I Z A Institute of Labor Economics

Initiated by Deutsche Post Foundation

DISCUSSION PAPER SERIES

IZA DP No. 15972

Temperature and Low-Stakes Cognitive Performance

Xin Zhang
Xi Chen
Xiaobo Zhang

FEBRUARY 2023

IZA Institute of Labor Economics

Initiated by Deutsche Post Foundation

DISCUSSION PAPER SERIES

IZA DP No. 15972

Temperature and Low-Stakes Cognitive Performance

Xin Zhang
Beijing Normal University

Xi Chen
Yale University and IZA

Xiaobo Zhang
Peking University and IFPRI

FEBRUARY 2023

ISSN: 2365-9793

IZA – Institute of Labor Economics

Schaumburg-Lippe-Straße 5–9
53113 Bonn, Germany

Phone: +49-228-3894-0
Email: publications@iza.org

www.iza.org

ABSTRACT

Temperature and Low-Stakes Cognitive Performance

This paper offers one of the first evidence in a developing country context that transitory exposure to high temperatures may disrupt low-stakes cognitive activities across a range of age cohorts. By matching eight years of repeated cognitive tests among all the participants in a nationally representative longitudinal survey in China with weather data according to the exact time and geographic location of their assessment, we show that exposure to a temperature above 32 °C on the test date, relative to a moderate day within 22–24 °C, leads to a sizable decline in their math scores by 0.066 standard deviations (equivalent to 0.23 years of education). Also, the effect on the math test scores becomes more pronounced as people age, especially for males and the less educated. However, the test takers living in hotter regions or those with air conditioning installed on site are less vulnerable to extreme high temperatures, indicating the role of adaptation.

JEL Classification: I24, Q54, Q51, D91, J14, J16

Keywords: cognition, high temperatures, climate change, adaptation, age gradients

Corresponding author:
Xi Chen
Department of Health Policy and Management
Department of Economics
Yale University
60 College St
New Haven, CT 06520
USA
E-mail: xi.chen@yale.edu

1. Introduction

Climate change has brought about more frequent extreme temperatures, such as heat waves and cold spells. The world's average temperature has increased 0.6 °C in the past three decades and 0.8 °C in the past century, and the trend is projected to continue (Hansen et al. 2006). The Intergovernmental Panel on Climate Change (IPCC) warns that, if greenhouse gas emissions continue at the current rate, by 2034 the atmosphere may warm up by as much as 1.5 °C (2.7 °F) above preindustrial levels (IPCC 2021). Along with the rising temperatures, heat waves are expected to occur more often.

There are several channels through which extreme temperatures may impede the cognitive performance of humans. First, cognitive activities often rely on regions in the brain sensitive to heat or cold weather, potentially causing impaired brain functioning (Hocking et al. 2001; Kiyatkin 2007). Second, exposure to heat waves may reduce the flow of blood to the brain (Kiyatkin 2007; Raichle and Mintun 2006) and therefore increase heat-related fatigue (McMorris et al. 2006; Nybo et al. 2014). Third, thermal stress may diminish a person's attention, working memory, information retention, and processing (Hocking et al. 2001; Vasmatzidis et al. 2002).

There is a growing body of literature assessing the impact of extreme temperatures, particularly heat waves, on cognitive performance. Some studies have examined the effect of exposure to heat waves on students' high-stakes exams (Park et al. 2020; Graff Zivin et al. 2020; Cho 2017; Park 2022; Park et al. 2021), while others have studied the impact of extreme temperatures on less challenging cognitive activities, often with a focus on children and young adults (Graff Zivin et al. 2018; Garg et al. 2020). It remains unclear, however, whether these findings hold true for the general population in low-stakes cognitive activities.

Our paper is among the first to provide evidence of the link between the transitory exposure to high temperatures and the performance of low-stakes cognitive activities, by leveraging a nationally representative longitudinal household survey in China that includes almost all age cohorts, and matching these with weather data according to the exact time and geographic location of the cognitive tests they undertook. By exploiting the variations in exposure to extreme temperatures for the same individuals over eight years (2010–2018), we show that exposure to heat waves impedes performance in math tests. Specifically, exposure to a mean temperature above 32 °C on the test date, relative to a day in the 22–24 °C range, leads to math test scores to decline by 0.066 standard deviations (SD), equivalent to a loss of 0.23 years of education. The effect on math test scores becomes more pronounced as people age, especially for males and the less educated. Test takers living in hotter regions or with air conditioning (AC) installed on site are less vulnerable to extreme high temperatures, indicating the role of adaptation. Some preliminary evidence suggest that our findings may not be driven by behavioral channels, such as the respondents being less cooperative, more impatient, or hastier in doing low-stakes cognitive assessments, but more plausibly by an impairment in the respondents' cognitive ability.

We contribute to the literature on several fronts. First, by including all groups above age 10 in low-stakes cognitive tests, we make the first attempt to identify heterogenous sensitivity to high temperatures by age; whereas the existing studies mainly focus on young children (Garg et al. 2020; Graff Zivin et al. 2018). Park et al. (2021) examine the age gradient of exposure to hot school days between students in elementary schools and those in middle schools. Our findings indicate that, while taking math tests, the impact of a day with a mean temperature above 32 °C, relative to a day in the 22–24 °C range, is on average 1.5 times as large on the elderly as that on middle-aged people. As our mechanism tests suggest, the pronounced impact seems not to occur through behavioral channels, such as by being less cooperative, more impatient, or hastier to finish low-stakes cognitive assessments, but more plausibly through a disruption in their cognitive ability.

Second, with detailed information on individual-level residential AC status, we accurately assess the role of residential AC in the linkage between extreme temperatures and cognitive performance. Previous studies have either relied on aggregated residential AC penetration data (e.g., Park et al. 2020) or imputed the probability of AC ownership based on social survey data (e.g., Graff Zivin et al. 2018). Our estimates indicate that the adoption of AC at home offsets the negative effects of hot days (>32 °C) on cognition by 36.6%. These findings accounting for avoidance behaviors and defensive investments suggest there is substantial scope for adaptation to higher temperatures. In line with recent work exploring the extent to which AC can mitigate the harmful effects of heat waves on mortality and labor productivity (Barreca et al. 2016; Behrer and Park 2017; Deschênes and Greenstone 2011; Heutel et al. 2021), our results provide insights into the potential offsetting effects of adaptive behaviors, which are expected could play a critical role in determining the ultimate impacts of higher temperatures in a changing climate.

Third, we are among the first to estimate the transitory impact of exposure to high temperatures on low-stakes cognitive performance in a developing country setting and the benefits of residential AC. Garg et al. (2020) offer another evaluation of the transitory effect in an agrarian context in a single state of India, though they did not assess the role of residential AC. The penetration rates of residential AC differ vastly between developed countries and developing countries. For example, survey evidence suggests that while 90% of US households have some form of AC, only 34% and 13% of households in China and Mexico, respectively, have AC (Park et al. 2021). Given that the effects of climatic shocks on health-related outcomes vary substantially by socioeconomic status (Park et al. 2021; Isen et al. 2017), and that defensive investments such as AC can be effective in attenuating the impacts (Barreca et al. 2016; Behrer and Park 2017; Park et al. 2020), it is important to verify the external validity of the evidence from high-income countries.

Fourth, building upon the three contributions summarized above, our improved understanding of heterogeneities in the temperature–cognition relationship by geographic region, level of adaptation, age cohort, and other key demographic factors may inform more accurate climate damage assessments in the long run. The existing assessments, however, have generally assumed a uniform relationship (Deschênes and Greenstone 2011; Hsiang et al. 2017) with few exceptions (e.g., Heutel et al. 2021). Overall, some previous climate damage assessments could deviate substantially from reality if the cognition effects of extreme temperatures vary geographically, change with the population aging, or if people adapt to their future climate.

Finally, our findings also shed light on the various consequences of extreme temperatures. Besides raising the mortality rate and disease burden (Deschênes and Greenstone 2011; Huang et al. 2012; Karlsson and Ziebarth 2018; Banerjee and Maharaj 2020; Lee and Li 2021), increasing the risk of mental illness (Obradovich et al. 2018; Mullins and White 2019) and suicide rates (Burke et al. 2018), and reducing labor supply (Deschenes 2014; Graff Zivin and Neidell 2014), as well as agricultural income and nutrition (Deschênes and Greenstone 2007; Shah and Steinberg 2017), we show that high temperatures may impair cognitive ability, which would deplete human capital and labor productivity, an important engine of economic growth. The impact is particularly acute for the elderly population. Older adults are often faced with high-stake decisions, such as on retirement timing, pension premiums, medical treatment, and choice of health insurance purchasing, which rely heavily on math skills. As global warming is projected to accelerate with more frequent high temperatures and with population aging in the coming decades, our findings suggest that a greater number of the elderly's high-stake decisions may be more frequently impaired by hot weather in the future. The total economic and social costs of heat waves would be larger than previously thought, if we take this toll into account.

The rest of the paper is organized as follows. Section 2 describes our data sources. Section 3 discusses our empirical strategy. Section 4 reports our findings, including baseline results, adaptation and stratification, as well as robustness checks. Finally, section 5 concludes the paper.

2. Data

2.1. Cognition data

Data on cognitive tests were obtained from the China Family Panel Studies (CFPS), a nationally representative biennial longitudinal household survey of Chinese families and individuals. CFPS is funded by Peking University and carried out by the university's Institute of Social Science Survey.[1] CFPS includes questions on a wide range of topics for families and individuals, including family dynamics and relationships, economic activities, health status, subjective well-being, and cognitive abilities.

The waves 2010, 2014, and 2018 of CFPS contain the same cognitive ability module, i.e., comprising 24 standardized mathematics questions and 34 word-recognition questions. The tests were conducted at respondents' homes. All the questions were obtained from standard textbooks and were sorted in ascending order of difficulty. The starting question depends on the respondent's education level.[2] The test ends when the individual

[1] The survey uses multistage probability proportional to size sampling with implicit stratification to better represent Chinese society. The 2010 CFPS baseline sample was drawn through three stages (i.e., county, village, and household) from 25 provinces. The 162 randomly chosen counties largely represent Chinese society (Xie and Hu 2014).

[2] Specifically, those whose education level is primary school or below start with the 1st question; those who

incorrectly answers three questions in succession. The final test score is defined as the rank of the hardest question a respondent can answer correctly. If the respondent fails to answer any questions, the score is assigned as the rank of the starting question minus one. For example, a respondent with middle school education would begin with the 9th question in the verbal test. If the hardest question the respondent can correctly answer is the 14th question, then the verbal test score would be 14. However, if the respondent fails to answer the 9th, 10th, and 11th questions consecutively, the verbal test score would be 8. Since the respondents do not know the testing rules prior to the interviews, there should be no incentive to manipulate their test performance on purpose.

CFPS is suitable for our study for several reasons. First, the survey includes several standardized cognitive tests. Second, the survey embodies information on residential AC ownership, allowing us to study adaptation behaviors. Third, exact information about the geographic locations and test dates for all the respondents is available, enabling precisely matching the individual test scores in the survey with local weather data. Further, the longitudinal data allow us to remove unobserved individual factors that may bias the estimates. Finally, because the cognitive tests are administered to all age cohorts older than 10 years old, we can study the effects of high temperatures across age groups.

2.2. Weather data

The weather data were provided by the China National Meteorological Data Service Center (CMDC) under the National Meteorological Information Center of China. The dataset contains daily weather records of 824 monitoring stations along with their longitudes and latitudes in China.[3] The key variable for our analysis was the daily mean temperature. Other weather controls include precipitation, wind speed, sunshine duration, and relative humidity. We interpolate the weather data from the stations into a $0.1° \times 0.1°$ grid level based on the inverse-distance weighting (IDW) method and extract the value of

---

attended middle school begin with the 9th question in the verbal test and the 13th question in the math test; and those who finished high school or above start with the 21st question in the verbal test and the 19th question in the math test.

[3] The spatial distribution of weather stations is displayed in Figure A1. Note, we are not allowed to mark the exact locations of the sampled counties in CFPS under Chinese privacy law.

the weather measures based on the boundaries of each county from the gridded data.[4]

As some previous studies have shown that air pollution is associated with bad performance in cognitive tests (Ebenstein et al. 2016; Zhang et al. 2018), we also control for air quality, which was collected from the air quality report published by the Chinese Ministry of Ecology and Environment (MEE).[5] Air quality is measured using the air pollution index (API), which ranges from 0 to 500, with larger values indicating worse air quality.[6] We match each CFPS county to the nearest API reporting city within 100 km according to the distance between the county centroid and the city boundaries.

Figure A2 depicts a histogram of the mean temperatures on the test dates in our sample. As most of the interviews were conducted in July and August when college students were employed as numerators (Figure A3), the distribution is skewed toward higher temperatures, with the mean being 24.40 °C. Following the general practice in the latest literature (Graff Zivin et al. 2018, 2020; Cho 2017), we use a state-of-the-art arrangement of 2 °C per bin indicators to allow for substantial flexibility and the nonlinear relationships between the cognitive performance and temperature exposure. Specifically, we divide the spectrum of temperatures into 12 bins, with the lowest bin including all temperatures below 12 °C and the highest bin including all temperatures above 32 °C, due to data sparseness at the extremities of the distribution. Figure 1 shows a plot of the percentage of days that fall into each bin, with 11.89% falling in the 22–24 °C range, 18.67% in the 28–30 °C bin, and 1.58% in the greater than 32 °C bin.

CFPS surveyed a panel of 49,652 individual respondents over 10 years old in 2010, 2014, and 2018, for a total of 96,990 observations with cognitive test scores. Of these observations, 1,728 are missing values for test dates or locations. Among the remaining

[4] The weather dataset provided by the CMDC has been widely used in the recent literature when studying weather/climate change in China (for example, Agarwal et al. 2021; Graff Zivin et al. 2020). As far as we know, the distribution of the weather stations in this dataset is finer than that of the gridded temperature products, which are typically at the 0.5°× 0.625° grid level. Interpolating the weather data from stations into the 0.1°× 0.1° grid level enables us to match the weather data following the exact boundaries for each county, which can help ameliorate the concerns about potential measurement errors caused by the relatively small matching radius in some geographically large counties in western China.

[5] The report includes 86 major cities in 2000 and covers most of the cities in China since 2014.

