nature human behaviour

ARTICLES
https://doi.org/10.1038/s41562-020-00959-9

Check for updates

Learning is inhibited by heat exposure, both internationally and within the United States

R. Jisung Park [1]✉, Joshua Goodman [2]✉ and A. Patrick Behrer [3]✉

Human capital generally, and cognitive skills specifically, play a crucial role in determining economic mobility and macroeconomic growth. While elevated temperatures have been shown to impair short-run cognitive performance, much less is known about whether heat exposure affects the rate of skill formation. We combine standardized achievement data for 58 countries and 12,000 US school districts with detailed weather and academic calendar information to show that the rate of learning decreases with an increase in the number of hot school days. These results provide evidence that climatic differences may contribute to differences in educational achievement both across countries and within countries by socioeconomic status and that may have important implications for the magnitude and functional form of climate damages in coupled human–natural systems.

Both across and within countries, people living in hotter climates complete less formal schooling, score lower on standardized tests and exhibit worse economic outcomes than those living in cooler climates[1–3]. Such associations are important given the growing role of cognitive skill in income mobility and economic growth[4–7], and because of current and expected changes to the Earth's climate, which appear to influence macroeconomic growth[8]. Whether and how climatic factors causally affect human capital development remain, however, debated, in part because so many other institutional and economic factors are correlated with a warmer historical climate. Some argue that initial conditions during colonization influenced the institutions created in hotter, more disease-prone climates, leading to lower levels of human capital today[9–11]. Others emphasize the role of correlated impediments to agricultural productivity[12,13], disease burden[14] and child nutrition and health[15], which may in turn change the incentive to pursue schooling[16,17].

We propose a more direct mechanism that may operate alongside institutional, agricultural or other factors. Across a range of laboratory and field environments, temperature has been shown to affect working memory, stamina and cognitive performance[18–20], and to lead individuals to reduce time spent engaging in labour activities[21]. This suggests that, in addition to the channels above, heat may directly affect students' capacity to learn or teachers' capacity to teach. Given vast international differences in thermal conditions experienced by students (Table 1), even small marginal effects of heat on learning could result in large educational disparities over time. Students in Indonesia and Thailand, for instance, experience over 200 days above 26.7 °C (80 °F) per school year, compared with approximately 40 such days in the United States and South Korea. Causal estimates of the returns to schooling suggest that small changes in educational achievement can result in persistent differences in lifetime earnings potential[22]. There is, however, limited evidence on how heat exposure affects the rate of learning and human capital accumulation in the context of formal schooling[3,23].

We provide evidence that heat exposure during learning periods directly impacts human capital accumulation, suggesting another channel through which climate is linked to macroeconomic development. To do so, we provide two sets of analyses, each using quasi-experimental research designs and incorporating region-specific academic calendars to measure temperature shocks that occur on school days preceding cognitive testing. The empirical designs focus on heat exposure during the school year (as opposed to momentary reductions in cognitive performance due to temperature on the day of assessment) and exploit year-to-year variation in weather within a given region to isolate the causal impact of hotter school years on learning.

The first analysis uses test score data from 58 developed and developing countries participating in the Programme for International Student Assessment (PISA) between 2000 and 2015. PISA's tests are designed to measure formal learning in mathematics, reading and science in nationally representative samples of 15 year olds. We find that students in school during hotter periods score worse on these exams than their peers in the same country who are schooled in cooler periods. The effect of years with more hot days (above 26.7 °C) on subsequent performance persists even when adding controls for changes in economic conditions (for example, per-capita income) and possible spurious correlation between regional time trends in warming and educational performance. To isolate the causal impact of heat exposure on learning, we link within-country temperature fluctuations over time to within-country fluctuations in test scores, controlling for country- and time-varying confounds. Regression equations, identifying assumptions and a series of robustness checks are presented in the Methods and Supplementary Information.

Exploiting variation in the timing of hot days within a given calendar year, we provide suggestive evidence on the potential mechanisms at play. Heat on school days before PISA exams lowers test scores while heat on non-school days (for example, weekends and summer vacation) has little effect, consistent with our hypothesis that heat directly interferes with learning time. In addition, including controls for potential correlated shocks to agricultural yields does not affect the magnitude or significance of these findings. Specifically, the effects are robust to controlling for hot days during region-specific rice-growing seasons as well as time-varying, country-level measures of agricultural employment, suggesting that the effects of hot temperature are not driven solely by correlated

[1]Department of Public Policy, University of California, Los Angeles, CA, USA. [2]Wheelock College of Education and Human Development / Department of Economics, Boston University, Boston, MA, USA. [3]Center for Food Security and the Environment, Stanford University, Palo Alto, CA, USA.
✉e-mail: jisungpark@luskin.ucla.edu; edecjosh@bu.edu; abehrer@stanford.edu

Table 1 | Heat exposure in selected PISA countries

Country	School days > 26.7 °C (80 °F)	Per-capita income (US$)	Average PISA score
Indonesia	243	2,180	−1.17
Thailand	203	3,937	−0.76
Brazil	122	7,043	−1.08
Mexico	140	8,160	−0.87
Vietnam	118	1,894	0.09
Israel	83	27,759	−0.40
United States	44	46,247	−0.08
South Korea	36	19,467	0.35
Spain	20	25,224	−0.14
Turkey	20	8,899	−0.59
France	11	34,616	−0.01
Netherlands	5	45,164	0.18

The school day measures report the average annual number of school days with a temperature above 26.7 °C (80 °F) experienced by each country during our sample period from 1995–2015. Per-capita income reports the average per-capita income in constant US$ over the same time period using data from the World Bank. The normalized PISA score reports the average normalized overall PISA score within each country over our sample period.

shocks to nutrition or time reallocation decisions in response to correlated changes in economic incentives to pursue schooling.

Even with a rich set of controls, the range of countries in our data implies that these effects could be driven by other correlated mechanisms noted above, particularly in lower-income, agrarian economies. The second analysis therefore focuses on the United States, a highly developed, non-agrarian setting where nutrition and agricultural income-related channels seem less likely to be empirically first order in explaining the impact of heat on achievement. We use district-level annual mathematics and English language arts (ELA) test scores from over 12,000 US school districts, taken from the Stanford Education Data Archive (SEDA). These tests are mandatory components of school accountability systems, so that the sample of test takers represents the near-universe of American students. The tests are deliberately aligned with school curricula to measure learning that is meant to occur during formal schooling. Similar to the international data, we link within-district temperature fluctuations over time to within-district fluctuations in test scores to isolate the causal effect of hotter temperature during the school year.

We find that US students in school during hotter years score worse than peers in the same district schooled during cooler periods. Consistent with the international evidence and the hypothesis that heat interferes with learning, we find that heat on school days entirely drives our results. These results are robust to the inclusion of controls for district-level changes in school funding and demographic composition, potential spurious correlation between regional warming patterns and trends in educational achievement, and controls for exam-day temperature.

Across both sets of analyses, we find that the marginal damage associated with hotter temperature appears to be larger for lower-income populations, consistent with previous work on climate adaptation[24]. These results suggest that the effects of hot temperature may be regressive not only across but also within countries, consistent with recent work[3,25]. In the United States, heat's effects appear to be larger for racial or ethnic minorities and students in lower-income school districts, who probably have less access to potentially compensatory resources. We also present evidence suggesting that the effect of heat exposure during learning periods on achievement is larger for younger students. The effect of heat on children may be more pronounced if children rely more heavily than adults on well-functioning institutions to enable effective avoidance behaviours or carry out necessary protective investments. These and other reasons suggest that children may be more susceptible to hyperthermia and heat exhaustion[26], but so far there has been little evidence regarding the differential impact of heat exposure on learning across age groups.

