JOIBS: November 2025. ISSN 2992-9253

A Dramatic Underestimate of the Performance Gap Between Elite Male and Female Athletes

Doug Rohrer, University of South Florida, USA. E-mail: drohrer@usf.edu

Author's Notes: The author thanks April Bleske-Rechek, Lee Jussim, and Joachim Krueger for their helpful comments on the original version of this paper.

Funding: The author received no specific funding for this work.

Competing interests: The author has declared that he has no competing interests.

Citation: Rohrer, D. (2025). A dramatic underestimate of the performance gap between elite male and female athletes. *Journal of Open Inquiry in the Behavioral Sciences, 4*(5). https:/doi.org/10.58408/issn.2992-9253.2025.04.05.0001

Supplementary materials: https://osf.io/qr6te/

Abstract

Strong evidence shows that anatomical and physiological sex differences favor men over women in nearly every athletic event, yet several researchers have published widely disseminated claims of an inherent advantage for women in some athletic events. The present study assessed the prevalence of this misconception with a survey of a diverse sample of 300 U.S. adults. For each of four events (1500-m run, 100-m dash, long jump, and 800-m swim), participants were asked, "The world record for professional women is closest to the world record for males of what age?" The correct answer for each event is age 14, yet, for each event, at least 64% of participants responded that the women's world record is closest to the record for males of age 18 or older (mean and median estimate ≈ age 20). This stark misconception suggests that many people in the U.S. and possibly elsewhere underestimate the contribution of biology to the male-female performance gap in athletics and, by extension, underrate the impact of allowing biological males to compete in female-only sports.

Keywords: Misconception, Performance, Gap, Sex, Male, Female, Athletics

Introduction

Elite male athletes outperform elite female athletes in events requiring speed, strength, and endurance, and this performance gap is widely attributed to anatomical and physiological sex differences that emerge with puberty (Cheuvront et al., 2005; Hunter et al., 2023; Seiler et al., 2007; Thibault et al., 2010). To be sure, the best women in any athletic event can easily defeat most men, but elite men markedly outperform elite women in nearly every event. The women's world records are also not improving more quickly than the men's records, as the world record gap ceased narrowing decades ago (Rohrer, 2024; Seiler et al., 2007). For instance, in every Olympic track event, the men's world record has remained approximately 10% quicker than the women's record since the 1980s, and, averaged across events, the men-women gap is slightly greater in 2025 than in 1990 (Figure 1).