[6] Carbon monoxide (CO), ozone (O_3), and particulate matter with a diameter smaller than 2.5 μm (PM2.5) were not added to the basket of the index until 2013. Because all the cognitive tests were administered between 2010 and 2018, we transform the air quality index (AQI) to the API in 2014 and 2018, and use the API based on sulfur dioxide (SO_2), nitrogen dioxide (NO_2), and particulate matter with a diameter smaller than 10 μm (PM10) in our paper.

95,262 observations, 70,771 observations could be matched to weather and API data. Due to there being some missing values for other control variables, the final dataset used in this study included 70,736 observations.[7]

3. Empirical strategy

Our baseline econometric specification is as follows:

$$Score_{ijt} = \sum_{k=1}^{12} \alpha_k TEMP_{jtk} + X'_{ijt}\beta + W'_{jt}\phi + \lambda_i + \delta_j + \eta_t + \varepsilon_{ijt} \quad (1)$$

The dependent variable $Score_{ijt}$ is the cognition test scores of respondent i in county j at date t. The two cognitive test scores we test in this paper are verbal test scores and math test scores. The key variables of interest $TEMP_{jtk}$ are a series of indicators for whether the mean temperature falls into temperature bin k (from 1 to 12) on the test date t in county j. We deploy 12 bins, i.e., lower than 12 °C bin, higher than 32 °C bin, and ten 2 °C-wide bins in between. We set the 22–24 °C temperature bin as the reference group as it is associated with the highest cognitive test scores. The vector X_{ijt} represents demographic correlates, including gender, age with its square term and education level. We also control for a vector of contemporaneous air quality and weather conditions W_{jt}, involving API, precipitation, wind speed, sunshine duration, and relative humidity in quadratic forms. λ_i denotes individual fixed effects. δ_j represents county fixed effects.[8] η_t indicates interview year, month, day-of-week, and hour-of-day fixed effects. ε_{ijt} is the error term. Standard errors are clustered at the county level. All the key variables and their summary statistics are described in Table 1.

By conditioning on the individual fixed effects and other sets of fixed effects listed above, the key parameters α_k are identified by making use of variations in exposure to temperatures for the same respondent in the three waves after controlling for seasonality and annual shocks. Due to the unpredictability of test dates and thus the random of

[7] Our further investigations confirm that the missing temperature data are not systematically correlated with household/county characteristics; the missing survey data are unrelated to temperature bins; and there are little systematic differences between the fixed effect sample and the full sample in terms of basic household/individual characteristics. The results are available upon request.

[8] The county fixed effects cannot be wiped out by individual fixed effects since some respondents do not live in the same counties across the three waves.

temperature fluctuations, it is reasonable to assume that this variation is orthogonal to the unobserved determinants of cognitive test scores.

4. Results

4.1. Baseline results and interpretations

Table A1 displays various specifications to our baseline results. Panel A corresponds to the verbal tests, and Panel B the math tests. The first column in each panel includes temperature exposure, demographic controls (i.e., gender, age with its square term, and education level) and environmental conditions (i.e., API, precipitation, wind speed, sunshine duration, and relative humidity in quadratic forms), and county fixed effects, as well as interview year, month, day-of-week, and hour-of-day fixed effects. We find a strong negative effect of exposure to high temperatures (>32 °C) on the math test scores. Also, the pattern continues to hold when individual fixed effects are further added in the second column of each panel.

Figure 2 plots the estimated results from the preferred specification in Columns (2) and (4) of Table A1. Figure 2A corresponds to the verbal test scores, while Figure 2B refers to the math test scores. Each figure reveals the estimated coefficients for 12 temperature bins (<12 °C, 12–14 °C, 14–16 °C, 16–18 °C, 18–20 °C, 20–22 °C, 22–24 °C, 24–26 °C, 26–28 °C, 28–30 °C, 30–32 °C, and >32 °C) in equation (1), together with their 90% and 95% confidence intervals (CIs). The temperature bin left out is 22–24 °C. Therefore, the coefficients for each bin measure the changes in test scores when the temperature falls into that bin relative to the reference bin.

As revealed in Figure 2A, there is no obvious association between temperature exposure and cognitive performance in the verbal tests. All the coefficients on the temperature bins are insignificant. Therefore, we can conclude that high temperatures seem to have little effect on the verbal test scores. Figure 2B further presents the estimated effect on the math test scores. We find a nonlinear relationship between the temperature and cognitive performance in the math tests, where high temperatures are associated with a decline in math test scores. Further, when exposed to temperatures higher than 32 °C, the negative effect is significant at the 5% level. Specifically, a test day with a mean temperature above 32 °C, relative to a day in the 22–24 °C range, leads to a reduction in math test scores by 0.420. To put this into context, note that the SD of the math test scores is 6.351. Therefore, the respondents' math test scores on a day with average temperatures above 32 °C are on average 0.066 SD lower than their scores on a day in the reference temperature bin (22–24 °C).

We compare the magnitude of our estimates with other similar studies that also use a series of 2 °C temperature bins, though their distributions of temperatures, highest temperature bins, and reference bins may vary due to the differences in countries and seasons of the surveys. The relevant studies are summarized in Table 2. To facilitate comparisons across studies, we report the effect sizes in SD change of the test scores per 1 °C higher temperature. As for our estimates, since the difference between the bins above 32 °C and 22–24 °C is approximately 9 °C, each 1 °C higher temperature decreases the math test scores by 0.0073 (0.066/9) SD.[9] Meanwhile, an increase in temperature by 1 °C decreases the test scores by 0.0120 (0.12/10) SD in Graff Zivin et al. (2018), 0.0320 (0.48/15) SD in Graff Zivin et al. (2020), and 0.0008 (0.0042/5) SD in Cho (2017), respectively. By comparison, our effect size is about two-thirds of that in Graff Zivin et al. (2018), who also use low-stakes test scores, and around one-third of that in Graff Zivin et al. (2020) for high-stakes Chinese college entrance exams. Notably, our effect size is much higher than that in Cho (2017) leveraging Korean college entrance exams.

As reviewed in Table 2, the effect size varies greatly across studies between low-stakes and high-stakes settings, and even between high-stake settings. On the one hand, if test takers exert more effort in high-stake settings than low-stake settings to offset the negative effect of high-temperature exposure, it is possible to observe a larger effect in low-stake settings. On the other hand, there are opposing factors at play that should be considered. First, the literature has mostly focused on school settings, in which high temperatures may impose both a transitory effect during the tests and longer-term effect on learning loss. Most of the respondents in our study, however, have already completed their education. In the absence of a learning loss channel for the adult population, our identified effect is therefore mainly transitory. Second, the penetration rate of AC in our sample is higher than

[9] Although our measure of the marginal effect (for each 1 °C increase in temperature) follows the existing studies (e.g., Graff Zivin et al. 2020), our linear approximation may understate the marginal effect of exposure to temperature extremes. Future work with larger sample and more statistical power may refine our bin classification so as to offer more accurate effect estimates over a wider range of temperature exposure.

in some high-stakes settings in public high schools (Park 2022; Graff Zivin et al. 2020), which partly mitigates the negative effect of high temperatures. Third, test takers often feel more pressure during high-stakes tests than low-stakes ones, which may affect their performance (Cai et al. 2019). High temperatures may worsen stress, which in turn may impede cognitive performance. Fourth, as illustrated earlier, our linear approximation may understate the real marginal effect of exposure to temperature extremes, if the true effect is nonlinear.

Our results by test subject are consistent with the literature in which the transitory effect of exposure to high temperatures is more often observed in math tests than in other subjects, like word recognition and reading comprehension (Graff Zivin et al. 2018; Park 2022; Garg et al. 2020). One potential explanation is that different regions of the brain perform distinct cognitive functions, and the regions responsible for solving math problems may be more sensitive to extreme temperatures than the regions in charge of reading functions (Hocking et al. 2001). These differential effects in the short run across cognitive tasks also provide strong evidence for the presence of a physiological channel connecting temperature exposure to cognitive performance.

To make the results more intuitive, we further interpret our findings by calculating the years of education lost based on the estimates, as cognition and educational attainment are highly correlated and intrinsically linked. Figure A4 plots the average years of education versus cognitive test scores for respondents, as well as their correlation coefficients. A one-point increase in verbal test scores corresponds to 0.321 years of education, while a one-point increase in math test scores is equivalent to 0.545 years of education. As calculated from Table A1, exposure to a mean temperature above 32 °C on the test date, relative to a day in the 22–24 °C range, leads to a sizable decrease in math test scores by 0.23 years of education.

4.2. Adaptation and stratification

We study adaptation to heat waves in two dimensions. First, we split the sample according to residential AC ownership.[10] As shown in Figure 3A for the verbal test scores

[10] The AC ownership information is only available in CFPS wave 2014. Therefore, we employ the same specification in Columns (1) and (3) of Table A1 by including demographic controls (i.e., gender, age with

and in Figure 3B for the math test scores, the negative effect of exposure to high temperatures is significant only for individuals taking math tests without AC. As reported in Columns (4)–(5) of Table A2, the adoption of AC offsets some of the negative effects of hot days (>32 °C) on cognition. The effect size of high temperatures on math test scores with AC is 36.6% (= (0.650-0.412)/0.650) smaller than that without AC. However, as revealed in Column (6) of Table A2, there is no statistically significant difference in the impact of high temperatures between these two subsamples.

Second, we repeat the exercises for cooler and hotter regions of China, classified according to the median of each county's average temperature in summer days (from July to August) during 2010–2018. The results are plotted in Figure 4, and the corresponding numerical results are displayed in Table A3. As revealed in Figure 4, high temperatures (>32 °C) are more harmful to the math test performance of respondents living in cooler regions than those in hotter regions. Nevertheless, the cross-equation test in Column (6) of Table A3 indicates that the difference between cooler and hotter regions is statistically insignificant. In the cooler regions, participants' math test scores, on a day with an average temperature above 32 °C, are on average 0.701 (0.110 SD) lower than their scores on a day in the reference temperature bin (22–24 °C). By contrast, people living in hotter regions are more sensitive to cold spells than those living in cooler regions, and the difference is significant at the 5% level.[11] Our findings are consistent with the literature on adaptation behaviors. For example, Cho (2017) shows that students in cities with relatively cool summers are affected more than students in cities with relatively hot summers. Behrer and Park (2017) find that hotter places in the US seem to better adapt to heat stress than cooler areas.

We also conduct stratification analyses by gender, age, education level, and workplace.

---

its square term and education level) and environmental conditions (i.e., API, precipitation, wind speed, sunshine duration, and relative humidity in quadratic forms), county fixed effects, as well as interview year, month, day-of-week, and hour fixed effects. We test the cross-sectional determinants of AC adoption in the sample. We find that higher income individuals, educated people and urban residents are more likely to have an AC at home, which corroborates the findings in the existing studies (Biddle 2008; Davis and Gertler 2015; Park et al. 2020).

[11] As the cooler regions in Western and Northern China generally have larger geographic areas, the matched temperature in cooler counties may have systematically larger measurement errors than warmer places. As a robustness check to mitigate this concern, we exclude CFPS respondents residing in a set of large counties with geographic area above 3,000 km^2. Our main findings still hold. The results are available upon request.

Figure 5A first presents the estimated results by gender. When exposed to temperatures higher than 32 °C, the negative effect on math test scores is only significant for males. However, the effect is muted for the cognitive performance in verbal tests. As reported in Figure 5A, spending a day with a temperature above 32 °C, relative to a day in the 22–24 °C range, decreases males' math test scores by 0.485 (0.076 SD).

To explore whether the effects of high temperatures on cognition differ across age cohorts, we divide the sample into three age groups (10–30, 31–59, and 60+ years old). Figure 5B plots the estimates and CIs on the temperature bins for the three age cohorts, separately. Compared to the younger cohort, hot days are more harmful to middle-aged or older adults taking math tests, especially among older adults. Specifically, a day with a mean temperature above 32 °C, relative to a day in the 22–24 °C range, is associated with a 0.467 (0.074 SD) and 0.718 (0.113 SD) decline in math test scores for the middle-aged (aged 31–59) and the old people (aged 60 or above), respectively. However, our small sample of school-age children lacks sufficient statistical power to identify any effect for this age group, preventing us from meaningfully comparing our results with the growing literature on this same age group taking high-stakes tests.

Furthermore, the education level may have a significant effect on the response to extreme high temperatures. Dividing the whole sample into two subgroups at 12 years of education, Figure 5C shows that high temperatures impose a significant effect on the math tests of less educated people. The estimated coefficients indicate that a day with a mean temperature above 32 °C, relative to a day in the 22–24 °C range, leads to a reduction in math test scores by 0.441 (0.069 SD) for respondents who had received high school education level or below.

Finally, we examine the differential impacts of high temperatures on outdoor and indoor workers. Figure 5D confirms that extreme high temperatures impair math test performance among workers who work outdoors and thus who are more likely to be exposed to heat waves at work. In terms of the magnitude of the effect, a day with a mean temperature above 32 °C, relative to a day in the 22–24 °C range, is associated with a reduction in math test scores by 0.406 (0.064 SD) for respondents working outdoors.[12]

---

[12] As education level and working indoors/outdoors are generally correlated, and it is difficult to disentangle their influences. Consequently, we also examine the stratified results on exposure to high temperatures and

4.3. Robustness checks

We first conduct a placebo test to address the concern over potential omitted variables. Following a common strategy in the literature (Cho 2017), we examine the effect of extreme temperature the day after the interview on cognitive test scores. If unobserved factors are correlated with both the time trend of extreme temperatures and the outcome variables, we should find similar effects when replacing the current exposure with later ones. As a placebo best, Columns (1)–(2) of Table A4 display the estimates from regressions of the verbal and math test scores on the temperature bins on the day after the interview. None of the coefficients is statistically different from zero, which largely dismisses the concern over omitted variables.

The transitory effect of high temperatures on cognitive performance may be driven by behavioral change. First, people may become less cooperative or more impatient when exposed to extreme high temperatures, thereby reducing their cognitive test scores. CFPS includes an evaluation of interviewees' level of cooperation (waves 2010 and 2014) and impatience (waves 2014 and 2018), as rated by the interviewers.[13] The ratings for cooperation and impatience are both scaled from 1 (low) to 7 (high). We explore the effect of exposure to extreme high temperatures on respondents' cooperation and impatience in Columns (3) and (4) of Table A4, respectively. We still employ equation (1) as our specification, except we replace the previous dependent variable with ratings on cooperation or impatience. The estimates indicate there is no significant association between extreme high temperatures (>32 °C) and respondents' cooperation or impatience, partially ruling out the behavioral channel.