We note three observations about these analyses. First, they study the impact of heat on learning, rather than momentary reductions in cognition that may arise from temperature stress. Existing evidence suggests that many factors, including temperature[20,23], air pollution[27], sleep deprivation[28] and attentional capture[29], can affect short-run cognition. The mechanism studied here does not operate through such short-term reductions in cognition during test taking or in the immediate lead up to test taking, and controls for the possibility of correlation between heat exposure during learning periods and hot temperature during a subsequent exam. The outcome measures are standardized assessments designed to capture cumulative learning throughout formal schooling, as opposed to tests of raw intelligence or cognitive capacity that are highly sensitive to test-taking conditions, in contrast with previous studies[23].

Second, these results encompass students in both the developing and developed world, presumably with varying levels of adaptation investment. Previous studies found that the effects of climatic shocks on health and economic outcomes vary substantially by income or previous exposure[2,24,30], and that investments such as air conditioning may be effective at mitigating heat-related impacts[31]. Given vast differences in the rate of air conditioning across countries, and notably between the United States and most other countries, it is important to assess the external validity of existing United States-based findings[3,23]. Recent survey evidence suggests that, whereas 90% of US households have some form of air conditioning, only 75, 19 and 13% of households in Australia, Sweden and Mexico, respectively, have air conditioning[32,33]. This study suggests that the smaller macro-level effects of temperature documented in developed economies[30] may mask substantial heterogeneity within these countries. Third, we suggest a seemingly universal physiological channel through which heat affects human capital accumulation, in contrast with an older and in some cases racially charged literature arguing that the association between climate and human capital is driven by genetic or cultural factors. Such literature claimed that those living in tropical countries were genetically and culturally lazy or otherwise disinclined to engage in cognitively intensive activities[34,35]. The unfortunate implications of this work may have inhibited discussion of a simpler and more policy-relevant explanation for the observed associations between heat and human capital. We suggest that the universal physiological burden of heat reduces students' capacity to learn and teachers' capacity to teach, independent of intelligence or disposition. Hotter climates may thus interfere with economic development by reducing the human capital stock of nations, which implies that investments aimed at protecting students from heat exposure may confer important economic benefits, particularly in hotter, poorer countries.

Results

International analysis. Our first analysis explored the relationship between heat exposure and standardized PISA performance. The sample comprised exam records from 58 countries that participated in PISA, which is administered by the Organisation for Economic Co-operation and Development and has provided internationally harmonized exams to nationally representative samples of 15 year olds every 3 years since 2000.

Our sample spanned a wide range of incomes and average climates, including poor tropical countries such as Vietnam and Thailand, as well as many richer temperate countries such as South Korea, France and New Zealand. The average per-capita income

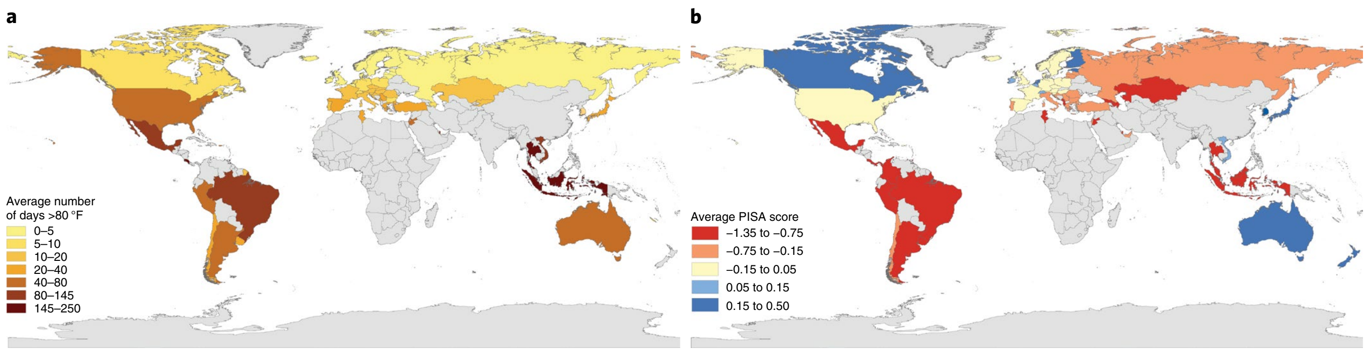

Fig. 1 | Temperature exposure and PISA scores. a, Average number of days with a maximum temperature above 26.7 °C (80 °F) over our sample period (1995–2015). **b**, Average normalized PISA score in terms of standard deviations across all subjects and years that countries reported scores for during our sample period. There were 67 countries reporting at least one PISA score in our sample period. In **a** and **b**, the coloured countries are those that take the PISA exam. World basemap reproduced with permission from Esri.

across the countries in our sample was US$25,962 in current US dollars (Supplementary Table 1), with some as low as $662 per capita (Kyrgyz Republic) and some as high as $80,857 (Luxembourg). The countries in our sample are plotted in Fig. 1, and represent approximately 144 million 15–19 year olds across the participating countries.

Our empirical design leveraged random variation in temperature within a given country over multiple years. While unobserved determinants of student achievement may be correlated with average climate in the cross-section, year-to-year fluctuations in temperature within a country are plausibly random, particularly when adjusting for correlated global or regional trends in warming and development. Our strategy compared deviations from country-specific averages in PISA performance with deviations from country-specific average temperature, controlling flexibly for other time-varying factors including precipitation and share of labour force in agriculture. We focused on the impact of the number of days with temperatures above 26.7 °C (80 °F), noting that previous studies of heat on cognitive performance and other behavioural outcomes found adverse impacts beginning around 26.7 °C (80 °F) (refs. [3,20,23]). Additional details regarding the data and empirical strategy are provided in the Supplementary Information.

We found that hotter temperatures in the years leading up to the PISA exam negatively impacted student performance. Each additional day above 26.7 °C (80 °F) during the 3 years preceding an exam lowered scores by 0.18% of a standard deviation ($P = 0.007$; 95% confidence interval (CI) = −0.22 to −0.04; Fig. 2). We measured hot days over 3 years to maintain consistency with the periodicity of the PISA exams. A one-standard-deviation increase in hot days conditional on country and year fixed effects amounted to 14 school days. Cold days had statistically insignificant impacts on performance ($\beta = 0.07$; $P = 0.517$; 95% CI = −0.14 to 0.28). These results were robust to the inclusion of continent-specific temperature trends, which suggests that they were not driven by spurious correlation between regional warming patterns and long-run trends in educational achievement, as well as specifications that allow for different functional forms of temperature.

To provide evidence on potential mechanisms, we assessed the impact of heat that occurred during three sets of mutually exclusive days of the year for each country in our sample: weekdays during the school year (henceforth, school days), weekends during the school year and summer vacation days. The effect of hot temperature on learning appeared to be driven almost exclusively by hot school days (Fig. 3a). Each additional hot school day lowered scores by 0.22 standard deviations ($P = 0.002$; 95% CI = −0.36 to −0.08; Supplementary Table 2). A Wald test indicated that the difference in the impact of hot school days and hot non-school days ($F_{1,57} = 3.41$; $P = 0.07$) was marginally significant.

To further probe whether heat impacts learning through other correlated shocks, including the effects of heat on agricultural productivity, we ran analyses that controlled for hot days during the rice-growing season, based on the observation that rice is a major staple crop in many of the poorer countries in our sample (Supplementary Table 4). In the countries for which we had data on growing seasons, we found that hot school days, controlling for the number of hot days during the rice-growing season, still appeared to reduce student performance by 0.31% of a standard deviation ($P = 0.013$; 95% CI = −0.55 to −0.07), whereas hot growing season days had statistically insignificant impacts ($\beta = -0.297$; $P = 0.322$; 95% CI = −0.90 to 0.31). Furthermore, the findings are robust to including controls for changes in per-capita income, share of labour force in agriculture, and female labour force participation, suggesting that they are probably not driven solely by correlated shocks to (gender-specific) economic incentives for educational investment[17].