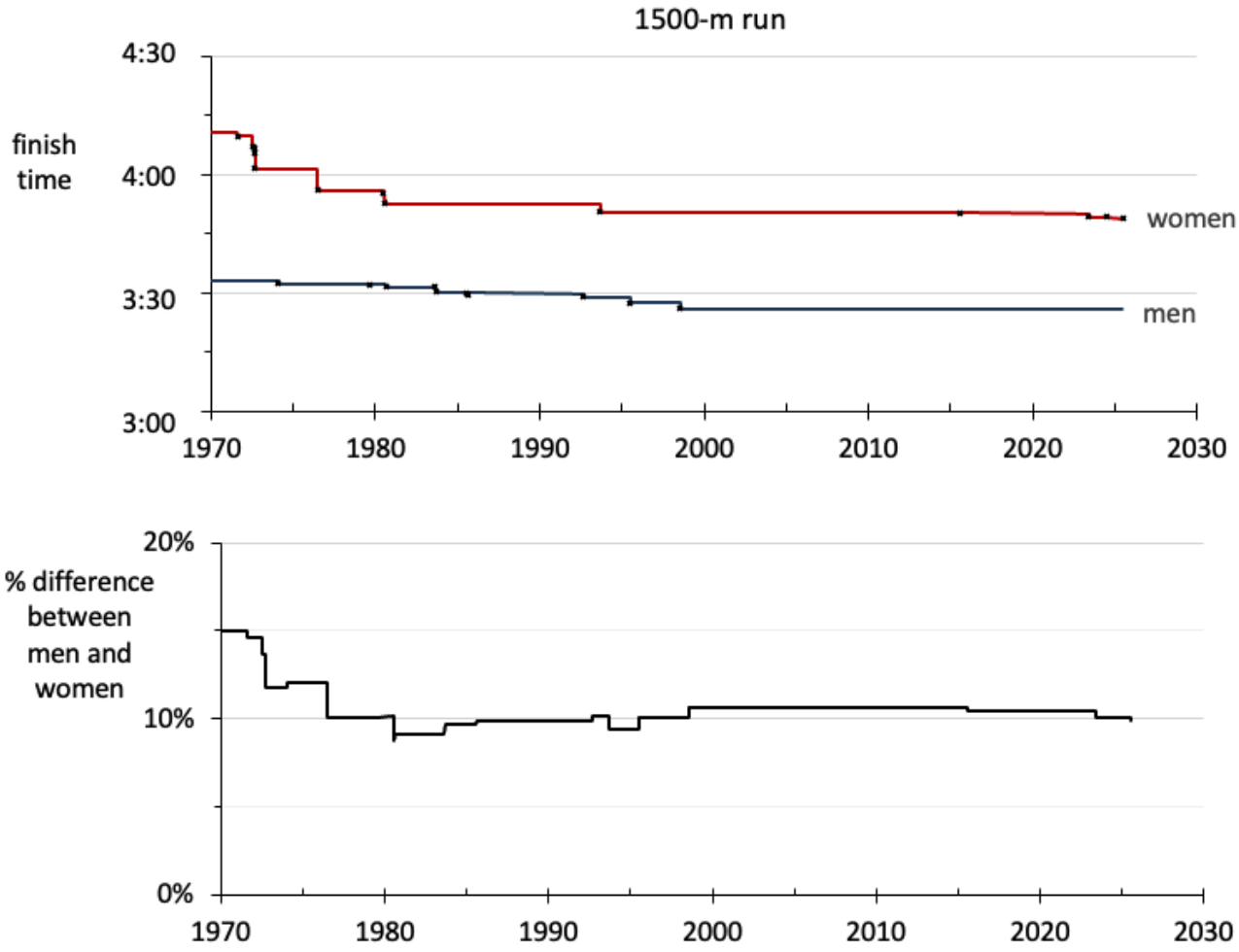

Figure 1. *Progression of world records and the men-women record gap in the 1500-m run.* Percent difference = (women's world record – men's world record) / women's world record. Records span 1.1.1970 through 7.7.2025.

The performance gap between elite men and women athletes is usually expressed as a percentage, but a percentage value can obscure the magnitude of the performance gap. For instance, elite men are about 10% quicker than elite women in most track and swimming events (Table A1), but this gap is larger than some people can appreciate, partly because many nonexperts lack a deep understanding of percentages. Alternatively, the magnitude of the performance gap is conveyed more concretely by comparing the records of elite women to the records of elite young males. For example, in most Olympic track and swimming events, the women's world record is approximately equal to the world record for boys of age 14 (Figure 2).