Another issue related to the interpretation of these results is potential fatigue and lower efforts during cognitive assessments on hot days, especially for low-stakes evaluations. In

---

the math test scores by education level and workplace. We find that only less educated people who are working outdoors are significantly affected by high temperatures (> 32 °C). The results are available upon request.

[13] Enumerators' assessments of cooperation and other behaviors can be subject to bias if enumerators themselves are impacted by high temperatures. Take ratings on cooperation as an example. As enumerators are more likely to give lower ratings on hot days, some of the identified negative effects of higher temperatures on cooperation ratings may stem from the enumerators, thus more likely overestimating the actual effect on respondents. The insignificant negative effect identified in Column (3) of Table A4 means that the net impact of high temperatures on respondents' cooperation behaviors is negligible, largely ruling out the potential bias from enumerators in the event of high temperatures.

particular, respondents may rush through the math tests on hot days, as they may feel the math tasks are more unpleasant. We do not have information on the completion time for each cognitive assessment other than the start and end time of the entire interview, which is an imprecise proxy for effort. Despite the shortcoming of the measure, we probe this channel using the time that each respondent takes to complete the whole questionnaire.[14] As shown in Column (5) of Table A4, there is no statistically significant relationship between temperature and the survey completion time, largely ruling out this channel.

The results are also robust to a wide range of alternative specifications. First, as revealed in Columns (1)–(2) of Table A5, the baseline results are robust to further controlling for county specific linear time trends and calendar date fixed effects. Second, we show in Table A4 that there is an insignificant effect of heat waves on completion time. Columns (3)–(4) of Table A5 further document that the results are essentially the same after adding completion time as a control variable. Third, considering that matching counties to the closest city with an air pollution reading may introduce measurement errors, we conduct a robustness check in Columns (5)–(6) by eliminating API from the regressions. We find that the estimated effects for math test scores even become slightly stronger. Additionally, as seen in Columns (7)–(8) of Table A5, our estimates appear to be unaffected by working outdoors.

Furthermore, Columns (1)–(2) of Table A6 show that our results are also robust to using the log form of the test scores as dependent variables. Moreover, as revealed in Columns (3)–(4) of Table A6, the estimated effects are qualitatively unchanged when only subjects with completed education levels are included in the analysis. Columns (5)–(6) of Table A6 indicate that migration, and thus location sorting, is unlikely to significantly bias our estimates. Columns (7)–(8) of Table A6 reveal that our findings still hold after we exclude ozone-dominated days, during which ozone may further interact with heat waves to impair cognition.

4.4. Cumulative effects

In the previous analysis, we found a significantly negative effect of transitory exposure to high temperatures on low-stakes math test performance. In this section, we add

[14] The average completion time in our sample is 52 minutes.

temperature bins for the month or year prior to the interview in the regression model. We first investigate the impact of exposure over the past month. We calculate the number of days falling in each temperature bin during the past 30 days, with Panel A of Figure A6 displaying the distribution. Panels A and B of Figure 10 plot the estimated coefficients associated with each temperature bin for the verbal and math test scores with 90% and 95% CIs. Neither of the coefficients on the highest temperature bin (>32 °C) is significant.[15]

Second, we also detect whether there are any long-run effects. For this, we divide the temperature spectrum into 10 bins (<-4 °C, -4–0 °C, 0–4 °C, 4–8 °C, 8–12 °C, 12–16 °C, 16–20 °C, 20–24 °C, 24–28 °C, and >28 °C) according to the distribution revealed in Panel B of Figure A7 and calculate the number of days falling in each temperature bin during the past year prior to the interview.[16] The temperature bin left out is 12–16 °C. The results in Panels C and D of Figure 10 indicate that exposure to both heat waves (>28 °C) and cold days (<-4 °C) in the past year impedes cognitive performance in math tests. The estimates reveal that spending 10 additional days in the year prior to the test with a daily average temperature above 28 °C, relative to a day in the 12–16 °C range, reduces math test scores by 0.093 (0.015 SD). These findings provide evidence that temperature extremes in the past year may impair cognition, a result consistent with prior findings (Park et al. 2020; Graff Zivin et al. 2020). It should be noted that this cumulative effect is likely mainly driven by the comparison of the >28 °C bin with the omitted temperature bin (12–16 °C) rather than some other bins.

4.5. Other mechanisms

The negative effect of high temperature on math test scores may also come through health channels. CFPS also contains some subjective health measures, such as sleeping

[15] We calculate the number of consecutive days above 32 °C immediately before the date of the survey and the largest number of consecutive hot days (> 32 °C) in the 30 days prior to the test date and study their effects on cognitive performance. We find that a one day increase in the largest number of consecutive hot days above 32 °C in the past 30 days leads to a reduction in math test scores by 0.069 (0.011 SD). The results are displayed in Table A7.

[16] The difference in bin size stems mainly from the difference in temperature distributions on the date of the interview, in the past 30 days, and in the past year. As most of the interviews were conducted in Summer, the distribution of temperatures on the interview date /in the past 30 days is skewed toward higher temperatures. Therefore, we divide the spectrum of temperatures into twelve bins, i.e., <12 °C, >32 °C, and ten 2 °C-wide bins in between. However, when it comes to the temperature throughout the past year, the range of days is significantly wider, spanning from extremely freezing winter days to extremely hot summer days. Hence, we divide the temperature spectrum into ten bins, i.e., <-4 °C, >28 °C, and eight 4 °C-wide bins in between.

patterns as well as overall well-being (physical and mental health). In this section, we examine the effect of extreme temperatures on a set of subjective health measures, including self-reported daily sleeping hours; self-rated health status, ranging from 1 (least healthy) to 5 (healthiest); K6 scores (0–24, the lower the better)[17]; life satisfaction, ranging from 1 (not satisfied at all) to 5 (very satisfied); and CES-D scores (0–24, the lower the better)[18]. For each measure, we study the transitory effect as well as the cumulative effects in the past month or year. The results are presented in Tables A8 through A10. We find that exposure to a temperature above 32 °C on the test date, relative to a day in the 22–24 °C range, reduces self-rated health status by 0.083 SD and increases CES-D scores by 0.108 SD. Besides, an additional day with a mean temperature above 32 °C in the past 30 days, relative to the reference temperature bin, is associated with a decline in self-rated health status by 0.009 SD. Furthermore, spending 10 additional days in the year prior to the test with a daily average temperature above 28 °C, relative to 10 days in the 12–16 °C range, leads to a reduction in daily sleeping hours by 0.027 SD.

5. Conclusions

By matching a nationally representative longitudinal survey with weather data according to the exact date and geographic location in China, this study examines the effect of transitory exposure to high temperatures on cognitive performance for people above 10 years old. Exploiting the longitudinal structure of CFPS and random fluctuations in weather across interviews, we identify the effect of temperatures in models with individual fixed effects. We find that exposure to a mean temperature above 32 °C on the test date, relative to a moderate day in the 22–24 °C range, leads to a decline in math test scores by 0.066

[17] The K6 scores are the total scores of six items in the Kessler Psychological Distress Scale (K6) developed by Kessler et al. (2003). In K6, interviewees are asked how often they have felt this way during the past month: 1) I felt so depressed that nothing could cheer me up; 2) I felt nervous; 3) I felt restless or fidgety; 4) I felt hopeless; 5) I felt that everything was an effort; 6) I felt that life was meaningless. The answers include 0) Never; 1) Sometimes; 2) Half the Time; 3) Often; 4) Almost Every Day.

[18] The Center for Epidemiologic Studies Depression scale (CES-D), which contains 20 items, is a self-reported scale designed to measure the current level of depressive symptomatology in the general population (Radloff 1977). CFPS chooses eight items from the standard 20-item CES-D scale to form a short version. The CES-D scores are the total scores of the eight items. In the short version, interviewees are asked how often they have felt this way during the past week: 1) I felt depressed. 2) I felt that everything I did was an effort. 3) My sleep was restless. 4) I was happy. 5) I felt lonely. 6) I enjoyed life. 7) I felt sad. 8) I could not get "going". The answers include 0) Rarely or none of the time; 1) Some or a little of the time; 2) Occasionally or a moderate amount of time; 3) Most or all the time.

SD, equivalent to a loss of 0.23 years of education. Further, the effect on the cognitive performance in math tests becomes more pronounced as people age, especially for males and the less educated. These results survive a placebo test and a set of robustness checks.

People living in hotter regions or with AC installed in their homes are not as vulnerable to extreme high temperatures, indicating some adaptation. Specifically, residential AC could mitigate the harmful effect of heat waves on math test scores by 36.6%. Yet, the adaptation is still limited. People residing in hotter regions are vulnerable to low temperatures, while people in cooler regions or without AC are susceptible to high temperatures, especially during math tests.

While this study mainly focuses on transitory exposures to heat events, the impact is sizable. First, compared to previous work, the cognitive tests in our study setting are close to our day-to-day, low-stakes cognitive activities. The salient effect in our setting suggests that the quality of routine math-related decision-making in our daily lives is compromised by temperature extremes. Second, as our identified effects of exposure to heat waves are both transitory and cumulative, more frequent high temperatures associated with climate change will likely impede cognitive performance in math domains more in the future. Finally, cognitive functions are essential for our everyday life. Damage to cognitive performance in math domains caused by extreme temperatures would compromise the quality of decision-making, generating inefficiencies and imposing additional costs on individual and social welfare. Previous studies evaluating the welfare cost of extreme temperatures have neglected its potential damage to cognition among older adults. As old people still need to make many critical decisions using math skills, the total social costs of heat waves, which are often inferred from the estimates on young people, are likely understated.

Acknowledgements

The IZA Discussion Paper Series serves as a preprint server to deposit latest research for feedback. We acknowledge the Institute of Social Science Survey at Peking University for providing us with the CFPS data. The study was approved by the Institutional Review Board (IRB) at Peking University (Approval No: IRB00001052-14010). All participants gave informed consent in accordance with policies of the IRB at Peking University. The authors acknowledge helpful comments by participants and discussants at various conferences, seminars, and workshops. Xi Chen acknowledges funding from the Yale Macmillan Center Faculty Research Fund, the US Federal PEPPER Center Scholar Award (P30AG021342), two NIH/NIA Grants (R01AG077529; K01AG053408), and Yale Alzheimer 's Disease Research Center (P30AG066508). Xiaobo Zhang and Xin Zhang respectively acknowledge financial support from the National Natural Science Foundation of China (72192844 and 72003014). The authors declare that they have no conflict of interest.

References

Agarwal, Sumit, Yu Qin, Luwen Shi, Guoxu Wei, and Hongjia Zhu. 2021. Impact of temperature on morbidity: New evidence from China. *Journal of Environmental Economics and Management* 109.

Banerjee, Rakesh, and Riddhi Maharaj. 2020. Heat, infant mortality, and adaptation: Evidence from India. *Journal of Development Economics* 143 (August 2019): 102378.

Barreca, Alan, Karen Clay, Olivier Deschenes, Michael Greenstone, and Joseph S Shapiro. 2016. Adapting to climate change: The remarkable decline in the US temperature-mortality relationship over the Twentieth Century. *Journal of Political Economy* 124 (1): 105–59.

Behrer, A Patrick, and Jisung Park. 2017. Will We Adapt? Temperature, Labor and Adaptation to Climate Change. *Working Paper*: 39p.

Biddle, Jeff. 2008. Explaining the spread of residential air conditioning, 1955-1980. *Explorations in Economic History* 45 (4): 402–23.

Burke, Marshall, Felipe González, Patrick Baylis, Sam Heft-Neal, Ceren Baysan, Sanjay Basu, and Solomon Hsiang. 2018. Higher temperatures increase suicide rates in the United States and Mexico. *Nature Climate Change* 8 (8): 723–29.

Cai, Xiqian, Yi Lu, Jessica Pan, and Songfa Zhong. 2019. Gender gap under pressure: Evidence from China's national college entrance examination. *Review of Economics and Statistics* 101 (2): 249–63.

Cho, Hyunkuk. 2017. The effects of summer heat on academic achievement: A cohort analysis. *Journal of Environmental Economics and Management* 83: 185–96.

Davis, Lucas W, and Paul J Gertler. 2015. Contribution of air conditioning adoption to future energy use under global warming. *Proceedings of the National Academy of Sciences of the United States of America* 112 (19): 5962–67.

Deschenes, Olivier. 2014. Temperature, human health, and adaptation: A review of the empirical literature. *Energy Economics* 46: 606–19.

Deschênes, Olivier, and Michael Greenstone. 2007. The economic impacts of climate change: Evidence from agricultural output and random fluctuations in weather. *American Economic Review* 97 (1): 354–85.

------. 2011. Climate change, mortality, and adaptation: Evidence from annual fluctuations in weather in the US. *American Economic Journal: Applied Economics* 3 (4): 152–85.

Ebenstein, Avraham, Victor Lavy, and Sefi Roth. 2016. The long-run economic consequences of high- stakes examinations: Evidence from transitory variation in pollution. *American Economic Journal: Applied Economics* 8 (4): 36–65.

Garg, Teevrat, Maulik Jagnani, and Vis Taraz. 2020. Temperature and Human Capital in India. *Journal of the Association of Environmental and Resource Economists* 7 (6): 1113–50.

Graff Zivin, Joshua, Solomon M Hsiang, and Matthew Neidell. 2018. Temperature and human capital in the short and long run. *Journal of the Association of Environmental and Resource Economists* 5 (1): 77–105.

Graff Zivin, Joshua, and Matthew Neidell. 2014. Temperature and the allocation of time: Implications for climate change. *Journal of Labor Economics* 32 (1): 1–26.

Graff Zivin, Joshua, Yingquan Song, Qu Tang, and Peng Zhang. 2020. Temperature and high-stakes cognitive performance: Evidence from the national college entrance examination in China. *Journal of Environmental Economics and Management* 104: 102365.

Hansen, James, Makiko Sato, Reto Ruedy, Ken Lo, David W Lea, and Martin Medina-Elizade. 2006. Global temperature change. *Proceedings of the National Academy of Sciences of the United States of America* 103 (39): 14288–93.