Splitting the sample into richer and poorer countries (above and below the mean per-capita income in 1995 in our sample; listed in Supplementary Table 5), we found that temperature exerts a significant impact in poorer countries ($\beta = -0.14$; $P = 0.001$; 95% CI = −0.22 to −0.07; Fig. 3a) but less so in richer ones ($\beta = -0.024$; $P = 0.733$; 95% CI = −0.17 to 0.12; Fig. 3a), consistent with lower levels of adaptation and/or other channels (for example, conflict[36]) through which heat can affect student outcomes in developing countries.

Taken together, these results provide further evidence consistent with the claim that hotter temperature during learning periods exerts a negative and casual impact on human capital accumulation. While these reduced form effects do not on their own demonstrate the mechanisms through which such impacts arise, they are consistent with the possibility that a proportion of the effect is driven through heat's disruptive impact on learning.

To better understand the extent to which our results are driven by physiological channels, we conducted a second set of analyses using more spatially resolved data from a highly developed, non-agrarian setting (where non-physiological factors are plausibly less influential) and with a richer set of demographic and location-specific characteristics.

US analysis. Our second analysis examined data on standardized student achievement for over 12,000 US school districts between 2009 and 2015 (Fig. 4). Drawn from SEDA[37], these records comprise the near-universe of state-administered standardized mathematics and verbal assessments for third to eighth graders, representing

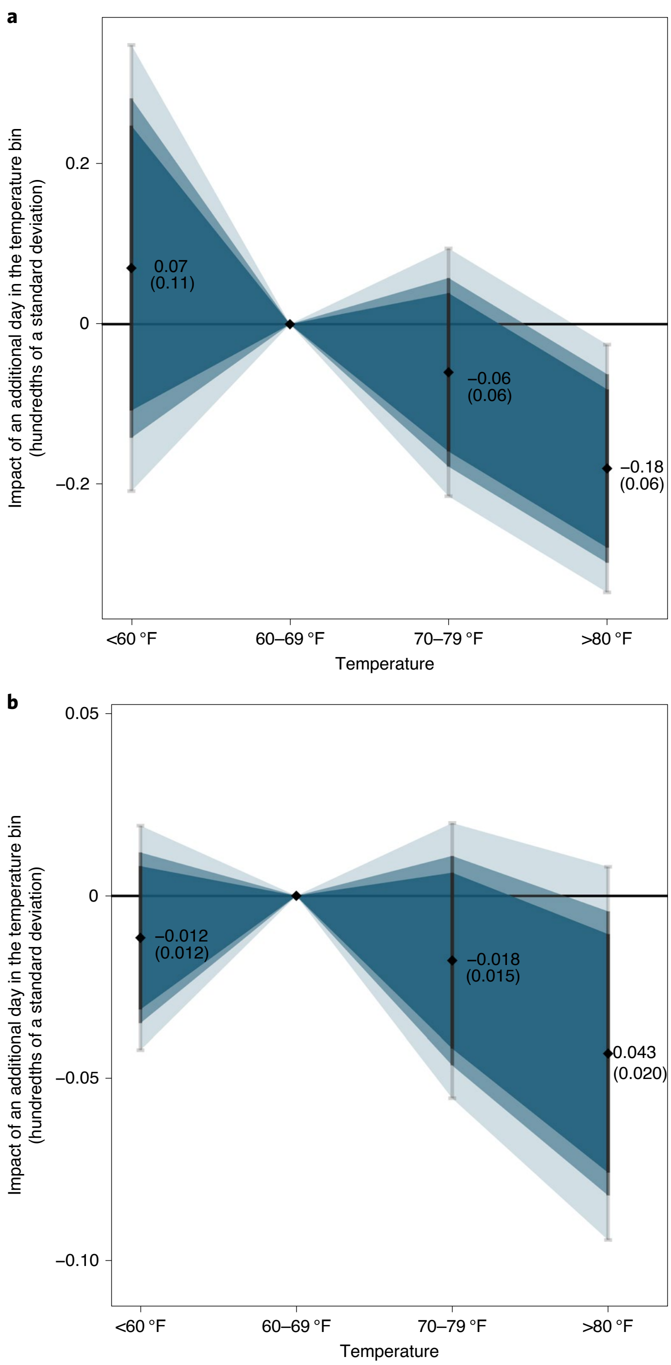

Fig. 2 | Impact of temperature on PISA and SEDA exam scores. **a**, Impact of days below 15.5 °C (+0.07 standard deviations; $P = 0.517$; 95% CI = −0.14 to 0.28), days between 21.1 and 26.7 °C (−0.06 standard deviations; $P = 0.316$; 95% CI = −0.17 to 0.56) and days greater than 26.7 °C (−0.18 standard deviations; $P = 0.007$; 95% CI = −0.31 to −0.05) on performance on the PISA exams. **b**, Impact of days below 15.5 °C (−0.012 standard deviations; $P = 0.335$; 95% CI = −0.035 to 0.012), days between 21.1 and 26.7 °C (−0.018 standard deviations; $P = 0.226$; 95% CI = −0.05 to 0.011) and days greater than 26.7 °C (−0.043 standard deviations; $P = 0.03$; 95% CI = −0.08 to −0.004) on performance on the SEDA exams. In **a** and **b**, the shaded areas connect coefficients representing the effect of an additional school day in each temperature bin on subsequent achievement in hundredths of a standard deviation, with light to dark shading corresponding to 90, 95 and 99% CIs, respectively. Standard errors are in parentheses. Sample sizes are $n = 281$ and $n = 825{,}416$ for **a** and **b**, respectively.

over 270 million test scores. These assessments, typically taken in March, April or May, vary across states but have been standardized by SEDA for national comparability. Similar to PISA exams, these tests are meant to capture cumulative learning specific to each state, grade and subject. These data are thus uniquely suited for assessing the effect of heat during formal instructional periods, in contrast with tests used in other US studies[3,23]. Our unit of observation is at the level of district by grade by subject by year, resulting in approximately 825,000 observations matched to district-level daily weather information using data from approximately 3,400 weather stations from the National Climatic Data Center. To account for possible differences in school year heat arising from regional differences in start and end dates, we used state-specific academic calendars, as represented by the largest urban district in each state.

We again exploit random variation in year-to-year temperature within a given district over time to account for potential correlation between unobserved determinants of educational achievement and average climates across districts. For instance, schools in the Southern United States typically perform worse than schools in the Northeast, but many factors other than climate, including teacher quality and legacies of segregation, may affect this cross-sectional relationship. The number of hot days during any given school year within a particular district, however, is plausibly exogenous, especially when taking aggregate (regional) warming patterns into account.

We found that students who experience hotter temperatures during the school year before their exams exhibit reduced learning. Each additional day with a temperature of 26.7 °C or hotter reduces achievement by approximately 0.04% of a standard deviation (Fig. 2 and Supplementary Table 6; $P = 0.071$; 95% CI = −0.07 to 0). Our measures of significance are robust to correlation in error terms within any given state, which typically holds over 200 school districts. This effect is concentrated among school days, with each additional hot school day lowering achievement by 0.07% of a standard deviation (Fig. 3b; $P = 0.01$; 95% CI = −0.12 to −0.02). Similar to the international analysis, heat on non-school days, such as weekends and summers, had no statistically significant impact on achievement ($\beta = 0.015$; $P = 0.561$; 95% CI = −0.04 to 0.06). A Wald test indicated a significant difference between the impact of hot school days and hot non-school days ($F_{1,3394} = 5.54$; $P = 0.019$). These estimates imply that a student who experiences an additional school week (five school days) with daily maximum temperatures above 26.7 °C will learn 0.35% of a standard deviation less than they otherwise would have during that school year, which is equivalent to reducing teacher quality by about 3–4% of a standard deviation[38].