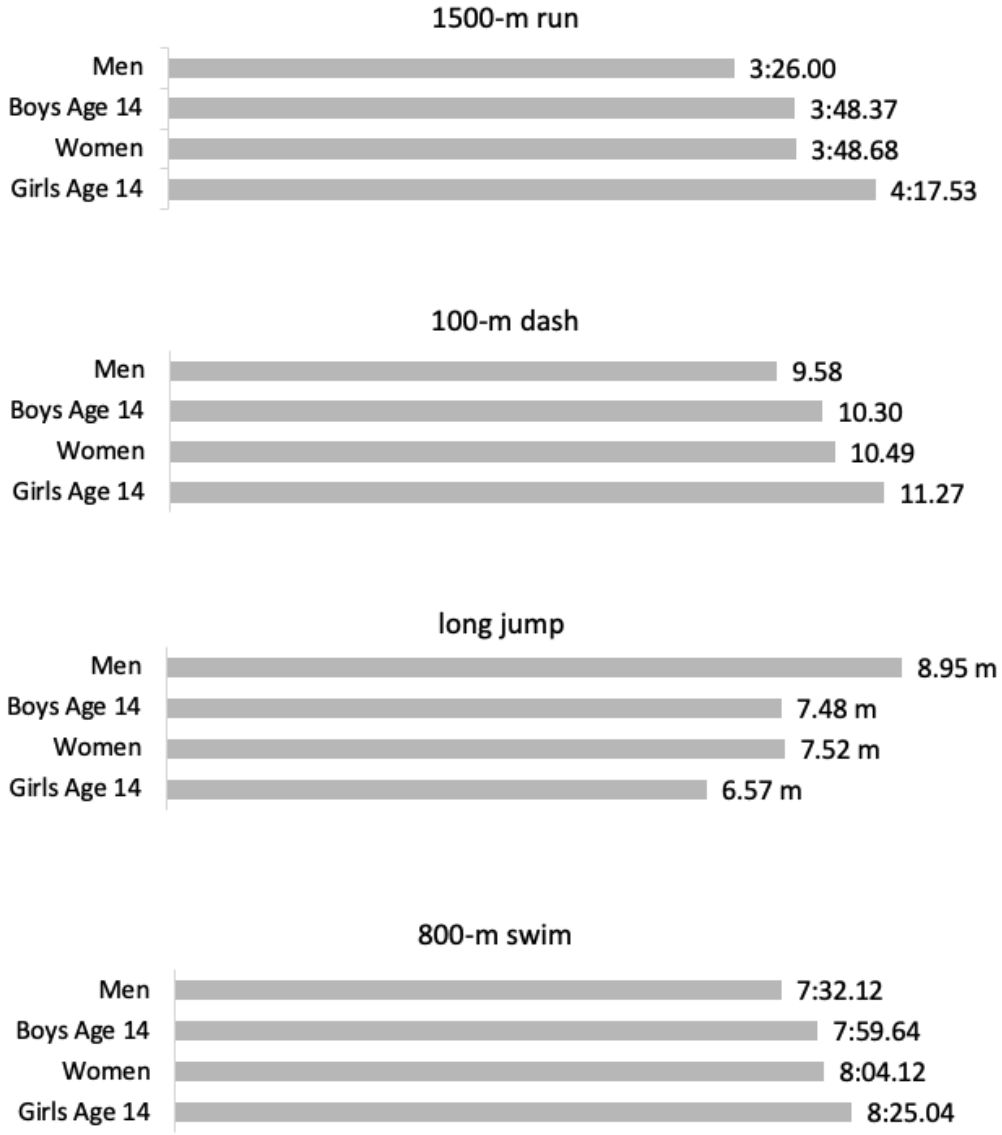

Figure 2. *World records for men, women, and 14-year-old boys and girls in four athletic events.* Additional information appears in Table A1.

Despite the enduring men-women world record gap and the established physiological evidence, some researchers have argued that sex differences favor women over men in certain athletic events. For instance, a few decades ago, two teams of researchers predicted that the men-women world record gap in several running events, including sprints, would narrow and ultimately reverse (Tatem et al., 2004; Whipp & Ward, 1992). Both predictions relied on inappropriate and deceptive analyses, and both predictions failed (Cheuvront et al., 2005; Reinboud, 2004; Rohrer, 2024). In fact, the men-women record gap in every Olympic running event ceased narrowing *before* either prediction was published (Seiler et al., 2007).

Other researchers have asserted that physiological sex differences favor women over men in long endurance contests such as ultramarathons, which are running races of at least 50 km (e.g., Beneke et al., 2005; Ocobock & Lacy, 2023). Yet these authors cited anecdotal race results and listed only those physiological sex differences that favor women over men in endurance events, without noting the male distance runners benefit from longer limbs, less body fat, and higher VO_2 max (Cheuvront et al., 2005; Hunter et al., 2023; Seiler et al., 2007). Of greater importance, the world records show that the men-women record gap is *greater* in ultramarathon events than in short races: the percentage gap is about 10% in track events yet exceeds 13% in most ultramarathon events (Tables A1 and A2). Still other researchers have argued that the men-women performance gap in *every* athletic event has no basis in biology (Bekker & Mumford, 2025). All these claims contradict the prevailing view of experts. In a recent consensus statement from the American College of Sports Medicine, a panel of 10 experts reviewed the evidence and concluded that, “In athletic events and sports relying on endurance, muscle strength, speed, and power, men typically outperform women because of fundamental sex differences dictated by sex chromosomes at puberty, in particular, testosterone” (Hunter et al., 2023, p. 2328).

Some of the published claims of an inherent advantage for women over in some athletic events may reflect bias. For instance, all the claims cited above were supported by cherry-picked evidence or inappropriate analyses (Reinboud, 2004; Rohrer, 2024), which are telltale signs of researcher bias (Gelman, 2018; Jussim & Honeycutt, 2023; Wicherts, 2017). The publishing of these claims also might signify either the bias of the reviewers who recommended publication or the bias of the editors who published these claims without inviting peer review (e.g., Clark et al., 2023; Jindra & Sakamoto, 2023).

