has been made using planar microelectrode arrays (*7*) and organic electrochemical transistor arrays (*5*). These tools have also been used to study islets derived from the differentiation of human stem cells in culture (*8*). Li *et al.* combined measurements of gene expression (transcriptomics) at cellular resolution and electrical field potentials, in an approach they called electro-sequencing. The authors' flexible bio-nanoelectronic platform enabled gene expression profiling and electrical recording at multiple sites at the same time in islet organoids. This approach made it possible to track islet functional maturation over weeks, whereas previ-

AI raises the productivity bar

Lingfei Wu and Bogdan Vasilescu

Science **391** (6787), . DOI: 10.1126/science.aef5239

Science (ISSN 1095-9203) is published by the American Association for the Advancement of Science. 1200 New York Avenue NW, Washington, DC 20005. The title *Science* is a registered trademark of AAAS.