The impact of heat exposure on learning is not confounded by precipitation, exam-day weather shocks, changing demographic compositions or resource levels of school districts, or spurious correlation between regional warming patterns and trends in educational achievement. That only hot weekdays during the school year reduce learning suggests once again that the set of mechanisms probably includes a reduction in contemporaneous educational inputs, whether in terms of the amount or intensity of learning time.

In the United States, the impact of heat on mathematics achievement is about three times larger than its impact on ELA achievement. Each additional hot school day lowers mathematics scores by 0.11% of a standard deviation (columns 3–4 of Supplementary Table 6b; $P = 0.002$; 95% CI = −0.17 to −0.04) but lowers ELA scores by less than 0.04% of a standard deviation (columns 5–6 of Supplementary Table 6b; $P = 0.099$; 95% CI = −0.09 to 0.01). There is little evidence that heat on non-school days affects achievement in either subject.

Importantly, hot temperature affects disadvantaged students much more than advantaged ones. Heat has substantially larger impacts on the achievement of students in lower-income school districts and little impact in higher-income districts, defined respectively as those in districts with federally subsidized lunch rates below and above 50%. Each additional hot school day lowers achievement in lower-income districts by 0.12% of a standard deviation but has little discernible effect on achievement in higher-income districts (Fig. 3b and Supplementary Table 7; $P=0.002$; 95% CI = −0.19 to −0.04). Each hot school day lowers the achievement of Black and Hispanic students by 0.10–0.12% of a standard deviation but has no statistically significant impact on non-Hispanic white students (Supplementary Table 8; Black students: $P=0.017$; 95% CI = −0.19 to −0.02; Hispanic students: $P=0.012$; 95% CI = −0.21 to −0.03; non-Hispanic white students: $\beta=-0.009$; $P=0.593$; 95% CI = −0.04 to 0.02). One week above 80 °F for the average Black or Hispanic student reduces learning by an amount equivalent to reducing teacher value added by 5–6% of a standard deviation.

The effect of hot school days is also larger for younger students than for older students. Each additional such day lowers the achievement of third to fifth graders by 0.08–0.13% of a standard deviation but has a statistically insignificant impact on those in grades six to eight (Fig. 3b and Supplementary Table 7; $P=0.003$; 95% CI = −0.17 to −0.03). This is consistent with previous evidence suggesting that younger children are likely to be more adversely affected by thermal stress, either due to physiology or behaviour[26]. However, this could be due to other factors, such as the potentially lower prevalence of school air conditioning in elementary schools relative to middle schools.

Discussion

Taken together, these results suggest a different perspective on how climate shapes human cognitive capacity. Thermal conditions in the physical learning environment appear to causally influence cumulative learning: a fact not yet documented in the voluminous literature on cross-country comparisons in student achievement[39]. It appears that heat exposure during the learning period, all else being equal, directly slows the rate of human capital formation, in part through persistent disruptions to the learning process. As noted above, the realized temperature environments facing students across the world vary dramatically, suggesting important implications for our understanding of differences in educational achievement and human capital.

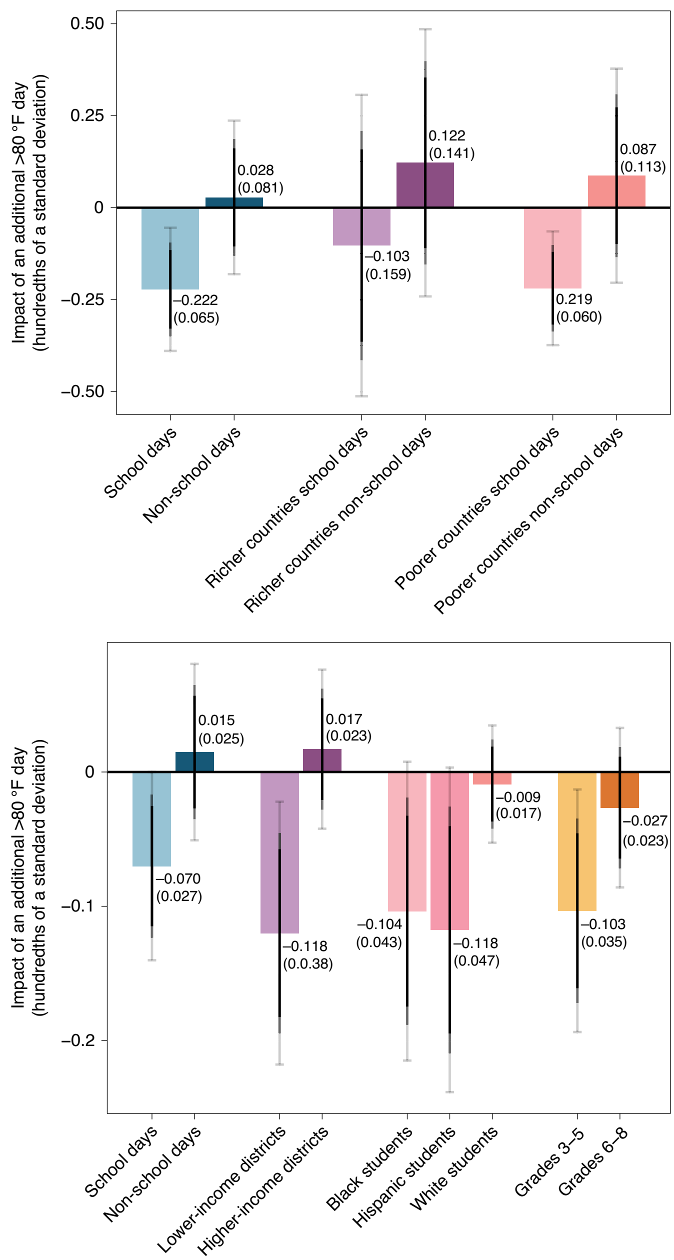

Fig. 3 | Heterogeneity of hot temperature impacts. a, The first two columns show the impact on subsequent standardized achievement of hot (≥26.7 °C) school days (that is, weekdays during the school year) versus hot weekends, holidays and summer vacation days in the 3 years leading up to any given PISA assessment for all participating countries in our sample over the period 2000–2015 ($n=271$). We show the impact of hotter school days (−0.22 standard deviations; $P=0.002$; 95% CI = −0.36 to −0.08) and hot non-school days (+0.03 standard deviations; $P=0.676$; 95% CI = −0.14 to 0.22) in the first two columns. Columns 3–6 show the corresponding effects for countries with above-mean (columns 3 and 4; $n=132$) and below-mean income (columns 5 and 6; $n=150$) in 1995 in our sample. We show the impact of hot school days (−0.099 standard deviations; $P=0.493$; 95% CI = −0.39 to 0.2) and hot non-school days (+0.127 standard deviations; $P=0.394$; 95% CI = −0.18 to 0.43) in richer countries, and of hot school days (−0.256 standard deviations; $P<0.001$; 95% CI = −0.36 to −0.15) and hot non-school days (+0.088 standard deviations; $P=0.342$; 95% CI = −0.10 to 0.28) in poorer countries. **b**, The first two columns show the impact on subsequent standardized achievement of hot (≥80 °F) school days (−0.070 standard deviations; $P=0.01$; 95% CI = −0.12 to −0.02) versus hot weekends, holidays and summer vacation days (+0.015 standard deviations; $P=0.974$; 95% CI = −0.06 to 0.07) for all US school districts over the period 2009–2015 ($n=825{,}416$). Columns 3–9 show the effect of hot school days in schools with a federally subsidized lunch rate below (−0.118 standard deviations; $P=0.002$; 95% CI = −0.19 to −0.04) and above 50% (+0.017 standard deviations; $P=0.456$; 95% CI = −0.03 to 0.06); for Black (−0.104 standard deviations; $P=0.017$; 95% CI = −0.19 to 0.02), Hispanic (−0.118 standard deviations; $P=0.012$; 95% CI = −0.21 to −0.03) and non-Hispanic white students (−0.009 standard deviations; $P=0.593$; 95% CI = −0.04 to 0.02) within each district; and for elementary (−0.103 standard deviations; $P=0.003$; 95% CI = −0.17 to −0.03) and middle-school students (−0.027 standard deviations; $P=0.249$; 95% CI = −0.07 to 0.02) in each district, respectively ($n=273{,}466$; 273,266; 183,060; 222,042; 733,219; 425,301 and 400,095 for columns 3–9, respectively). In all columns in **a** and **b**, the shaded lines show the confidence intervals of our estimates, with light to dark shading corresponding to 90, 95 and 99% CIs, respectively. Consistent with the existing literature (for example, ref. [49]), all coefficients in **a** and **b** can be interpreted as the effect relative to an additional day with a temperature in the range 15.5–21.1 °C (60–70 °F) on combined mathematics, verbal and science scores (**a**) or combined mathematics and ELA scores (**b**). In **a** and **b**, we provide our coefficient estimates and standard errors (in parentheses) beside each bar.