To be clear, arguments for female superiority have been put forth by only a small number of researchers, often without peer review, but their claims have been widely disseminated through English-language popular media. The aforementioned predictions of elite women overtaking elite men appeared in the scientific journal *Nature* (Tatem et al., 2004; Whipp & Ward, 1992), and the predictions received worldwide attention in the news media. In a recent cover story in the magazine *Scientific American*, two academics contended that women are better suited to endurance events than are men, partly because of estrogen (Ocobock & Lacy, 2023). A story in *The New York Times* cited a research study that found “… the longer the distance of a race, the shorter the gap between the pace of men and women” (Minsberg, 2021). In a *Washington Post* story, a physiologist states that, “Females are closing the performance gap in ultras, and the longer the distance, the smaller the gap” (Loudin, 2020).

The false claims may also distort the responses of some Artificial Intelligence models. In July 2025, the author asked three popular large language models, "Who has greater endurance, men or women?" Each model began its response with a misleading statement: "In general, women tend to have greater endurance than men…" (ChatGPT), "Women generally have greater endurance than men…" (Perplexity), and "The question of who has greater endurance between men and women is complex…" (Gemini).

Are these misconceptions shared by laypeople? The present study tested the preregistered hypothesis that nonexperts will underestimate the men-women performance gap in four athletic events. A diverse sample of 300 U.S. adults completed an online survey. Although the online format allowed participants to cheat by obtaining answers from the internet or another person, cheating would have improved performance and thus worked *against* the hypothesis. For each event, participants were asked: "In [this event], the world record for professional women is closest to the world record for males of what age?" The correct answer for each event is age 14.

Method

Transparency

The study was approved by the Institutional Review Board at the University of South Florida. The hypothesis, method, and analyses were preregistered (https://osf.io/7vgt4). Materials and data are publicly available (https://osf.io/qr6te/). All data are reported accurately. None of the writing was plagiarized from work by a nonauthor. The study was conducted twice because the first iteration inadvertently excluded two elements of the preregistered method (question order was not randomized, and participant sample was not stratified by political affiliation). Both iterations showed the same misconception, and its magnitude was unreliably greater in the first iteration.

Participants

A sample of 300 adults was obtained from the company Prolific, an opt-in platform for participants who complete studies in return for payment. Another 15 participants began the survey but did not finish, and their data were excluded from all analyses. Prolific reported that all participants were English-speaking U.S. residents. Prolific also indicated that the sample was stratified by age, sex, and political affiliation (Democrat, Independent, or Republican) in proportion to the population of U.S. adults. The partitioning by age and sex align with the U.S. census, but the partitioning by political affiliation was not verified (Table A3).

Materials

Participants were asked about the 1500-run, 100-m dash, long jump, and 800-m swim, which demand varying mixtures of speed, power, and endurance. Events excluded from consideration include track races with hurdles or barriers (which are higher for men than women), the shotput and discus (which are heavier for men than for women), track races of at least 5000 m or the marathon (in which young teenagers rarely compete), and events in which the women's world record is closest to the record for boys of age 13 rather than the record for boys of age 14, as the former would have likely led to larger errors and thus worked in favor of the hypothesis.

Procedure

Participants were required to complete the survey with a computer or tablet rather than a phone. The order of the four questions was randomized for each participant. Participants were forced to answer each question before they saw the next question, and they could not return to a prior question. Each question was followed by a list of answer choices ranging from ages 1 through 27 in 1-year increments, which means that the middle choice equaled the correct answer of 14. The survey concluded with an attention check item that all participants answered correctly. Screenshots of the survey appear in the supplementary materials (https://osf.io/qr6te/).

Results

For each of the four events, 91%–94% of participants indicated that the women's world record is closest to the world record for males of an age that is older than the correct answer of age 14, and 64%–70% chose an age of 18 or older (Table 1, Figure 3).

Table 1

Participants' Estimates (years of age)

	mean	SD	Cohen's *d*	median	interquartile range
1500-m run	20.08	4.09	1.49	20	16 – 24
100-m dash	19.68	4.20	1.35	19	16 – 23
long jump	19.74	4.12	1.39	19	16 – 23
800-m swim	19.85	3.87	1.51	20	17 – 23

Note. Cohen's *d* = difference between correct answer (14) and the mean, measured in SD units.