Heutel, Garth, Nolan H Miller, and David Molitor. 2021. Adaptation and the Mortality Effects of Temperature Across U.S. Climate Regions. *The Review of Economics and Statistics*: 1–14.

Hocking, Chris, Richard B Silberstein, Wai Man Lau, Con Stough, and Warren Roberts. 2001. Evaluation of cognitive performance in the heat by functional brain imaging and psychometric testing. *Comparative Biochemistry and Physiology - A Molecular and Integrative Physiology* 128 (4): 719–34.

Hsiang, Solomon, Robert Kopp, Amir Jina, James Rising, Michael Delgado, Shashank Mohan, DJ Rasmussen, Robert Muir-Wood, Paul Wilson, Michael Oppenheimer, Kate Larsen, and Trevor Houser. 2017. Estimating economic damage from climate change in the United States. *Science* 356 (6345): 1362–69.

Huang, Cunrui, Adrian G Barnett, Xiaoming Wang, and Shilu Tong. 2012. The impact of temperature on years of life lost in Brisbane, Australia. *Nature Climate Change* 2 (4): 265–

70.
IPCC. 2021. *Climate Change 2021: The Physical Science Basis*. Ed. V. Masson-Delmotte, P. Zhai, A. Pirani, S.L. Connors, C. Péan, S. Berger, N. Caud, Y. Chen, L. Goldfarb, M.I. Gomis, M. Huang, K. Leitzell, E. Lonnoy, J.B.R. Matthews, T.K. Maycock, T. Waterfield, O. Yelekçi, R. Yu, and B. Zhou. Cambridge University Press. In Press.

Isen, Adam, Maya Rossin-Slater, Reed Walker, and V Kerry Smith. 2017. Relationship between season of birth, temperature exposure, and later life wellbeing. *Proceedings of the National Academy of Sciences of the United States of America* 114 (51): 13447–52.

Karlsson, Martin, and Nicolas R Ziebarth. 2018. Population health effects and health-related costs of extreme temperatures: Comprehensive evidence from Germany. *Journal of Environmental Economics and Management* 91: 93–117.

Kessler, RC, PR Barker, LJ Colpe, JF Epstein, JC Gfroerer, E Hiripi, MJ Howes, SL Normand, RW Manderscheid, EE Walters, and AM Zaslavsky. 2003. Screening for serious mental illness in the general population. *Archives of General Psychiatry* 60: 184–89.

Kiyatkin, Eugene A. 2007. Brain temperature fluctuations during physiological and pathological conditions. *European Journal of Applied Physiology* 101 (1): 3–17.

Lee, Wang Sheng, and Ben G Li. 2021. Extreme weather and mortality: Evidence from two millennia of Chinese elites. *Journal of Health Economics* 76: 102401.

McMorris, Terry, Jon Swain, Marcus Smith, Jo Corbett, Simon Delves, Craig Sale, Roger C Harris, and Julia Potter. 2006. Heat stress, plasma concentrations of adrenaline, noradrenaline, 5-hydroxytryptamine and cortisol, mood state and cognitive performance. *International Journal of Psychophysiology* 61 (2): 204–15.

Mullins, Jamie T, and Corey White. 2019. Temperature and mental health: Evidence from the spectrum of mental health outcomes. *Journal of Health Economics* 68.

Nybo, Lars, Peter Rasmussen, and Michael N Sawka. 2014. Performance in the heat-physiological factors of importance for hyperthermia-induced fatigue. *Comprehensive Physiology* 4 (2): 657–89.

Obradovich, Nick, Robyn Migliorini, Martin P Paulus, and Iyad Rahwan. 2018. Empirical evidence of mental health risks posed by climate change. *Proceedings of the National Academy of Sciences of the United States of America* 115 (43): 10953–58.

Park, Jisung, Joshua Goodman, Michael Hurwitz, and Jonathan Smith. 2020. Heat and

Learning. *American Economic Journal: Economic Policy* 12 (2): 306–39.

Park, R Jisung. 2022. Hot Temperature and High Stakes Performance. *Journal of Human Resources* 57 (2): 400–434.

Park, R Jisung, A Patrick Behrer, and Joshua Goodman. 2021. Learning is inhibited by heat exposure, both internationally and within the United States. *Nature Human Behaviour* 5 (1): 19–27.

Radloff, Lenore Sawyer. 1977. The CES-D Scale : A Self-Report Depression Scale for Research in the General Population. *Applied Psychological Measurement* 1 (3): 385–401.

Raichle, Marcus E, and Mark A Mintun. 2006. Brain Work and Brain Imaging. *Annual Review of Neuroscience* 29 (1): 449–76.

Shah, Manisha, and Bryce Millett Steinberg. 2017. Drought of opportunities: Contemporaneous and long-term impacts of rainfall shocks on human capital. *Journal of Political Economy* 125 (2): 527–61.

Vasmatzidis, Ioannis, Robert E Schlegel, and Peter A Hancock. 2002. An investigation of heat stress effects on time-sharing performance. *Ergonomics* 45 (3): 218–39.

Xie, Yu, and Jingwei Hu. 2014. An Introduction to the China Family Panel Studies (CFPS). *Chinese Sociological Review* 47 (1): 3–29.

Zhang, Xin, Xi Chen, and Xiaobo Zhang. 2018. The impact of exposure to air pollution on cognitive performance. *Proceedings of the National Academy of Sciences of the United States of America* 115 (37): 9193–97.

Figure 1 Distribution of daily mean temperature on the test date

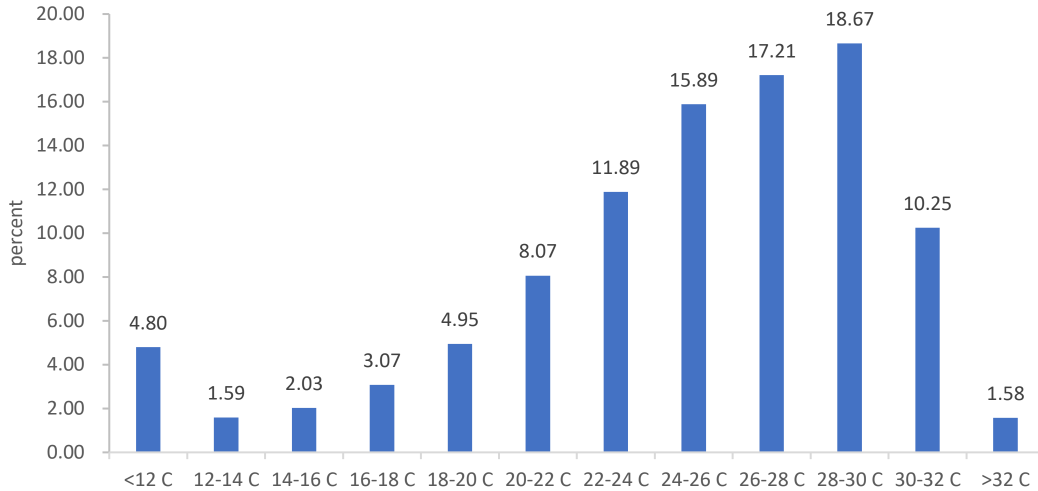
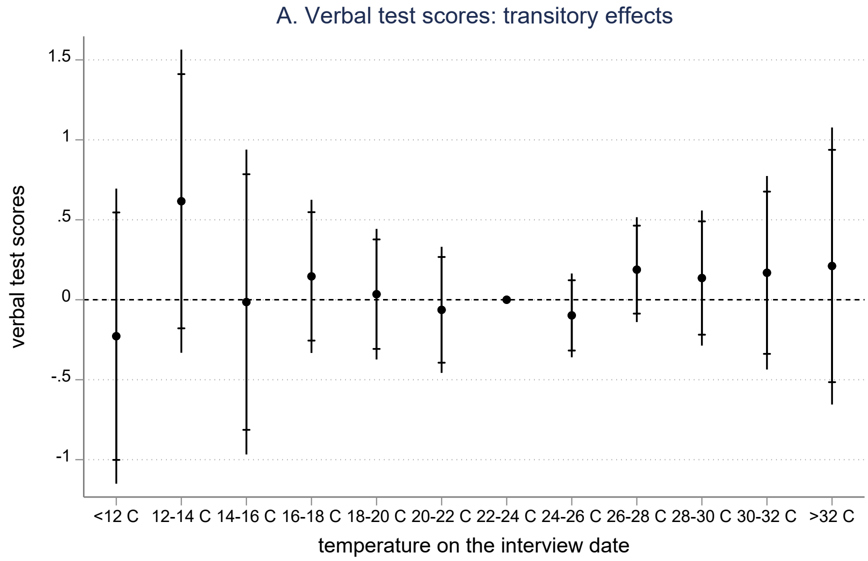


Figure 2 Transitory effects of temperatures on cognitive test scores



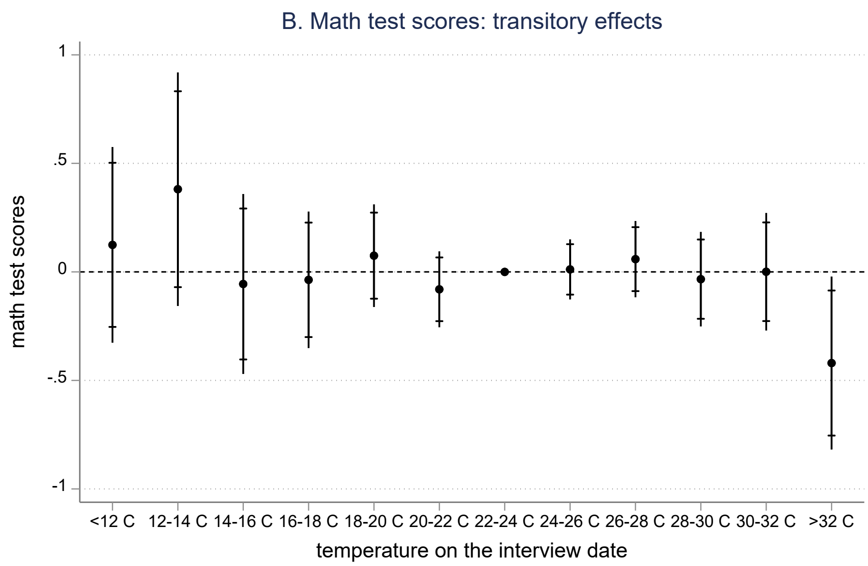
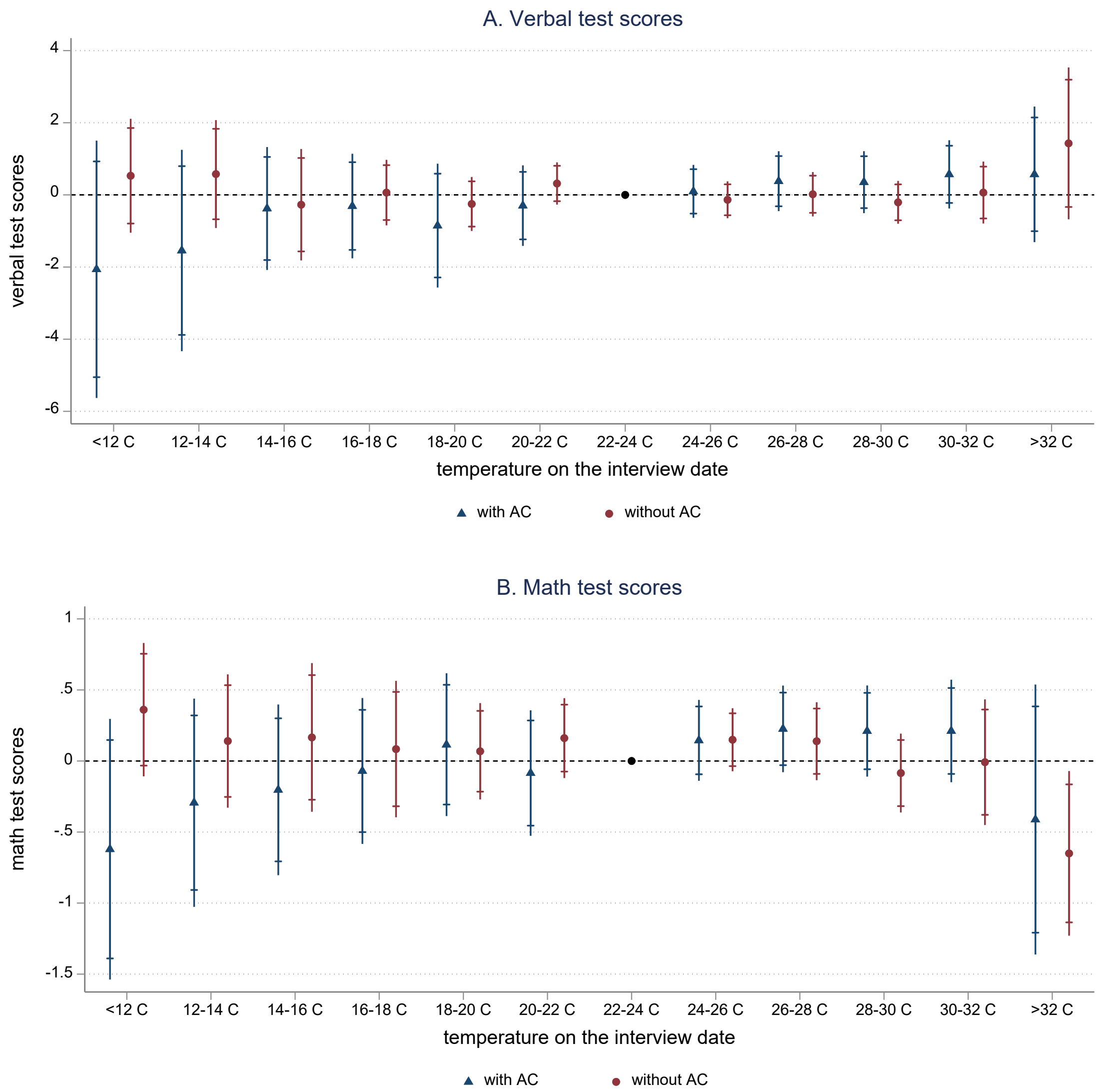
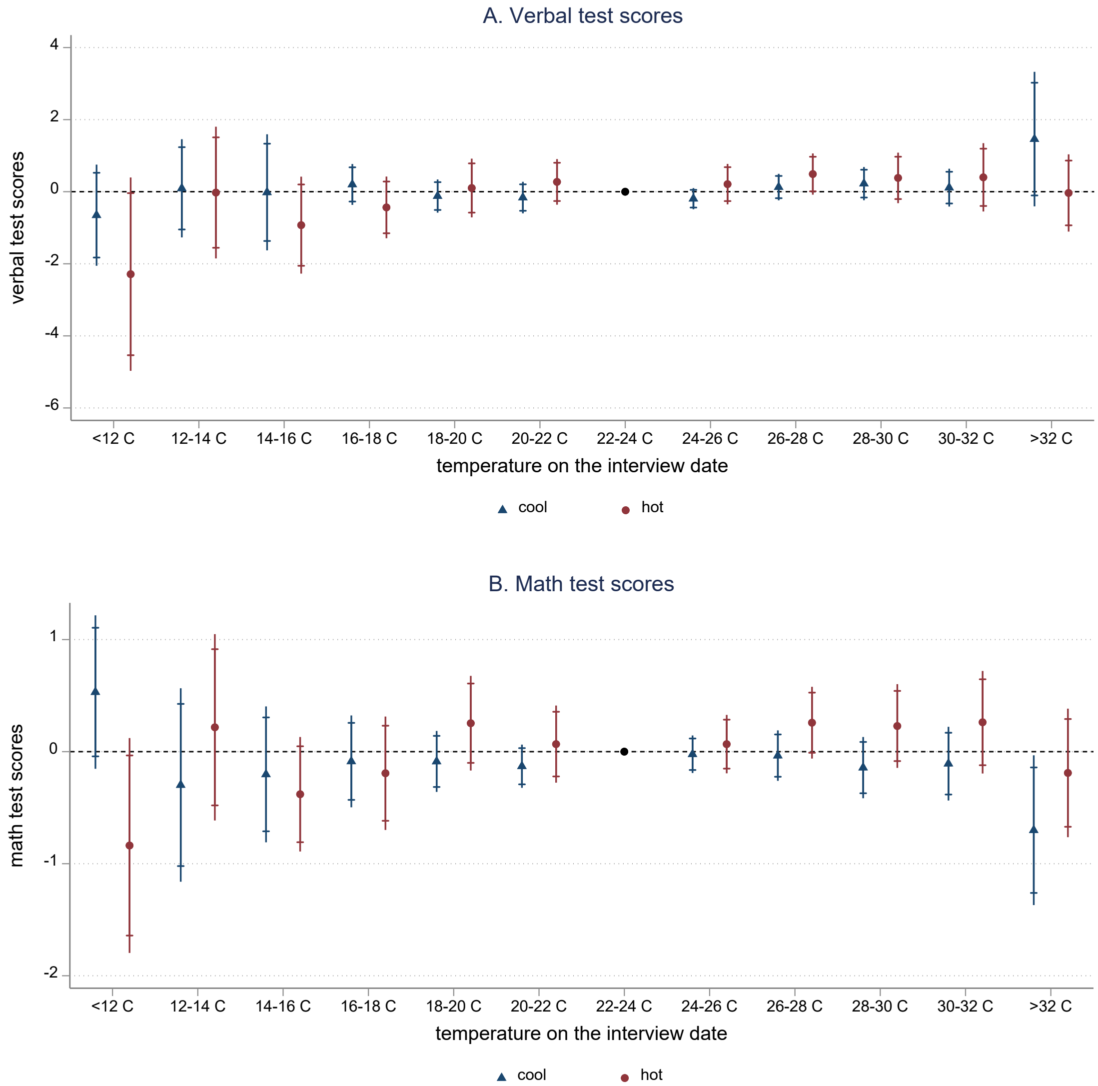