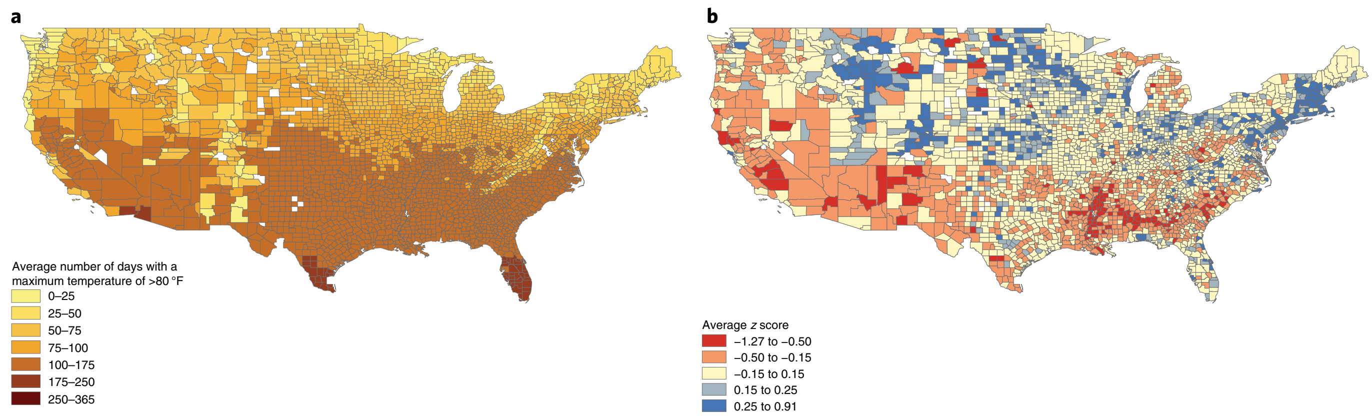

Fig. 4 | Temperature exposure and SEDA scores. a, Average number of days with a maximum temperature above 26.7 °C (2009–2015) in each US county in our sample. **b**, Average *z* score across all subjects and years within each US county in our sample. Missing counties (white) are those omitted from the SEDA data for confidentiality reasons. Basemap data come from the US Census. World basemap reproduced with permission from Esri.

We find heat exposure to be a plausible mechanism. It matches emerging findings on the effects of temperature on labour capacity[40,41], morbidity and mortality[42,43] and short-run cognition[20,23]. However, we note that this analysis does not imply that heat exposure is the only mechanism at play: many others are probably relevant in explaining the relationship between climate and levels of human capital across countries. Teasing apart the potential mechanisms in greater detail (for instance, determining whether hotter temperatures drive student/teacher absenteeism and whether poor nutrition and hunger exacerbate heat-induced cognitive impacts, and understanding the extent to which these mechanisms interact) are important questions for future work.

Importantly, the magnitude of these disruptions appears to vary greatly across socioeconomic groups, both across and within countries. As shown in Fig. 3b, the effect of an additional 26.7 °C day in US school districts in the lower third of average income is approximately −0.12% of a standard deviation ($P = 0.002$; 95% CI = −0.19 to −0.04), while the effect in the top third is statistically indistinguishable from zero. Impacts are also larger for some racial or ethnic minorities, particularly Black and Hispanic students. This is consistent with evidence from the United States suggesting that school and home air conditioning status is correlated with student race/ethnicity and income[3], and suggests that climatic factors may contribute to longstanding racial/ethnic achievement gaps.

How large are these effects? Suppose we take the US estimates as a lower bound for the rest of the world, given relatively high rates of air conditioning there. Education researchers have, for instance, examined the impact of improving teacher quality or reducing class sizes on learning outcomes. Our US analyses suggest that, even with relatively high levels of air conditioning, a school year with 30 additional days above 26.7 °C reduces learning by approximately 2.1% of a standard deviation. This is large enough to offset the gains of reducing class sizes by approximately 3–4%, or to offset improving teacher quality by 20% of a standard deviation. For lower-income students, the effect of the same temperature event appears to be nearly three times larger. These sizable magnitudes suggest that the learning impacts of a hotter climate could result in large real consequences, especially given that students in many tropical economies regularly experience more than 100 such days per school year (Table 1). Put differently, greater heat exposure during the school year may lead students in Brazil to learn 6% less than their South Korean counterparts per year, which, over time, might explain around one-third of the difference in their PISA performance (see Supplementary Information for details of this calculation).

This perspective has important policy implications. It suggests that climate may have a more direct and persistent influence on economic growth than was previously appreciated. Human capital accumulation is central to national economic growth and individual economic mobility[4,7], and current climatic conditions appear to slow the rate of human capital accumulation for some more than others. This suggests that policies aimed at improving physical learning environments, whether in the form of electric infrastructure or low-income energy assistance, may pay larger dividends over time than was previously appreciated. These pro-growth, pro-adaptation policies may or may not include school air conditioning, which may improve student cognition as well as teacher attendance/retention, but which may also exacerbate the climate externality. Making such investments to facilitate learning in hotter environments may be particularly important in light of evidence suggesting that education itself may be an important climate adaptation strategy[44].

It also suggests that current estimates of the social costs of carbon may be understated. Existing integrated assessment models do not include direct impacts on human capital, and often model climate impacts as a non-accumulating reduction in the level of gross domestic product as opposed to cumulative growth rate effects. Adding these arguments to the damage function would probably shift the entire distribution of estimates to be more negative[45]. Accounting for within-country regressivity of these impacts, as suggested by our findings, may also imply larger social costs of carbon estimates, regardless of one's choice of pure rate of time preference or discount rate[46].

Methods

Data description. *Global temperature data.* We used separate temperature datasets for our global and domestic analyses, given varying geographic and temporal coverage. For the global analysis, which used PISA test scores from many different countries, we started with data from the National Oceanic and Atmospheric Administration's Global Historical Climatology Network. This provided us with daily data from a network of more than 100,000 stations located in approximately 180 countries. The data provided included daily maximum and minimum temperatures and total daily precipitation. We collected data starting in 1995 and pulled all of the available data for the countries that appeared in our PISA sample.

To count school versus non-school days, we excluded weekend days from the school days and assigned each country a dummy called summer on the days that students in that country were typically on summer vacation. When schools started on a range of days (for example, the first 2 weeks of September), we choose a date at or adjacent to the midpoint of the range. We separately identified weekends that occurred during the school year and those that occurred during the summer so that we could examine whether heat on non-school days during the school year had different effects from heat on non-school days outside of the school year.