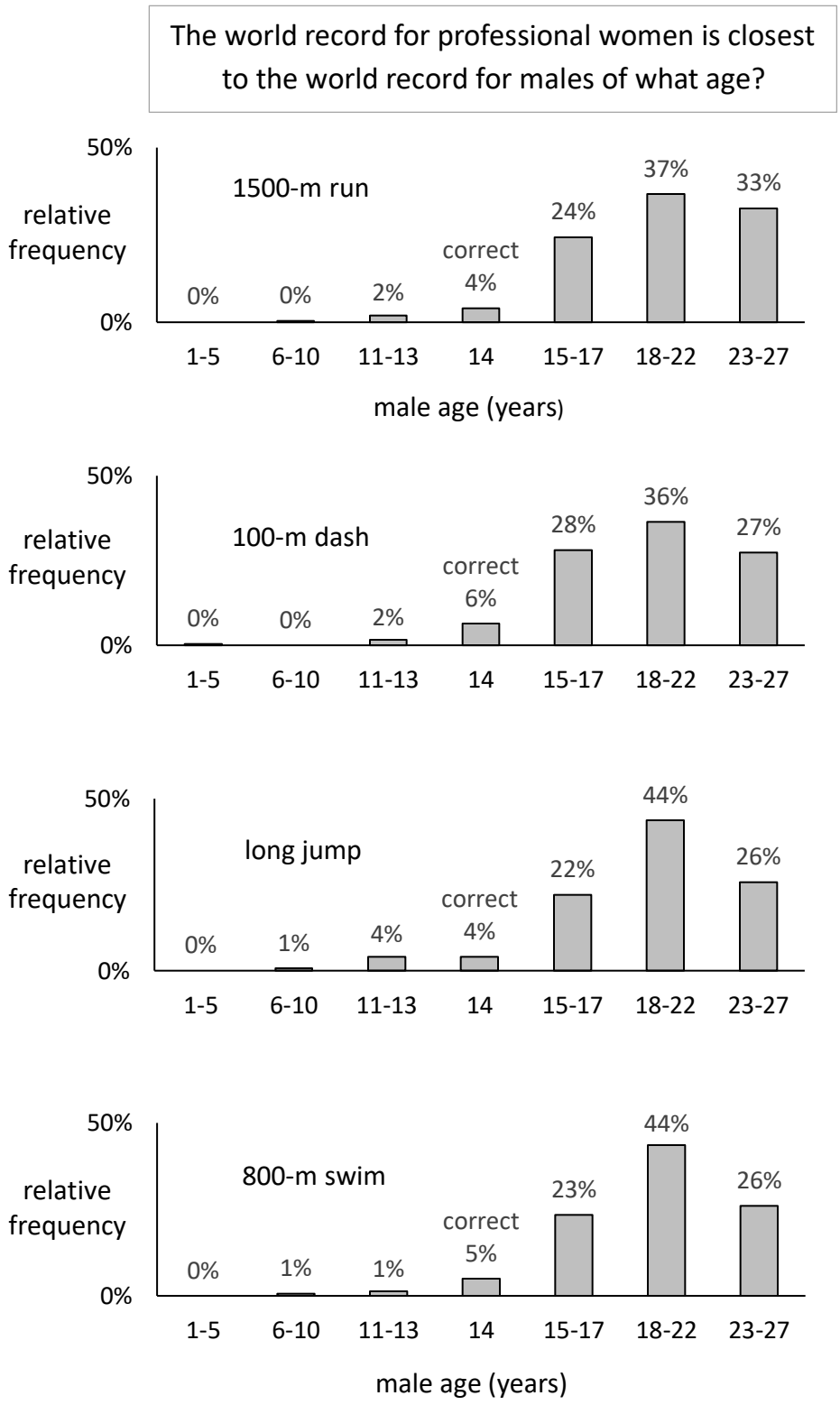

Figure 3. *Participants' estimates*. Values are rounded to the nearest one percent and do not necessarily sum to 100%. Summary statistics appear in Table 1.

Participants who gave large overestimates in one event typically gave large overestimates in other events (Table 2).

Table 2

Pearson Correlations between Participants' Estimates for each Pair of Events (n = 300)

	1500-m run	100-m dash	long jump
1500-m run	—		
100-m dash	.62**	—	
long jump	.66**	.63**	—
800-m swim	.64**	.55**	.63**

** $p < .01$

Correlations between participants' estimates and participants' age, sex, and political affiliation were small or null (Table 3). Their estimates did not correlate with their age, and the overestimate was slightly smaller for men (M = 19.4 years) than for women (M = 20.3 years). Errors were also smaller for Republicans (M = 18.9 years) than for Independents (M = 20.1 years) and Democrats (M = 20.4 years).