Figure 3 Transitory effects of temperatures on cognitive test scores, by residential AC ownership



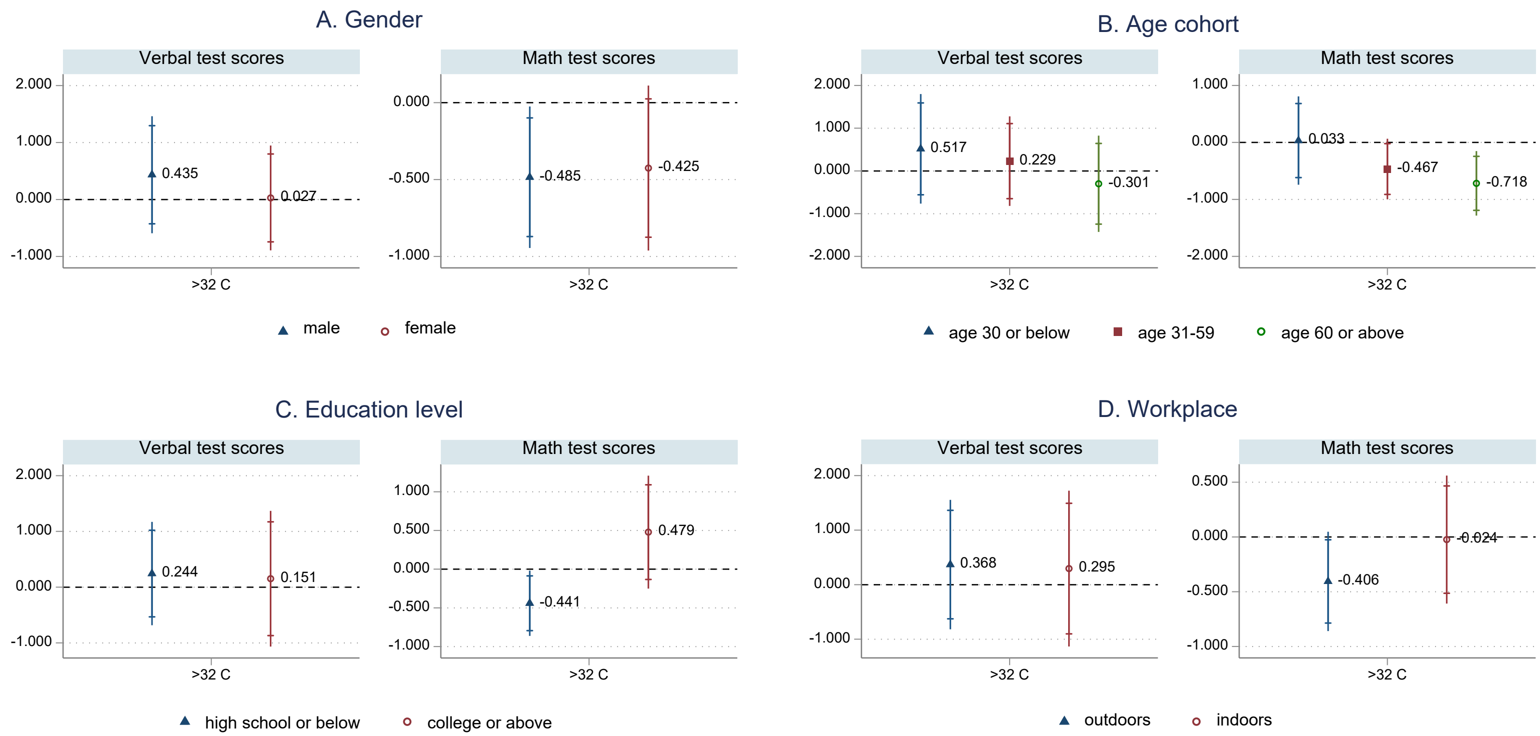
Note: The figures plot the estimated coefficients on temperature bins for the households with and without AC based on the results in Columns (1)-(2) and (4)-(5) of Table A2. Both 90 (short caps) and 95 percent (long lines) confidence intervals are displayed. The left-out temperature bin is 22-24 °C. The coefficients can be interpreted as effects of a day in the corresponding temperature bin on cognitive test scores relative to the reference temperature category. Panel A refers to verbal test scores, while Panel B refers to math test scores.

Figure 4 Transitory effects of temperatures on cognitive test scores, by region



Note: The figures plot the estimated coefficients on temperature bins for the cool and hot regions based on the results in Columns (1)-(2) and (4)-(5) of Table A3. Both 90 (short caps) and 95 percent (long lines) confidence intervals are displayed. The left-out temperature bin is 22-24 °C. The coefficients can be interpreted as effects of a day in the corresponding temperature bin on cognitive test scores relative to the reference temperature category. Panel A refers to verbal test scores, while Panel B refers to math test scores.

Figure 5 Stratification effects of temperatures on cognitive test scores



Note: Panels A-D plot the stratification effects of high temperatures on cognitive test scores by gender, age cohort, education level and workplace, respectively. All the regressions control the temperature bins "<12 °C, 12-14 °C, 14-16 °C, 16-18 °C, 18-20 °C, 20-22 °C, 22-24 °C, 24-26 °C, 26-28 °C, 28-30 °C, 30-32 °C, and >32 °C" and the figure only plots the estimated coefficients on the temperature bin ">32 °C". The left-out temperature bin is 22-24 °C. All the regressions include individual fixed effects, county fixed effects, year, month, day-of-week, and hour-of-day fixed effects. Demographic controls include age with its square term and education level. Environmental controls include air pollution index (API), total precipitation, wind speed, sunshine duration, and relative humidity in quadratic forms. Both 90 (short caps) and 95 percent (long lines) confidence intervals are displayed. The coefficients can be interpreted as effects of a day in the corresponding temperature bin on cognitive test scores relative to the reference temperature category.

Table 1 Summary statistics

Variable	whole sample		normal temp. (< 32°C)		high temp. (> 32 °C)		difference
	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.	p-value
CFPS data							
verbal test scores	18.536	10.539	18.528	10.541	19.031	10.424	0.113
math test scores	10.220	6.351	10.229	6.353	9.659	6.172	0.003
gender	0.487	0.500	0.488	0.500	0.456	0.498	0.036
age	44.609	18.544	44.566	18.532	47.238	19.106	0.000
years of education	7.433	4.620	7.435	4.619	7.363	4.664	0.604
Temperature							
mean temperature, °C	24.397	6.309	24.260	6.265	32.887	0.834	0.000
indicator for days >32 °C	0.016	0.125	0.000	0.000	1.000	0.000	-
Environmental controls							
API	64.557	35.281	64.585	35.506	62.785	15.921	0.090
precipitation, *mm*	4.281	12.382	4.345	12.464	0.320	3.197	0.000
wind speed, *m/s*	2.102	1.061	2.101	1.065	2.172	0.747	0.027
sunshine duration, *hour*	6.383	4.223	6.313	4.215	10.719	1.697	0.000
relative humidity, *%*	74.602	12.599	74.805	12.555	61.975	8.137	0.000

Table 2 Relevant studies on the effects of high temperatures on cognitive performance

Study	Country	Years	Outcomes	Temperature bins	Matching method	Effects
Graff Zivin et al. (2018) JAERE	the United States (951 counties)	1988-2006	mathematics, reading recognition, and reading comprehension from NLSY79	indicators for temperature in 2 °C-wide bins from 12 °C to 32 °C, with 20-22 °C as the reference category	linearly interpolate temperatures at each county centroid using readings from the seven nearest stations	Changing the daily mean temperature from 20 °C–22 °C to 30 °C–32 °C decreases a child's math scores by 0.12 SD. The effect of temperature bin above 32 °C on math test scores is insignificant *each 1 °C higher temperature decreases test scores by 0.0120 (0.12/10) SD*
Graff Zivin et al. (2020) JEEM	China (2227 counties)	2005-2011	total scores for National college entrance examination	indicators for temperature in 2 °C-wide bins from 12 °C to 28 °C, with 12-14 °C as the reference category	calculate weather for a given county based on inverse-distance weighted averages of readings from all weather stations within a 200 km radius of the county centroid	Exposure to a daily mean temperature above 28 °C, relative to a day in the 12–14 °C range, leads to a reduction in total exam scores by 0.0553 log points (convert to 0.48 SD) *each 1 °C higher temperature decreases test scores by 0.0320 (0.48/15) SD*
Cho (2017) JEEM	Korea (164 cities)	2009-2013	reading, math and English test scores for Korean college entrance exam	indicators for maximum temperature in 2 °C-wide bins from 22 °C to 34 °C, with 28-30 °C as the reference category	match weather data based on school location at the city level (weather data from adjacent cities is used for eight cities without weather information)	An additional day with a maximum daily temperature above 34 °C during the summer, relative to a day with a maximum daily temperature in the 28–30 °C range, reduces the math test scores by 0.0042 SD *each 1 °C higher temperature decreases test scores by 0.0008 (0.0042/5) SD*
Garg et al. (2020) JAERE	India	2006-2014; 2002-2011	math and reading test scores for children in primary and secondary school	number of days in the calendar year prior to the year of the test falling in 2 °C-wide bins from 13 °C to 29 °C, with 15-17 °C as the reference category	construct an inverse-distance weighted average of all the weather grid points (on a 1×1° latitude-longitude grid level) within a 100-kilometer range of the district centroid	10 extra days with average daily temperature above 29°C (85°F) during the prior year, relative to 15°–17°C (59°–63°F), reduce math and reading test scores by 0.03 and 0.02 SD, respectively
Park (2022) JHR	the United States (New York City)	1998-2011	high stakes exam scores, likelihood of passing exam, educational attainment	indicators for temperature in >90 °F, 80-90 °F, 70-80 °F; >90 °F, 85-90 °F, 80-85 °F, 75-80 °F, 70-75 °F	match schools to the nearest weather station in the NYC area	Taking an exam when outdoor temperatures are 90 °F reduces performance by approximately 13 percent of a SD relative to a temperature of 75 °F, and results in a roughly 10 percent lower likelihood of passing a particular subject.
Park et al. (2020) AEJ EP	the United States	1998-2012	the PSAT math and reading scores	number of days in the year prior to the test falling in 10 °F-wide bins from 40 °F to 100 °F, with 60-70 °F as the reference category	assign each high school to the nearest weather station, the average matching distance is 9.7 miles	Three additional days above 90 °F in the year prior to the test lower achievement by 0.002 SD

Note: Our estimates indicate that each 1 °C higher temperature decreases math test scores by 0.0073 (0.066/9) SD. For the studies closest to ours, we mark their identified effect sizes in *italic fonts* using a comparable metric based on a linear approximation.