To create our temperature bins, we counted the number of days with a maximum temperature, in 5.5 °C (10 °F) bins, from −17.7 to 60 °C (0–140 °F), by station. We grouped all days below –17.7 °C (0 °F) into a single bin. Each country was then assigned the weighted average number of days (across all stations) in each temperature bin in each year. Weights were based on the population living within 15 km of the station, as measured by LandScan population data. We weighted stations based on their population in 2000, at the beginning of our sample. We also created lagged variables that counted the number of days in each bin in each of five lagged years, as well as the cumulative days in each bin over the previous 1–5 years. The cumulative lag variable did not count the number of days in a given bin in the contemporaneous year. For additional information about the weather data used in the PISA sample, see the Supplementary Information.

US temperature data. Daily temperature data came from the National Oceanic and Atmospheric Administration's Daily Global Historical Climatology Network, which includes station-level data for thousands of weather stations across the United States. We focused on the subset of nearly 3,400 weather stations with daily temperature data available for at least 95% of the days from 1 July 2004 to 30 June 2015 (the time period covering the potential test-taking dates of our sample). Doing so allowed us to assign each school district a single, stable weather station over the entire time period, which avoided endogeneity concerns driven by the possibility that stations coming online or going offline were somehow correlated with local population growth, economic conditions or temperature conditions in ways that might contaminate our estimates[47]. We imputed the small proportion of missing daily observations with those from the nearest stations with non-missing data.

We assigned each school district to the weather station nearest to that district's centroid, resulting in an average distance of 9.6 miles between each district's centroid and the weather station being used to measure temperature at that district. We defined our primary heat exposure variable as the number of days for which the average daily maximum temperature exceeded a given multiple of 5.5 °C (10 °F) from 1 June to 28 February in the year before the test. We used daily maximum temperature because schooling occurs during the daytime when such temperatures usually occur. Of course, to the extent that daytime maximum and night-time minimum temperature is correlated, some of our effects may have been driven by disrupted sleep. We have not taken a stand on whether sleep is a factor or not, as both in-class and at-home disruptions through learning that are brought about by the physiological effects of heat are of interest.

We used the June to February time period because the exact timing of SEDA's standardized exams varies by state and year but almost always occurs between March and May. We focused particularly on temperature experienced on school days, treating non-school days (weekends and all summer days between 15 June and 15 August) as separate sources of variation. We also used the weather stations to construct data on test date temperature, rain and snowfall, as well as cumulative rain and snowfall exposure, over the year before the test, which helped account for potential independent effects of such precipitation.

PISA data. PISA assessments are designed to capture cumulative skills developed during formal schooling (for example, arithmetic, basic scientific concepts and reading comprehension), and to be comparable across countries. Our data on average PISA scores by country came from the National Center for Education Statistics (NCES) International Data Explorer tool. NCES assembles average country scores by year in mathematics, science and reading from the PISA microdata provided by the Organisation for Economic Co-operation and Development. We followed the advice of NCES and did not compare mathematics and science scores from 2000 or 2003 (for science) with scores from later years because of changes in the PISA methodology. We did not modify the raw PISA data from NCES except to drop countries from the sample for which we did not have temperature data or for which we only had 1 year of PISA data. We excluded PISA data from sub-national units (from individual states within the United States, for example). A minimum of 5,000 students were sampled in each country that participates unless the total population of 15-year-old students was less than 5,000, in which case all students were tested. Some large countries sample more students. In total, more than 500,000 students took a PISA exam across all participating countries in 2015.

PISA scores were designed to have a global average of 500 and a student-level standard deviation of 100, which we used to compute standardized versions of each country's mathematics, science and reading scores. In any given year, there was wide variation in performance across countries. On the 2009 PISA exam, for example, South Korean students averaged 546 points in mathematics while Indonesian students averaged 371 points. Our primary outcome measure was the average of each country's three subject scores in any given year, standardized so that effects could be interpreted in terms of student-level standard deviations (similar to SEDA).

Summary statistics for the PISA sample. See Supplementary Table 1 for summary statistics for the PISA sample. On average, countries in our sample were hotter than the United States (experiencing 114 school days over the previous 3 years above 26.7 °C versus only 97 such days in the United States) and poorer (per-capita income = \$26,000 versus \$42,000) with lower PISA scores (normalized score = −0.28 versus −0.08). We also split the sample into rich and poor countries based on where a country's per-capita income in 1995 ranked in our sample. We defined rich as countries that had a per-capita income in 1995 above \$14,000 (roughly the average in our sample for that year). Splitting the sample into rich and poor indicated that the rich sample was substantially cooler and wealthier and had a lower population of test takers than the poor countries. PISA scores were substantially better in the rich sample on average, with lower variance within the sample.

SEDA data. Data from SEDA were based on the standardized accountability tests in mathematics and ELA, administered annually by each state to all public school students in grades 3–8. SEDA combines information on the test scores in each school district with information from the National Assessment of Educational Progress, creating scores that are nationally comparable across districts in different states.

Our version of SEDA's data spanned the school years ending 2009–2015 and contained elementary and middle-school students from approximately 12,000 school districts across all 50 states. We observed a standardized measure of both mathematics and ELA achievement at the district-by-grade-by-year level. We observed this measure averaged across all test takers in a school district, as well as for some demographic subgroups. The particular standardization used implied that effect sizes could be interpreted in student-level standard deviations.

We used SEDA's mean scores across all students and across racial/ethnic subsets of students. We focused on scores for non-Hispanic white, Black and Hispanic students, which SEDA reports separately for all school districts with a sufficient number of such students. Student race/ethnicity was based on the information that state education agencies receive from school districts, which in turn relies on students' self-identification. We split school districts into lower or higher income based on whether more or fewer than half of the students received federally subsidized school lunches, which generally indicated a family income below the poverty line.

Other international data. In addition to temperature and PISA performance data, we collected data on a set of potentially relevant covariates for the countries in our sample. All of these data came from the World Bank's World Development Indicators archive. We collected time-varying measures of the share of total employment in agriculture, per-capita income, the share of male and female employment in agriculture, total population and the share of the population made up by 15–19 year olds. The only data we modified related to the 15- to 19-year-old population share, which we combined with the total population to estimate the absolute number of 15–19 year olds in each country-year. We matched all of the data to temperature and PISA country-years using country International Organization for Standardization codes.

Empirical approach. Our econometric approach exploited the quasi-random variation within a given geography's total exposure to days above 26.7 °C (80 °F) in the years between test takes. The geographic unit in PISA is a country, whereas in SEDA it is a school district. The time between test takes is 3 years in PISA and 1 year in SEDA. To account for serial correlation in temperature shocks across geographies, we clustered standard errors at the relevant geographic unit. In all statistical tests, we assumed normality but did not formally test for it. All tests of significance were two tailed.

We estimated several versions of the base model:

$$\bar{Z}_{it} = \sum_{k=1}^{9} \beta_k \mathbf{TMAX}_{ikg} + \sigma \mathbf{X}_{it} + \gamma_t + \delta_i + \omega_{ct} + \epsilon_{it} \quad (1)$$

where $\bar{Z}_{it}$ is the normalized PISA score in country i and year t. TMAX_{ikg} is the total number of days with a maximum temperature in each of k degree bins in geography i in the gap g between exam takes. $\mathbf{X}_{it}$ is a vector of geography- and year-specific controls, including the total annual precipitation in the year of the exam as well as the gap year(s), the same set of k degree bins in the year of the exam and, in the case of the PISA data, the controls from the World Bank described above. The parameters δ_i and γ_t are geography and year fixed effects. ω_{ct} is a continent-specific time trend included in the PISA regressions. ϵ_{it} is the error term. We weighted each geography by the total number of 15–19 year olds in that country in the exam year in the PISA data, as calculated from the World Bank data, and by the students in each district taking the exam in the SEDA data.