Table 3

Pearson Correlations between Participants' Estimates and their Sex, Age, and Political Affiliation (n = 300)

	1500-m run	100-m dash	long jump	800-m swim
Age	−.01	.00	−.06	−.06
Sex (1 = men 2 = women)	.10	.14*	.10	.10
Political Affiliation (1 = Dem, 2 = Ind, 3 = Rep)	−.17**	−.12*	−.11	−.15**

* $p < .05$

** $p < .01$

Discussion

For each of four athletic events, about 2 in 3 participants indicated that the women's world record is closest to the world record for men of age 18 or older, and the mean estimate of about age 20 exceeded the correct answer of age 14 by nearly 1.5 standard deviations. This glaring misconception suggests that many people in the U.S. and perhaps elsewhere underestimate the performance gap between elite male and female athletes, possibly because they do not fully appreciate the anatomical and physiological sex differences that underlie this gap.

The misconception observed here has several explanations, none of which exclude another. As detailed in the Introduction, the misconception might partly reflect the broad dissemination of some researchers' claims that females have an inherent advantage over males in some athletic events. The misconception also might be due to the broader misconception that nearly all sex differences in behavior or ability have little or no basis in biology. Or some people might hold the default view that men and women have generally equal abilities. Whatever the cause, the present finding provides no evidence for any causal relationship.

The misconception reported here may have limited generality. For instance, despite the stratification of the participant sample by age, sex, and political affiliation, the sample may not represent the population in other respects, such as educational level or wealth. Indeed, few wealthy people might choose to complete surveys in return for low pay. The misconception also might not generalize to domains in which elite men outperform elite women, including some athletic events and even nonathletic contests such as chess.

Perhaps the most important caveat is that the present study measured the men-women performance gap with only one question (elite males of what age perform as well as elite women?). This question does not allow participants to indicate that the best woman outperforms the best man and thus might have worked in favor of the observed finding. To remedy this, participants could have instead estimated the men-women performance gap by providing a percentage (e.g., men are 10% quicker than women, or women are 10% quicker than men). However, this measure has drawbacks as well. Providing a percentage requires that participants have a conceptual understanding of percentages, which many adults do not have. Another approach is to have participants estimate the *absolute* difference between the men's and women's world records (e.g., men are 20 seconds quicker in the 1500-m run), but the correct answer to this question varies by event and requires familiarity with the performance of elite athletes in the event. For example, when estimating the men-women world record gap in the 800-m swim, providing an answer close to the correct answer of 32 s is easier if one knows that elite swimmers complete this race in about 8 minutes rather than, say, 25 minutes. That is, a gap of 32 s is relatively large if the event lasts only 8 minutes (32 s $\rightarrow$ 7% gap), but it is relatively small if the race lasts 25 minutes (32 s $\rightarrow$ 2% gap). In short, it might be that any measure has its weaknesses, but the use of multiple measures nevertheless might have provided a richer understanding of the misconception.

Finally, the misconception reported here has obvious implications for the controversy surrounding biological males who identify as females and compete in female-only sports, as people who underestimate the male-female performance gap in athletics might also underrate the impact of allowing males to compete in female-only sports. This misconception observed here also raises the possibility that some people might alter their views on this contentious issue if

they fully appreciated the magnitude of the male-female performance gap in athletics. For example, those who believe that athletic competition between 13-year-old and 17-year-old boys is a mismatch might be more likely to support segregation by biological sex if they knew that the performance gap between 13- and 17-year-old boys is *smaller* than the gap between boys and girls of age 17. To be sure, this politically charged issue is a question of policy rather than science, but the public can better evaluate the impact of a policy when scientists faithfully report the evidence.

References

Bekker, S., & Mumford, S. (2025). Open play: The case for feminist sport. Reaktion.