Appendix A: Supplementary Figures and Tables

Figure A1 The distribution of weather stations

Note: This figure is plotted using ArcMap 10.3.1.

Figure A2 Histogram of mean temperature (°C) on the test date

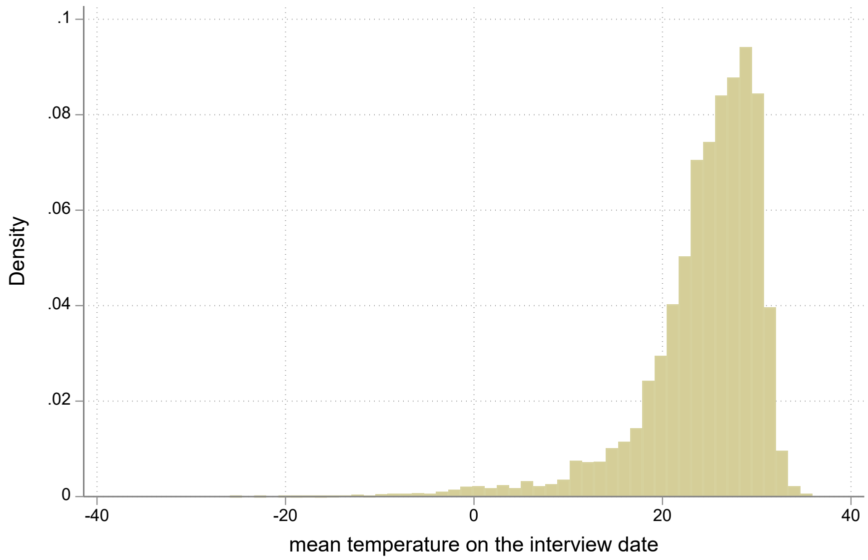


Figure A3 Distribution of interview months in 2010, 2014 and 2018

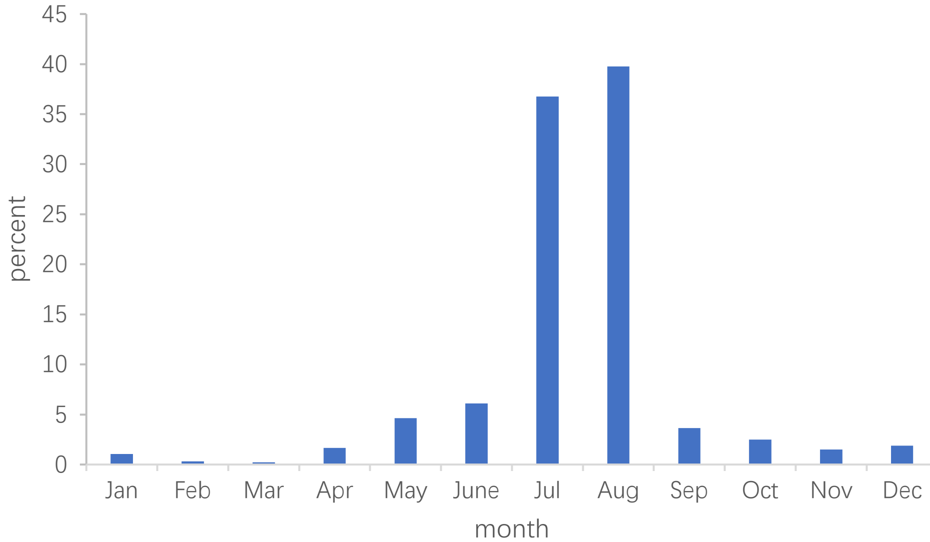
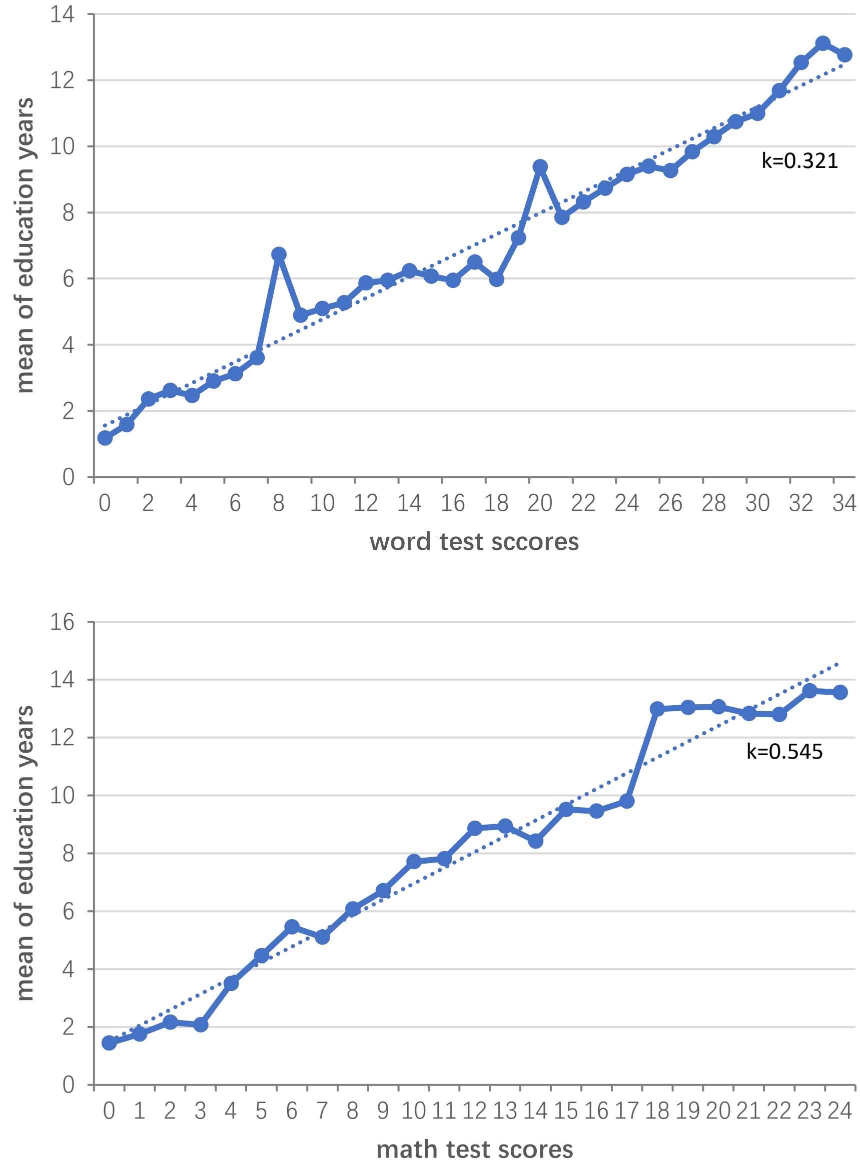


Figure A4 Relationship between cognitive test scores and mean values of education years



Note: k values indicate the coefficients from regressing mean values of education years on verbal test scores/math test scores.

Figure A5 Distribution of mean temperature (°C) in the past 30 days and the past year

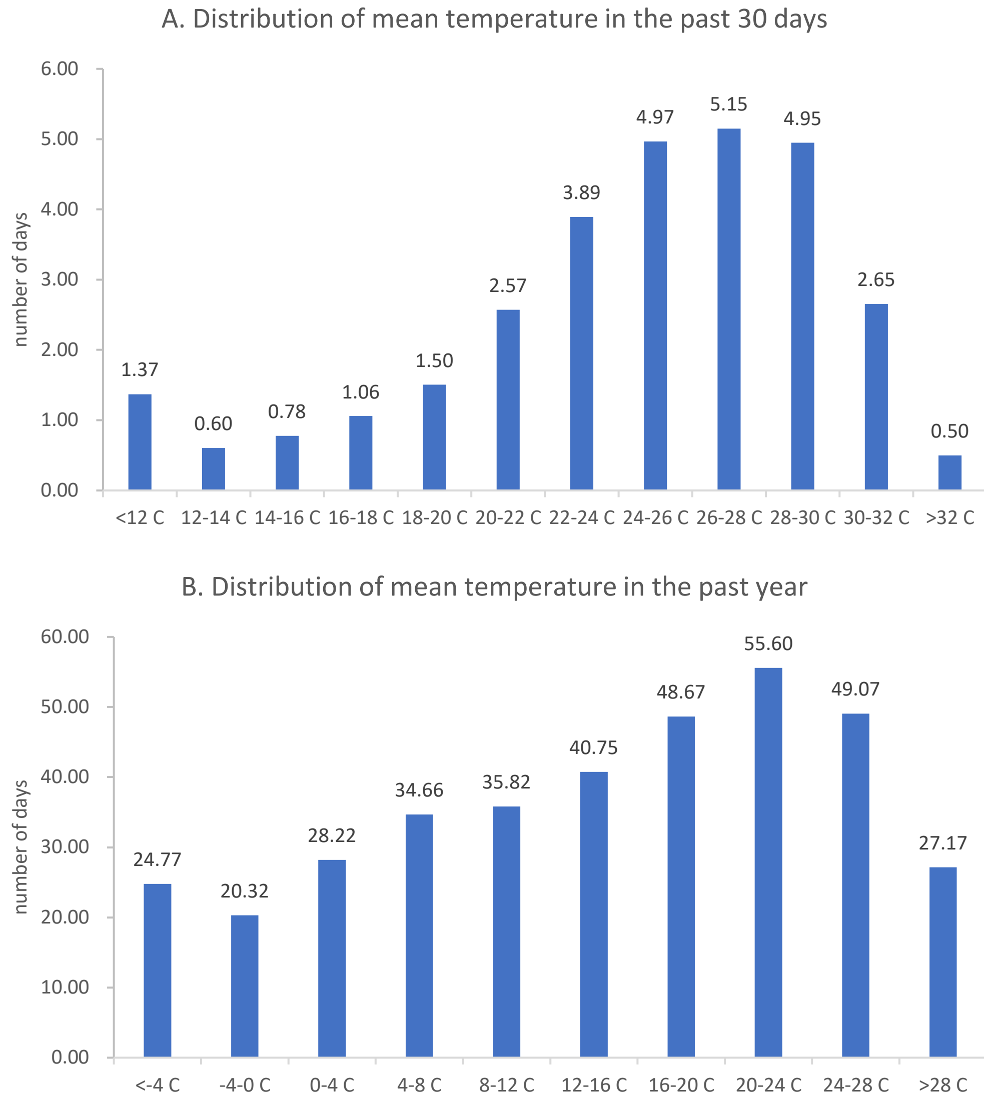
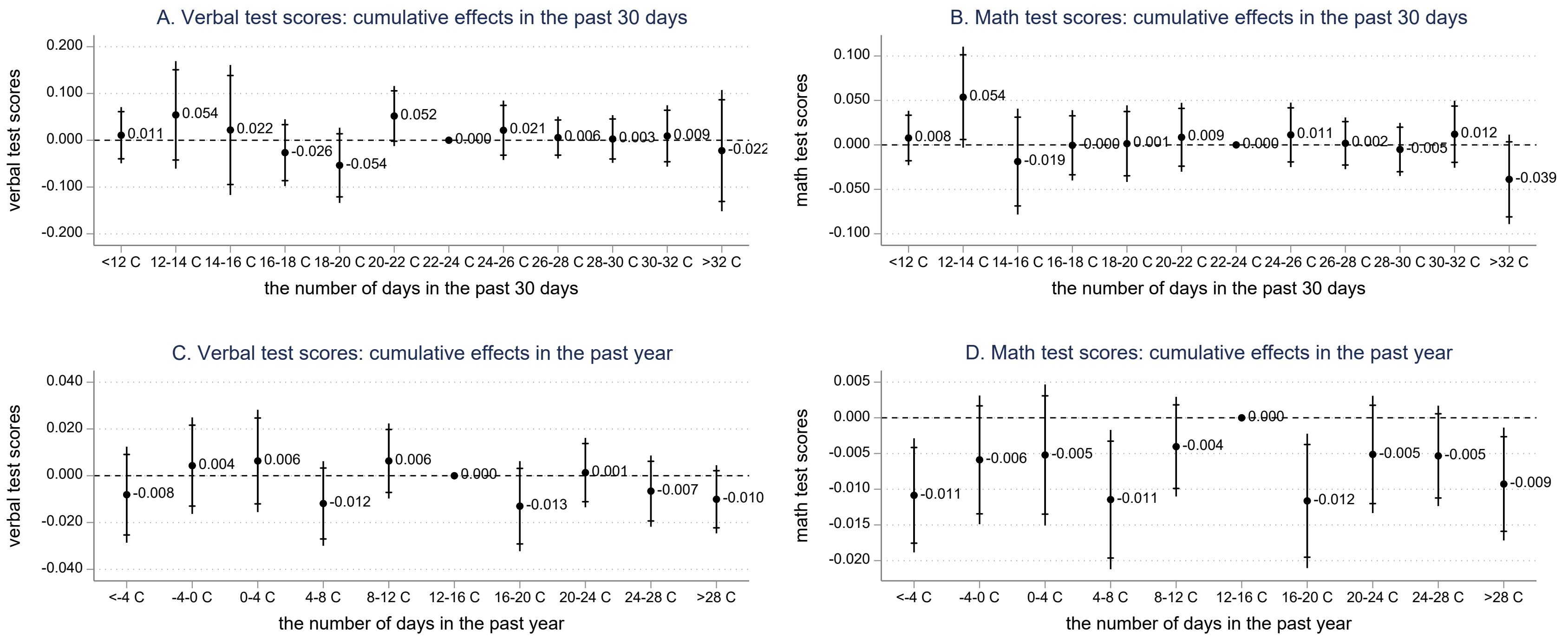


Figure A6 Cumulative effects of temperatures on cognitive test scores



Note: Panels A and B plot the estimated coefficients on the number of days in each temperature bin "<12 °C, 12-14 °C, 14-16 °C, 16-18 °C, 18-20 °C, 20-22 °C, 22-24 °C, 24-26 °C, 26-28 °C, 28-30 °C, 30-32 °C, and >32 °C" during the past 30 days. The left-out temperature bin is 22-24 °C. Panels C and D plot the estimated coefficients on the number of days in each temperature bin "<-4 °C, -4-0 °C, 0-4 °C, 4-8 °C, 8-12 °C, 12-16 °C, 16-20 °C, 20-24 °C, 24-28 °C, and >28 °C" during the past year. The left-out temperature bin is 12-16 °C. Both 90 (short caps) and 95 percent (long lines) confidence intervals are displayed. The coefficients can be interpreted as effects of an additional day in the corresponding temperature bin on cognitive test scores relative to the reference temperature category. All the regressions control the 12 temperature bins ("<12 °C, 12-14 °C, 14-16 °C, 16-18 °C, 18-20 °C, 20-22 °C, 22-24 °C, 24-26 °C, 26-28 °C, 28-30 °C, 30-32 °C, and >32 °C") on the interview date. Other controls include individual fixed effects, county fixed effects, year, month, day-of-week, and hour-of-day fixed effects. Demographic controls include age with its square term and education level. Environmental controls include air pollution index (API), total precipitation, wind speed, sunshine duration, and relative humidity in quadratic forms. Robust standard errors are clustered at the county level.