Our variable of interest was β_9 for the bin representing days over 26.7 °C (80 °F). Because we omitted the 15.5–21.1 °C (60–70 °F) bin from our set of controls, the coefficient β_k should be interpreted as exchanging 1 day over the relevant gap in the 15.5–21.1 °C (60–70 °F) bin for one >26.7 °C (80 °F).

Identification rests on the assumption that the number of days in any given temperature bin, and therefore the >26.7 °C (80 °F) bin we are interested in, varies randomly within a geographical region from year to year. This year-to-year variation results in random variation in the aggregate exposure that students experience in the lead up to their exams. To account for possible spurious correlation between regional warming trends and secular changes in educational outcomes, we included continent-specific trends in all regressions. We also include country-specific trends in Supplementary Table 9. Our approach is analogous to the now widely used binning of annual temperatures first described in ref. [48].

School versus non-school days estimation. In our primary specifications, we binned all days in a year together. We also separately reported the results of the effect of school days above 80 °F and non-school days above 26.7 °C (80 °F). There, we estimated the following model:

$$\bar{Z_{it}} = \sum_{k=1}^{9} \beta_k \mathbf{TMAX}_{ikg}^{\mathrm{school}} + \sum_{k=1}^{9} \psi_k \mathbf{TMAX}_{ikg}^{\mathrm{non-school}} + \sigma \mathbf{X}_{it} + \gamma_t + \delta_i + \omega_{ct} + \epsilon_{it} \quad (2)$$

where the variables are as before but β reports the estimates of the impact of days while school is in session, while ψ reports the effects of non-school days.

Summer versus non-summer estimation. We distinguished school year days further by separating school year weekend days from school year non-weekend days. We estimated:

$$\bar{Z_{it}} = \sum_{k=1}^{9} \beta_k \mathbf{TMAX}_{ikg}^{\mathrm{school}} + \sum_{k=1}^{9} \psi_k \mathbf{TMAX}_{ikg}^{\mathrm{summer}} + \sum_{k=1}^{9} \phi_k \mathbf{TMAX}_{ikg}^{\mathrm{school\ weekend}} + \sigma \mathbf{X}_{it} + \gamma_t + \delta_i + \omega_{kt} + \epsilon_{ijt} \quad (3)$$

where the variables are as before but β reports the estimates of the impact of days while school is in session, while ψ reports the effects of non-school days.

Subject-specific estimation. Finally, we estimated subject-specific effects. There, we returned to the original estimating equation:

$$\bar{Z_{its}} = \sum_{k=1}^{9} \beta_k \mathbf{TMAX}_{ikg} + \sigma \mathbf{X}_{it} + \gamma_t + \delta_i + \omega_{ct} + \epsilon_{it} \quad (4)$$

However, we replaced $\bar{Z_{it}}$ with the subject-specific normalized score, $\bar{Z_{its}}$ for each of reading, science and mathematics in PISA and ELA and mathematics in SEDA. In each case, we calculated the normalized score in the way described above. In the PISA data, for both mathematics and science, we used the shorter panel to avoid data comparability issues due to changes in the PISA methodology in those subjects. To estimate the subject-specific effects of school and non-school days, we substituted $\bar{Z_{its}}$ into equation (2). For additional details on the empirical strategy, including robustness checks and regression tables, see the Supplementary Information.

Reporting Summary. Further information on research design is available in the Nature Research Reporting Summary linked to this article.

Data availability
The weather data that support the findings of this study are available from the National Oceanic and Atmospheric Administration (https://www.ncdc.noaa.gov/). The international assessment data are available from PISA (https://www.oecd.org/pisa/data/). The US assessment data are available through NCES, compiled by district, grade, subject and year at SEDA (https://exhibits.stanford.edu/data/catalog/db586ns4974). Additional data at the country level, including employment shares and per-capita income, are available at the World Bank's World Development Indicators archives (https://datatopics.worldbank.org/world-development-indicators/wdi-archives.html).

Code availability
Custom code that supports the findings of this study is available from the corresponding author upon request.

Received: 11 October 2019; Accepted: 27 August 2020;
Published online: 05 October 2020

References
1. de Secondat Montesquieu, C. *The Spirit of Laws*. vol. 2. (J. Nourse and P. Vaillant, 1773).
2. Dell, M., Jones, B. F. & Olken, B. A. Temperature shocks and economic growth: evidence from the last half century. *Am. Econ. J. Macroecon.* **4**, 66–95 (2012).
3. Park, R. J., Goodman, J. & Goodman, M. & Smith, J.Heat and learning. *Am. Econ. J. Econ. Policy* **12**, 306–339 (2020).
4. Goldin, C. D. & Katz, L. F. *The Race Between Education and Technology* (Harvard Univ. Press, 2009).
5. Chetty, R., Hendren, N., Kline, P. & Saez, E. Where is the land of opportunity? The geography of intergenerational mobility in the United States. *Q. J. Econ.* **129**, 1553–1623 (2014).
6. David, H. Skills, education, and the rise of earnings inequality among the 'other 99 percent'. *Science* **344**, 843–851 (2014).
7. Hanushek, E. A. & Woessmann, L. Knowledge capital, growth, and the East Asian miracle. *Science* **351**, 344–345 (2016).
8. Marshall, B. & Emerick, K. Adaptation to climate change: evidence from US agriculture. *Am. Econ. J. Econ. Policy* **8**, 106–140 (2016).
9. Acemoglu, D., Johnson, S. & Robinson, J. A. The colonial origins of comparative development: an empirical investigation. *Am. Econ. Rev.* **91**, 1369–1401 (2001).
10. Nathan, N. & Wantchekon, L. The slave trade and the origins of mistrust in Africa. *Am. Econ. Rev.* **101**, 3221–3252 (2011).
11. Nathan, N. The historical roots of economic development. *Science* **367**, eaaz9986 (2020).
12. Gallup, J. L., Sachs, J. D. & Mellinger, A. D. Geography and economic development. *Int. Reg. Sci. Rev.* **22**, 179–232 (1999).
13. Schlenker, W., Hanemann, W. M. & Fisher, A. C. The impact of global warming on US agriculture: an econometric analysis of optimal growing conditions. *Rev. Econ. Stat.* **88**, 113–125 (2006).
14. Gallup, J. L. & Sachs, J. D. The economic burden of malaria. *Am. J. Trop. Med. Hyg.* **64**, 85–96 (2001).
15. Currie, J. Healthy, wealthy, and wise: socioeconomic status, poor health in childhood, and human capital development. *J. Econ. Lit.* **47**, 87–122 (2009).
16. Maccini, S. & Yang, D. Under the weather: health, schooling, and economic consequences of early-life rainfall. *Am. Econ. Rev.* **99**, 1006–1026 (2009).
17. Shah, M. & Steinberg, B. M. Drought of opportunities: contemporaneous and long-term impacts of rainfall shocks on human capital. *J. Political Econ.* **125**, 527–561 (2017).
18. Ramsey, J. D. Task performance in heat: a review. *Ergonomics* **38**, 154–165 (1995).
19. Hocking, C., Silberstein, R. B., Lau, W. M., Stough, C. & Roberts, W. Evaluation of cognitive performance in the heat by functional brain imaging and psychometric testing. *Comp. Biochem. Physiol. A Mol. Integr. Physiol.* **128**, 719–734 (2001).
20. Park, R. J. Hot temperature and high stakes performance. *J. Hum. Res.* https://doi.org/10.3368/jhr.57.2.0618-9535R3 (2020).
21. Graff Zivin, J. & Neidell, M. Temperature and the allocation of time: implications for climate change. *J. Labor Econ.* **32**, 1–26 (2014).
22. Acemoglu, D. & Autor, D. in *Handbook of Labor Economics* Vol. 4, 1043–1171 (Elsevier, 2011).
23. Graff Zivin, J. S., Hsiang, S. M. & Neidell, M. J.Temperature and human capital in the short- and long-run.*J. Assoc. Environ. Res. Econ.* **5**, 77–105 (2017).
24. Carleton, T. A. et al. *Valuing the Global Mortality Consequences of Climate Change Accounting for Adaptation Costs and Benefits* Working paper (NBER, 2018).
25. Hsiang, S., Oliva, P. & Oliva, R.The distribution of environmental damages.*Rev. Environ. Econ. Pol.* **13**, 83–103 (2018).
26. Rowland, T. Thermoregulation during exercise in the heat in children: old concepts revisited. *J. Appl. Physiol.* **105**, 718–724 (2008).
27. Ebenstein, A., Lavy, V. & Roth, S. The long-run economic consequences of high-stakes examinations: evidence from transitory variation in pollution. *Am. Econ. J. Appl. Econ.* **8**, 36–65 (2016).
28. Paula, A. & Päivi, P.-K. Sleep deprivation: impact on cognitive performance. *Neuropsychiatr. Dis. Treat.* **3**, 553–567 (2007).
29. Mani, A., Mullainathan, S., Shafir, E. & Zhao, J. Poverty impedes cognitive function. *Science* **341**, 976–980 (2013).
30. Burke, M., Hsiang, S. M. & Miguel, E. Global non-linear effect of temperature on economic production. *Nature* **527**, 235–239 (2015).
31. Barreca, A., Clay, K., Deschenes, O., Greenstone, M. & Joseph, J. S. Adapting to climate change: the remarkable decline in the US temperature–mortality relationship over the twentieth century. *J. Political Econ.* **124**, 105–159 (2016).
32. Davis, L. W. & Gertler, P. J. Contribution of air conditioning adoption to future energy use under global warming. *Proc. Natl Acad. Sci. USA* **112**, 5962–5967 (2015).
33. Randazzo, T., De Cian, E. & Mistry, M. Air conditioning and electricity expenditure: the role of climate in temperate countries. *Econ. Model.* **90**, 273–287 (2020).
34. Gilfillan, S. C. The coldward course of progress. *Polit. Sci. Q.* **35**, 393–410 (1920).
35. Huntington, E. *Civilization and Climate* (Yale Univ. Press, 1922).
36. Hsiang, S. M., Burke, M. & Miguel, E. Quantifying the influence of climate on human conflict. *Science* **341**, 1235367 (2013).
37. Reardon, S. F., Kalogrides, D. & Shores, K. *The Geography of Racial/Ethnic Test Score Gaps*. CEPA working paper no. 16-10 (Stanford Center for Education Policy Analysis, 2017).
38. Chetty, R. et al. How does your kindergarten classroom affect your earnings? Evidence from project star. *Q. J. Econ.* **126**, 1593–1660 (2011).
39. Woessmann, L. The importance of school systems: evidence from international differences in student achievement. *J. Econ. Perspect.* **30**, 3–31 (2016).
40. Kjellstrom, T. & Crowe, J. Climate change, workplace heat exposure, and occupational health and productivity in Central America. *Int. J. Occup. Environ. Health* **17**, 270–281 (2011).
41. Graff-Zivin, J. & Neidell, M. The impact of pollution on worker productivity. *Am. Econ. Rev.* **102**, 3652–3673 (2012).