Beneke, R., Leithäuser, R. M., & Doppelmayr, M. (2005). Women will do it in the long run. *British Journal of Sports Medicine*, *39*, 410. https://doi.org/10.1136/bjsm.2004.017574

Cheuvront, S. N., Carter, R., DeRuisseau, K. C., & Moffatt, R. J. (2005). Running performance differences between men and women. *Sports Medicine*, *35*, 1017–1024. https://doi.org/10.2165/00007256-200535120-00002

Clark, C. J., Jussim, L., Frey, K., Stevens, S. T., al-Gharbi, M., Aquino, K., Bailey, J. M., Barbaro, N., Baumeister, R. F., Bleske-Rechek, A., Buss., D., Ceci, S., Del Giudice, M., Ditto, P. H., Forgas, J. P., Geary, D. C., Geher, G., Haider, S., Honeycutt, N., ... & von Hippel, W. (2023). Prosocial motives underlie scientific censorship by scientists: A perspective and research agenda. *Proceedings of the National Academy of Sciences*, *120*(48), e2301642120. https://doi.org/10.1073/pnas.2301642120

Gelman, A. (2018). Ethics in statistical practice and communication: Five recommendations. *Significance*, *15(5)*, 40–43. https://doi.org/10.1111/j.1740-9713.2018.01193.x

Hunter, S. K., Angadi, S. S., Bhargava, A., Harper, J., Hirschberg, A. L., Levine, B. D., Moreau, K. L., Nokoff, N. J., Stachenfeld, N. S., & Bermon, S. (2023). The biological basis of sex differences in athletic performance: Consensus statement for the American College of Sports Medicine. *Medicine and Science in Sports and Exercise*, *55*(12), 2328–2360. https://doi.org/10.1249/MSS.0000000000003300

Jindra, M., & Sakamoto, A. (2023). When ideology drives social science: Statistical cherry-picking and malfeasance are rife. *Chronicle of Higher Education*. https://www.chronicle.com/article/when-ideology-drives-social-science

Jussim, L., & Honeycutt, N. (2023). Psychology as science and as propaganda. Psychology *Learning and Teaching*, *22*(3), 237–244. https://doi.org/10.1177/14757257231195347

Loudin, S. (2020, September 26). More women gain ground in ultramarathons, other long-distance races. *Washington Post*. https://www.washingtonpost.com/health/women-long-distance-runners/2020/09/25/bffdda10-1871-11ea-a659-7d69641c6ff7_story.html

Minsberg, T. (2021, July 11). In a 100-mile race that demands both physical and mental fortitude, women set the pace. *New York Times*. https://www.nytimes.com/2021/07/11/sports/western-states-women.html

Ocobock, C., & Lacy, S. (2023). The theory that men evolved to hunt and women evolved to gather is wrong. *Scientific American*, *329(4)*, 22–29. www.scientificamerican.com/article/the-theory-that-men-evolved-to-hunt-and-women-evolved-to-gather-is-wrong1/

Pew Research Center (2025, July 23). The National Public Opinion Reference Survey (NPORS).

https://www.pewresearch.org/methods/fact-sheet/national-public-opinion-reference-survey-npors/

Reinboud, W. (2004). Linear models can't keep up with sport gender gap. *Nature*, *432*, 147. https://doi.org/10.1038/432147a

Rohrer, D. (2024). Researcher bias and the enduring gap between the world's fastest men and women. *Frontiers in Physiology*, *15*, 1360731. https://doi.org/10.3389/fphys.2024.1360731

Seiler, S., De Koning, J. J., & Foster, C. (2007). The fall and rise of the gender difference in elite anaerobic performance 1952-2006. *Medicine and Science in Sports and Exercise*, *39*(3), 534–540. https://doi.org/10.1249/01.mss.0000247005.17342.2b

Tatem, A. J., Guerra, C. A., Atkinson, P. M., & Hay, S. I. (2004). Momentous sprint at the 2156 Olympics? *Nature*, *431*(7008), 525. https://doi.org/10.1038/431525a

Thibault, V., Guillaume, M., Berthelot, G., El Helou, N., Schaal, K., Quinquis, L., Nassif, H., Tafflet, M., Escolano, S., Hermine, O., & Toussaint, J.-F. (2010). Women and men in sport performance: the gender gap has not evolved since 1983. *Journal of Sports Science & Medicine*, *9*, 214–223.

U.S. Census Bureau (2021). Sex by age. Table B01001. *American Community Survey*. https://data.census.gov/table/ACSDT5YSPT2021.B01001?t=-00. Accessed 5 July 2025.

Whipp, B. J., & Ward, S (1992). Will women soon outrun men? *Nature*, *355(6355)*, 25. https://doi.org/10.1038/355025a0

Wicherts, J. M. (2017). The weak spots in contemporary science (and how to fix them). *Animals 7*(12), 90. https://doi.org/10.3390/ani7120090