Table A1 Transitory effects of temperatures on cognitive test scores

Dependent variable	A. verbal test scores		B. math test scores	
	(1)	(2)	(3)	(4)
temperature bins				
<12 °C	-0.263	-0.228	0.307	0.125
	(0.436)	(0.469)	(0.231)	(0.229)
12-14 °C	0.402	0.616	0.300	0.381
	(0.420)	(0.482)	(0.203)	(0.274)
14-16 °C	-0.400	-0.014	0.047	-0.056
	(0.374)	(0.485)	(0.147)	(0.211)
16-18 °C	0.197	0.146	0.156	-0.037
	(0.242)	(0.244)	(0.125)	(0.160)
18-20 °C	-0.005	0.035	0.066	0.075
	(0.192)	(0.208)	(0.097)	(0.120)
20-22 °C	-0.033	-0.063	-0.062	-0.080
	(0.177)	(0.200)	(0.069)	(0.089)
22-24 °C				
24-26 °C	-0.007	-0.098	0.020	0.012
	(0.125)	(0.133)	(0.054)	(0.070)
26-28 °C	0.155	0.188	0.036	0.059
	(0.143)	(0.167)	(0.064)	(0.089)
28-30 °C	-0.005	0.136	-0.050	-0.033
	(0.161)	(0.215)	(0.074)	(0.111)
30-32 °C	0.108	0.169	-0.002	0.001
	(0.219)	(0.308)	(0.088)	(0.138)
>32 °C	0.315	0.211	-0.342**	-0.420**
	(0.426)	(0.440)	(0.160)	(0.203)
demographic controls	Yes	Yes	Yes	Yes
environmental controls	Yes	Yes	Yes	Yes
individual fixed effects	No	Yes	No	Yes
county fixed effects	Yes	Yes	Yes	Yes
interview year, month, day-of-week, and hour-of-day fixed effects	Yes	Yes	Yes	Yes
Observations	70,736	70,736	70,736	70,736
Adjusted *R*-squared	0.622	0.072	0.751	0.178

Note: The left-out temperature bin is 22-24 °C. The coefficients can be interpreted as effects of a day in the corresponding temperature bin on cognitive test scores relative to the reference temperature category. Demographic controls include gender, age with its square term and education level. Environmental controls include air pollution index (API), total precipitation, wind speed, sunshine duration, and relative humidity in quadratic forms. Robust standard errors, clustered at the county level, are presented in parentheses. * 10% significance level; ** 5% significance level; *** 1% significance level.

Table A2 Transitory effects of temperatures on cognitive test scores, by AC ownership

Dependent variable	verbal test scores			math test scores		
	with AC	without AC	difference	with AC	without AC	difference
	(1)	(2)	(3)	(4)	(5)	(6)
temperature bins						
<12 °C	-2.063	0.530	-2.593	-0.621	0.361	-0.982*
	(1.811)	(0.802)	(2.062)	(0.466)	(0.238)	(0.550)
12-14 °C	-1.542	0.577	-2.120	-0.294	0.140	-0.434
	(1.417)	(0.760)	(1.475)	(0.372)	(0.238)	(0.456)
14-16 °C	-0.377	-0.272	-0.105	-0.203	0.166	-0.369
	(0.865)	(0.785)	(1.131)	(0.305)	(0.266)	(0.390)
16-18 °C	-0.309	0.064	-0.373	-0.070	0.084	-0.154
	(0.736)	(0.461)	(0.801)	(0.261)	(0.244)	(0.350)
18-20 °C	-0.851	-0.251	-0.600	0.115	0.068	0.046
	(0.871)	(0.380)	(0.977)	(0.255)	(0.172)	(0.326)
20-22 °C	-0.298	0.316	-0.614	-0.085	0.161	-0.246
	(0.567)	(0.297)	(0.628)	(0.224)	(0.143)	(0.272)
22-24 °C						
24-26 °C	0.098	-0.136	0.233	0.145	0.150	-0.004
	(0.373)	(0.260)	(0.429)	(0.144)	(0.113)	(0.173)
26-28 °C	0.380	0.018	0.362	0.226	0.139	0.087
	(0.421)	(0.312)	(0.485)	(0.155)	(0.139)	(0.197)
28-30 °C	0.353	-0.206	0.559	0.211	-0.085	0.296
	(0.436)	(0.301)	(0.493)	(0.163)	(0.141)	(0.211)
30-32 °C	0.570	0.065	0.504	0.211	-0.008	0.220
	(0.480)	(0.435)	(0.638)	(0.183)	(0.225)	(0.301)
>32 °C	0.570	1.429	-0.859	-0.412	-0.650**	0.238
	(0.954)	(1.069)	(1.384)	(0.482)	(0.294)	(0.571)
Observations	11,052	13,807		11,052	13,807	
Adjusted *R*-squared	0.604	0.619		0.759	0.748	

Note: Based on data from CFPS 2014. The left-out temperature bin is 22-24 °C. The coefficients can be interpreted as effects of a day in the corresponding temperature bin on cognitive test scores relative to the reference temperature category. All the regressions include county fixed effects, year, month, day-of-week, and hour-of-day fixed effects. Demographic controls include gender, age with its square term and education level. Environmental controls include air pollution index (API), total precipitation, wind speed, sunshine duration, and relative humidity in quadratic forms. The results in Columns (3) and (6) indicate the differences of temperature bins between the AC and without AC groups. Robust standard errors, clustered at the county level, are presented in parentheses. * 10% significance level; ** 5% significance level; *** 1% significance level.

Table A3 Transitory effects of temperatures on cognitive test scores, by region

Dependent variable	verbal test scores			math test scores		
	cool	hot	difference	cool	hot	difference
	(1)	(2)	(3)	(4)	(5)	(6)
temperature bins						
<12 °C	-0.651	-2.288*	1.637	0.532	-0.838*	1.369**
	(0.712)	(1.362)	(1.541)	(0.347)	(0.487)	(0.599)
12-14 °C	0.094	-0.024	0.118	-0.298	0.217	-0.515
	(0.691)	(0.928)	(1.159)	(0.438)	(0.422)	(0.610)
14-16 °C	-0.017	-0.927	0.910	-0.203	-0.380	0.177
	(0.817)	(0.683)	(1.067)	(0.308)	(0.259)	(0.403)
16-18 °C	0.202	-0.435	0.637	-0.087	-0.193	0.106
	(0.287)	(0.435)	(0.522)	(0.208)	(0.257)	(0.331)
18-20 °C	-0.120	0.103	-0.223	-0.087	0.254	-0.341
	(0.234)	(0.414)	(0.476)	(0.138)	(0.215)	(0.256)
20-22 °C	-0.163	0.273	-0.435	-0.130	0.067	-0.197
	(0.222)	(0.321)	(0.391)	(0.098)	(0.175)	(0.201)
22-24 °C						
24-26 °C	-0.194	0.210	-0.404	-0.023	0.067	-0.090
	(0.149)	(0.285)	(0.321)	(0.085)	(0.132)	(0.157)
26-28 °C	0.129	0.490*	-0.362	-0.035	0.259	-0.294
	(0.187)	(0.291)	(0.346)	(0.114)	(0.163)	(0.199)
28-30 °C	0.225	0.383	-0.159	-0.143	0.229	-0.371
	(0.234)	(0.356)	(0.427)	(0.139)	(0.190)	(0.236)
30-32 °C	0.113	0.399	-0.286	-0.107	0.262	-0.370
	(0.266)	(0.481)	(0.551)	(0.167)	(0.232)	(0.287)
>32 °C	1.460	-0.035	1.495	-0.701**	-0.190	-0.511
	(0.947)	(0.544)	(1.094)	(0.339)	(0.291)	(0.448)
Observations	20,833	20,498		35,311	35,425	
Adjusted *R*-squared	0.067	0.083		0.181	0.181	

Note: The left-out temperature bin is 22-24 °C. The coefficients can be interpreted as effects of a day in the corresponding temperature bin on cognitive test scores relative to the reference temperature category. All the regressions include individual fixed effects, county fixed effects, year, month, day-of-week, and hour-of-day fixed effects. Demographic controls include age with its square term and education level. Environmental controls include air pollution index (API), total precipitation, wind speed, sunshine duration, and relative humidity in quadratic forms. The results in Columns (3) and (6) indicate the differences of temperature bins between the cool and hot regions. Robust standard errors, clustered at the county level, are presented in parentheses. * 10% significance level; ** 5% significance level; *** 1% significance level.

Table A4 Placebo tests and behavioral channels

Dependent variable	Placebo test		Ruling out other behavioral channels		
	verbal	math	cooperation	impatience	completion time (min)
	(1)	(2)	(3)	(4)	(5)
Dep. Var. mean			*5.723*	*2.301*	*52.084*
temperature bins					
<12 °C	-0.264	0.299	-0.290	-0.172	3.292
	(0.457)	(0.237)	(0.195)	(0.364)	(4.570)
12-14 °C	0.196	0.086	-0.297*	0.252	1.896
	(0.428)	(0.211)	(0.176)	(0.354)	(4.025)
14-16 °C	0.507*	0.178	-0.056	-0.003	2.468
	(0.278)	(0.149)	(0.164)	(0.387)	(3.626)
16-18 °C	0.407*	0.135	-0.150	-0.434*	2.940
	(0.232)	(0.160)	(0.125)	(0.255)	(4.002)
18-20 °C	-0.152	0.090	0.009	-0.227	0.285
	(0.226)	(0.099)	(0.105)	(0.195)	(2.154)
20-22 °C	-0.102	0.132	-0.100	0.093	1.693
	(0.178)	(0.100)	(0.068)	(0.136)	(1.297)
22-24 °C					
24-26 °C	-0.098	0.102	-0.113**	0.171	1.282
	(0.143)	(0.081)	(0.055)	(0.132)	(0.985)
26-28 °C	-0.025	0.015	-0.090	0.094	1.828
	(0.181)	(0.089)	(0.090)	(0.152)	(1.318)
28-30 °C	0.184	0.119	-0.066	0.184	1.682
	(0.184)	(0.099)	(0.097)	(0.169)	(1.329)
30-32 °C	0.213	0.003	-0.129	0.161	2.165
	(0.261)	(0.121)	(0.116)	(0.204)	(2.259)
>32 °C	-0.029	-0.263	-0.284	-0.068	-3.807
	(0.321)	(0.160)	(0.181)	(0.298)	(2.559)
Observations	70,710	70,710	47,028	48,592	70,736
Adjusted R-squared	0.072	0.178	0.045	0.042	0.047

Note: Results in Column (3) are based on data from CFPS 2010 and 2014. Results in Column (4) is based on data from CFPS 2014 and 2018. The placebo test is conducted using temperature exposure the next day. The left-out temperature bin is 22-24 °C. The coefficients can be interpreted as effects of a day in the corresponding temperature bin on cognitive test scores relative to the reference temperature category. All the regressions include individual fixed effects, county fixed effects, year, month, day-of-week, and hour-of-day fixed effects. Demographic controls include gender, age with its square term and education level. Environmental controls include air pollution index (API), total precipitation, wind speed, sunshine duration, and relative humidity in quadratic forms. Robust standard errors, clustered at the county level, are presented in parentheses. * 10% significance level; ** 5% significance level; *** 1% significance level.

Table A5 Robustness checks

Dependent variable	Adding county specific linear time trends and calendar date fixed effects		Adding completion time		Removing API		Adding an indicator for working outdoors	
	verbal	math	verbal	math	verbal	math	verbal	math
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
temperature bins								
<12 °C	0.710	0.211	-0.239	0.119	-0.162	0.133	-0.040	0.256
	(0.565)	(0.306)	(0.470)	(0.229)	(0.474)	(0.226)	(0.476)	(0.225)
12-14 °C	0.653	0.178	0.610	0.378	0.691	0.391	0.734	0.363
	(0.591)	(0.358)	(0.482)	(0.274)	(0.485)	(0.273)	(0.475)	(0.274)
14-16 °C	-0.337	-0.314	-0.023	-0.060	0.042	-0.048	0.095	0.025
	(0.621)	(0.277)	(0.486)	(0.210)	(0.483)	(0.212)	(0.438)	(0.205)
16-18 °C	0.086	-0.048	0.136	-0.042	0.179	-0.032	0.216	-0.018
	(0.260)	(0.191)	(0.246)	(0.161)	(0.238)	(0.157)	(0.229)	(0.137)
18-20 °C	-0.079	-0.037	0.034	0.074	0.054	0.077	0.051	0.096
	(0.219)	(0.127)	(0.208)	(0.120)	(0.207)	(0.120)	(0.215)	(0.122)
20-22 °C	-0.147	-0.088	-0.069	-0.083	-0.063	-0.080	-0.047	-0.069
	(0.184)	(0.088)	(0.199)	(0.088)	(0.202)	(0.089)	(0.202)	(0.093)
22-24 °C								
24-26 °C	-0.085	-0.003	-0.102	0.009	-0.105	0.011	-0.155	-0.006
	(0.144)	(0.078)	(0.133)	(0.070)	(0.133)	(0.071)	(0.140)	(0.073)
26-28 °C	0.252	0.031	0.182	0.055	0.170	0.056	0.127	0.019
	(0.182)	(0.094)	(0.166)	(0.089)	(0.164)	(0.090)	(0.172)	(0.093)
28-30 °C	0.109	-0.073	0.130	-0.037	0.112	-0.036	0.087	-0.083
	(0.229)	(0.121)	(0.215)	(0.111)	(0.211)	(0.111)	(0.216)	(0.105)
30-32 °C	0.134	-0.096	0.161	-0.003	0.140	-0.003	0.119	-0.025
	(0.318)	(0.167)	(0.308)	(0.138)	(0.304)	(0.137)	(0.307)	(0.138)
>32 °C	0.337	-0.383*	0.224	-0.413**	0.183	-0.424**	0.154	-0.484**
	(0.409)	(0.213)	(0.440)	(0.202)	(0.433)	(0.201)	(0.434)	(0.197)
Observations	70,736	70,736	70,736	70,736	70,736	70,736	67,421	67,421
Adjusted R-squared	0.111	0.210	0.072	0.179	0.071	0.178	0.039	0.122

Note: The left-out temperature bin is 22-24 °C. The coefficients can be interpreted as effects of a day in the corresponding temperature bin on cognitive test scores relative to the reference temperature category. All the regressions include individual fixed effects, county fixed effects, year, month, day-of-week, and hour-of-day fixed effects. Demographic controls include age with its square term and education level. Environmental controls include air pollution index (API), total precipitation, wind speed, sunshine duration, and relative humidity in quadratic forms. Robust standard errors, clustered at the county level, are presented in parentheses. * 10% significance level; ** 5% significance level; *** 1% significance level.