42. Deschênes, O. & Greenstone, M. Climate change, mortality, and adaptation: evidence from annual fluctuations in weather in the US. *Am. Econ. J. Appl. Econ.* **3**, 152–185 (2011).
43. Anderson, G. B. et al. Heat-related emergency hospitalizations for respiratory diseases in the Medicare population. *Am. J. Resp. Crit. Care Med.* **187**, 1098–1103 (2013).
44. Lutz, W., Muttarak, R. & Striessnig, E. Universal education is key to enhanced climate adaptation. *Science* **346**, 1061–1062 (2014).
45. Greenstone, M., Kopits, E. & Wolverton, A. Developing a social cost of carbon for US regulatory analysis: a methodology and interpretation. *Rev. Environ. Econ. Policy* **7**, 23–46 (2013).
46. Anthoff, D. & Emmerling, J. Inequality and the social cost of carbon. *J. Assoc. Environ. Resour. Econ.* **6**, 29–59 (2019).
47. Auffhammer, M. & Mansur, E. T. Measuring climatic impacts on energy consumption: a review of the empirical literature. *Energy Econ.* **46**, 522–530 (2014).
48. Deschênes, O. & Greenstone, M. The economic impacts of climate change: evidence from agricultural output and random fluctuations in weather. *Am. Econ. Rev.* **97**, 354–385 (2007).
49. Barreca, A., Clay, K., Deschênes, O., Greenstone, M. & Shapiro, J. S. Convergence in adaptation to climate change: evidence from high temperatures and mortality, 1900–2004. *Am. Econ. Rev.* **105**, 247–251 (2015).

Acknowledgements
We acknowledge P. Stainier for excellent research assistance. We received no specific funding for this work.

Author contributions
R.J.P., J.G. and A.P.B. designed the research, performed the research, analysed the data and wrote the paper.

Competing interests
The authors declare no competing interests.

Additional information
Supplementary information is available for this paper at https://doi.org/10.1038/s41562-020-00959-9.

Correspondence and requests for materials should be addressed to R.J.P., J.G. or A.P.B.

Peer review information Primary Handling Editor: Aisha Bradshaw.

Reprints and permissions information is available at www.nature.com/reprints.

nature research

Corresponding author(s): R. Jisung Park

Last updated by author(s): Apr 21, 2020

Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Research policies, see our Editorial Policies and the Editorial Policy Checklist.

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a	Confirmed	
☐	☒	The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement
☐	☒	A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☐	☒	The statistical test(s) used AND whether they are one- or two-sided *Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
☐	☒	A description of all covariates tested
☐	☒	A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
☐	☒	A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
☐	☒	For null hypothesis testing, the test statistic (e.g. *F*, *t*, *r*) with confidence intervals, effect sizes, degrees of freedom and *P* value noted *Give P values as exact values whenever suitable.*
☒	☐	For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
☒	☐	For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
☒	☐	Estimates of effect sizes (e.g. Cohen's *d*, Pearson's *r*), indicating how they were calculated

Our web collection on statistics for biologists contains articles on many of the points above.

Software and code

Policy information about availability of computer code

Data collection	No software was used.
Data analysis	Analysis was conducted in Stata15 and R.

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors and reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Research guidelines for submitting code & software for further information.

Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A list of figures that have associated raw data
- A description of any restrictions on data availability

The weather data that support the findings of this study are available from NOAA.
The international assessment data are available from PISA.
The US assessment data are available through the National Center for Educational Statistics, compiled by district-grade-subject-year at the Stanford Educational Data Archives. Additional data at the country level including employment shares and per capita income are available at the World Bank WDI archives.

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☐ Life sciences ☒ Behavioural & social sciences ☐ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/documents/nr-reporting-summary-flat.pdf

Behavioural & social sciences study design

All studies must disclose on these points even when the disclosure is negative.

Study description	Quantitative using quasi-experimental observational methods.
Research sample	There are two samples in the work. The universe of countries that take the PISA exams from 2000 to 2015 and the students in the Stanford Educational Database (SEDA) from 2009-2015. The PISA scores are the full sample of countries that take that exam. The SEDA data is a state-representative sample of students from 3rd to 8th grade.
Sampling strategy	N/A
Data collection	N/A
Timing	N/A
Data exclusions	N/A
Non-participation	N/A
Randomization	No randomization by authors.

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

n/a	Involved in the study	
☒	☐	Antibodies
☒	☐	Eukaryotic cell lines
☒	☐	Palaeontology and archaeology
☒	☐	Animals and other organisms
☒	☐	Human research participants
☒	☐	Clinical data
☒	☐	Dual use research of concern

Methods

n/a	Involved in the study	
☒	☐	ChIP-seq
☒	☐	Flow cytometry
☒	☐	MRI-based neuroimaging