Table A6 Robustness checks (continued)

Dependent variable	log form of test scores		Using subjects with time-invariant education levels		Using non-migrants only		Excluding ozone dominated days	
	verbal	math	verbal	math	verbal	math	verbal	math
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
temperature bins								
<12 °C	0.000	0.026	-0.044	0.018	-0.161	0.158	-0.294	0.124
	(0.052)	(0.042)	(0.530)	(0.237)	(0.488)	(0.240)	(0.493)	(0.261)
12-14 °C	0.057	0.010	1.030*	0.365	0.656	0.365	0.408	0.326
	(0.045)	(0.047)	(0.572)	(0.275)	(0.492)	(0.286)	(0.539)	(0.312)
14-16 °C	0.000	-0.012	0.303	0.037	0.039	-0.077	0.055	-0.064
	(0.055)	(0.045)	(0.471)	(0.205)	(0.517)	(0.225)	(0.564)	(0.231)
16-18 °C	-0.017	-0.032	0.450*	-0.096	0.111	-0.055	-0.002	-0.146
	(0.026)	(0.029)	(0.252)	(0.149)	(0.260)	(0.169)	(0.318)	(0.180)
18-20 °C	-0.008	0.011	0.161	0.063	0.047	0.087	0.062	0.128
	(0.025)	(0.024)	(0.226)	(0.132)	(0.214)	(0.123)	(0.262)	(0.133)
20-22 °C	-0.009	-0.022	-0.105	-0.089	-0.028	-0.073	-0.112	-0.101
	(0.026)	(0.019)	(0.214)	(0.107)	(0.206)	(0.095)	(0.235)	(0.099)
22-24 °C								
24-26 °C	-0.014	-0.000	-0.110	-0.033	-0.089	0.005	0.010	0.037
	(0.015)	(0.014)	(0.155)	(0.079)	(0.138)	(0.074)	(0.173)	(0.089)
26-28 °C	0.006	0.007	0.169	-0.003	0.186	0.063	0.245	0.074
	(0.019)	(0.018)	(0.190)	(0.094)	(0.174)	(0.091)	(0.215)	(0.115)
28-30 °C	-0.000	-0.009	0.123	-0.094	0.145	-0.027	0.026	-0.072
	(0.025)	(0.022)	(0.223)	(0.108)	(0.226)	(0.115)	(0.278)	(0.144)
30-32 °C	0.007	0.002	0.038	-0.009	0.176	0.018	0.230	0.145
	(0.035)	(0.030)	(0.307)	(0.138)	(0.319)	(0.141)	(0.454)	(0.187)
>32 °C	-0.013	-0.068**	0.356	-0.416*	0.156	-0.436**	-0.328	-0.546**
	(0.045)	(0.034)	(0.448)	(0.227)	(0.463)	(0.209)	(0.514)	(0.251)
Observations	70,728	70,735	57,705	57,705	67,322	67,322	54,664	54,664
Adjusted R-squared	0.039	0.048	0.037	0.124	0.069	0.172	0.072	0.159

Note: The left-out temperature bin is 22-24 °C. The coefficients can be interpreted as effects of a day in the corresponding temperature bin on cognitive test scores relative to the reference temperature category. All the regressions include individual fixed effects, county fixed effects, year, month, day-of-week, and hour-of-day fixed effects. Demographic controls include age with its square term and education level. Environmental controls include air pollution index (API), total precipitation, wind speed, sunshine duration, and relative humidity in quadratic forms. Robust standard errors, clustered at the county level, are presented in parentheses. * 10% significance level; ** 5% significance level; *** 1% significance level.

Table A7 Cumulative effects of heat waves on cognitive performance

	the number of consecutive days above 32 °C immediately before the survey		the largest number of consecutive days above 32 °C in the past 30 days	
	verbal test scores	math test scores	verbal test scores	math test scores
	(1)	(2)	(3)	(4)
the number of consecutive days above 32 °C before the survey	-0.032	-0.070		
	(0.082)	(0.049)		
the largest number of consecutive days above 32 °C in the past 30 days			-0.060	-0.069**
			(0.091)	(0.030)
demographic controls	Yes	Yes	Yes	Yes
environmental controls	Yes	Yes	Yes	Yes
individual fixed effects	Yes	Yes	Yes	Yes
county fixed effects	Yes	Yes	Yes	Yes
interview year, month, day-of-week, and hour-of-day fixed effects	Yes	Yes	Yes	Yes
Observations	70,738	70,738	70,738	70,738
Adjusted R-squared	0.071	0.178	0.071	0.178

Note: Demographic controls include age with its square term and education level. Environmental controls include air pollution index (API), total precipitation, wind speed, sunshine duration, and relative humidity in quadratic forms. Robust standard errors, clustered at the county level, are presented in parentheses. * 10% significance level; ** 5% significance level; *** 1% significance level.

Table A8 Mechanism test – transitory effect of temperatures on other outcomes

Dependent variable	daily sleeping hours	self-rated health status	K6 scores	life satisfaction	CES-D scores
	(1)	(2)	(3)	(4)	(5)
Dep. Var. mean	*7.929*	*2.628*	*2.973*	*3.798*	*5.461*
SD	*1.470*	*1.288*	*3.824*	*1.023*	*4.047*
temperature bins					
<12 °C	0.306**	-0.020	-0.250	0.011	-0.538
	(0.125)	(0.059)	(0.348)	(0.067)	(0.477)
12-14 °C	0.206*	-0.098	0.250	-0.024	-0.918
	(0.121)	(0.070)	(0.387)	(0.071)	(0.560)
14-16 °C	0.085	-0.045	0.017	0.105*	-0.207
	(0.094)	(0.062)	(0.341)	(0.057)	(0.416)
16-18 °C	0.095	0.006	0.195	0.054	-0.027
	(0.100)	(0.045)	(0.237)	(0.051)	(0.236)
18-20 °C	0.035	-0.002	0.160	0.031	-0.214
	(0.077)	(0.034)	(0.180)	(0.037)	(0.171)
20-22 °C	0.027	-0.046*	0.086	0.043	-0.033
	(0.045)	(0.027)	(0.136)	(0.035)	(0.202)
22-24 °C					
24-26 °C	0.032	-0.055**	-0.131	0.009	0.246
	(0.035)	(0.028)	(0.143)	(0.025)	(0.163)
26-28 °C	-0.024	-0.019	0.000	0.004	0.211
	(0.048)	(0.028)	(0.140)	(0.030)	(0.175)
28-30 °C	0.012	-0.032	-0.068	-0.037	0.157
	(0.055)	(0.032)	(0.157)	(0.032)	(0.193)
30-32 °C	0.026	-0.018	-0.350*	-0.046	0.347
	(0.060)	(0.036)	(0.208)	(0.031)	(0.211)
>32 °C	0.083	-0.107**	-0.103	-0.059	0.437*
	(0.087)	(0.054)	(0.337)	(0.050)	(0.251)
Observations	54,675	67,413	43,472	65,365	23,594
Adjusted *R*-squared	0.065	0.423	0.021	0.098	0.101

Note: Results in Column (5) are based on data from CFPS 2018. The left-out temperature bin is 22-24 °C. The coefficients can be interpreted as effects of a day in the corresponding temperature bin on cognitive test scores relative to the reference temperature category. All the regressions include individual fixed effects (except for Column 5), county fixed effects, year, month, day-of-week, and hour-of-day fixed effects. Demographic controls include gender, age with its square term and education level. Environmental controls include air pollution index (API), total precipitation, wind speed, sunshine duration, and relative humidity in quadratic forms. Robust standard errors, clustered at the county level, are presented in parentheses. * 10% significance level; ** 5% significance level; *** 1% significance level.

Table A9 Mechanism test – cumulative effect of temperatures in the past 30 days on other outcomes

Dependent variable	daily sleeping hours	self-rated health status	K6 scores	life satisfaction	CES-D scores
	(1)	(2)	(3)	(4)	(5)
Dep. Var. mean	*7.929*	*2.628*	*2.973*	*3.798*	*5.462*
SD	*1.470*	*1.288*	*3.824*	*1.023*	*4.047*
temperature bins					
<12 °C	0.009	0.002	0.007	0.002	-0.040
	(0.009)	(0.004)	(0.021)	(0.004)	(0.035)
12-14 °C	-0.014	0.007	0.016	0.005	0.122*
	(0.016)	(0.008)	(0.037)	(0.006)	(0.069)
14-16 °C	0.035**	-0.013*	0.074*	0.011	-0.119
	(0.016)	(0.007)	(0.039)	(0.007)	(0.082)
16-18 °C	-0.019	0.006	-0.024	-0.002	0.002
	(0.012)	(0.007)	(0.046)	(0.008)	(0.065)
18-20 °C	0.003	0.000	0.012	0.008	0.004
	(0.011)	(0.005)	(0.031)	(0.005)	(0.039)
20-22 °C	0.006	-0.007	0.011	0.006	-0.013
	(0.013)	(0.005)	(0.029)	(0.005)	(0.040)
22-24 °C					
24-26 °C	0.001	-0.001	0.005	0.002	-0.033
	(0.009)	(0.004)	(0.031)	(0.005)	(0.032)
26-28 °C	0.003	-0.000	-0.016	0.005	0.017
	(0.007)	(0.003)	(0.022)	(0.004)	(0.025)
28-30 °C	0.002	-0.004	0.013	0.004	-0.025
	(0.007)	(0.003)	(0.026)	(0.004)	(0.023)
30-32 °C	-0.002	0.002	-0.026	-0.004	0.012
	(0.008)	(0.004)	(0.029)	(0.004)	(0.025)
>32 °C	-0.010	-0.012*	-0.034	0.002	0.069
	(0.011)	(0.007)	(0.049)	(0.006)	(0.042)
Observations	54,677	67,415	43,472	65,367	23,596
Adjusted *R*-squared	0.066	0.423	0.022	0.099	0.101

Note: Results in Column (5) are based on data from CFPS 2018. The left-out temperature bin is 22-24 °C. The coefficients can be interpreted as effects of an additional day in the corresponding temperature bin on cognitive test scores relative to the reference temperature category. All the regressions include individual fixed effects (except for Column 5), county fixed effects, year, month, day-of-week, and hour-of-day fixed effects. Demographic controls include gender, age with its square term and education level. Environmental controls include air pollution index (API), total precipitation, wind speed, sunshine duration, and relative humidity in quadratic forms. Robust standard errors, clustered at the county level, are presented in parentheses. * 10% significance level; ** 5% significance level; *** 1% significance level.

Table A10 Mechanism test – cumulative effect of temperatures in the past year on other outcomes

Dependent variable	daily sleeping hours	self-rated health status	K6 scores	life satisfaction	CES-D scores
	(1)	(2)	(3)	(4)	(5)
Dep. Var. mean	*7.929*	*2.628*	*2.973*	*3.798*	*5.462*
SD	*1.470*	*1.288*	*3.824*	*1.023*	*4.047*
temperature bins					
<-4 °C	0.001	0.003*	-0.006	0.000	-0.010
	(0.003)	(0.001)	(0.007)	(0.001)	(0.032)
-4-0 °C	0.003	0.004***	0.014	-0.000	-0.053
	(0.002)	(0.001)	(0.009)	(0.002)	(0.038)
0-4 °C	-0.001	-0.001	0.005	0.002*	-0.038
	(0.003)	(0.002)	(0.008)	(0.001)	(0.042)
4-8 °C	-0.000	-0.000	-0.005	0.000	0.004
	(0.002)	(0.001)	(0.007)	(0.001)	(0.039)
8-12 °C	-0.002	-0.000	-0.005	-0.001	-0.024
	(0.003)	(0.001)	(0.007)	(0.001)	(0.048)
12-16 °C					
16-20 °C	-0.002	-0.001	-0.007	-0.000	-0.013
	(0.002)	(0.001)	(0.008)	(0.001)	(0.036)
20-24 °C	-0.002	-0.001	-0.003	0.000	-0.022
	(0.002)	(0.001)	(0.007)	(0.001)	(0.026)
24-28 °C	-0.003	0.001	-0.004	-0.000	-0.028
	(0.002)	(0.001)	(0.005)	(0.001)	(0.025)
>28 °C	-0.004*	-0.001	-0.004	-0.001	-0.015
	(0.002)	(0.001)	(0.007)	(0.001)	(0.027)
Observations	54,677	67,415	43,472	65,367	23,596
Adjusted R-squared	0.066	0.423	0.024	0.099	0.101

Note: Results in Column (5) are based on data from CFPS 2018. The left-out temperature bin is 12-16 °C. The coefficients can be interpreted as effects of an additional day in the corresponding temperature bin on cognitive test scores relative to the reference temperature category. All the regressions include individual fixed effects (except for Column 5), county fixed effects, year, month, day-of-week, and hour-of-day fixed effects. Demographic controls include gender, age with its square term and education level. Environmental controls include air pollution index (API), total precipitation, wind speed, sunshine duration, and relative humidity in quadratic forms. Robust standard errors, clustered at the county level, are presented in parentheses. * 10% significance level; ** 5% significance level; *** 1% significance level.