



Open Research Online

The Open University's repository of research publications and other research outputs

Is 8:30 a.m. Still Too Early to Start School? A 10:00 a.m. School Start Time Improves Health and Performance of Students Aged 13-16.

Journal Item

How to cite:

Kelley, Paul; Lockley, Steven W; Kelley, Jonathan and Evans, Mariah D R (2017). Is 8:30 a.m. Still Too Early to Start School? A 10:00 a.m. School Start Time Improves Health and Performance of Students Aged 13-16. *Frontiers in human neuroscience*, 11 p. 588.

For guidance on citations see [FAQs](#).

© [not recorded]

Version: Version of Record

Link(s) to article on publisher's website:

<http://dx.doi.org/doi:10.3389/fnhum.2017.00588>

Copyright and Moral Rights for the articles on this site are retained by the individual authors and/or other copyright owners. For more information on Open Research Online's data [policy](#) on reuse of materials please consult the policies page.

oro.open.ac.uk



Is 8:30 a.m. Still Too Early to Start School? A 10:00 a.m. School Start Time Improves Health and Performance of Students Aged 13–16

Paul Kelley^{1*}, Steven W. Lockley^{2,3}, Jonathan Kelley⁴ and Mariah D. R. Evans⁵

¹ Sleep, Circadian and Memory Neuroscience, The Open University, Milton Keynes, United Kingdom, ² Division of Sleep and Circadian Disorders, Brigham and Women's Hospital, Boston, MA, United States, ³ Division of Sleep Medicine, Harvard Medical School, Harvard University, Boston, MA, United States, ⁴ International Survey Center, Sociology, University of Nevada, Reno, NV, United States, ⁵ Sociology and Applied Statistics Program, University of Nevada, Reno, NV, United States

OPEN ACCESS

Edited by:

Peter Sörös,
University of Oldenburg, Germany

Reviewed by:

John F. Araujo,
Federal University of Rio Grande do
Norte, Brazil
Luigi De Gennaro,
Sapienza Università di Roma, Italy

*Correspondence:

Paul Kelley
drpaulkelley@gmail.com

Received: 29 August 2017

Accepted: 20 November 2017

Published: 08 December 2017

Citation:

Kelley P, Lockley SW, Kelley J and Evans MDR (2017) Is 8:30 a.m. Still Too Early to Start School? A 10:00 a.m. School Start Time Improves Health and Performance of Students Aged 13–16. *Front. Hum. Neurosci.* 11:588. doi: 10.3389/fnhum.2017.00588

While many studies have shown the benefits of later school starts, including better student attendance, higher test scores, and improved sleep duration, few have used starting times later than 9:00 a.m. Here we report on the implementation and impact of a 10 a.m. school start time for 13 to 16-year-old students. A 4-year observational study using a before-after-before (A-B-A) design was carried out in an English state-funded high school. School start times were changed from 8:50 a.m. in study year 0, to 10 a.m. in years 1–2, and then back to 8:50 a.m. in year 3. Measures of student health (absence due to illness) and academic performance (national examination results) were used for all students. Implementing a 10 a.m. start saw a decrease in student illness after 2 years of over 50% ($p < 0.0005$ and effect size: Cohen's $d = 1.07$), and reverting to an 8:50 a.m. start reversed this improvement, leading to an increase of 30% in student illness ($p < 0.0005$ and Cohen's $d = 0.47$). The 10:00 a.m. start was associated with a 12% increase in the value-added number of students making good academic progress (in standard national examinations) that was significant (<0.0005) and equivalent to 20% of the national benchmark. These results show that changing to a 10:00 a.m. high school start time can greatly reduce illness and improve academic performance. Implementing school start times later than 8:30 a.m., which may address the circadian delay in adolescents' sleep rhythms more effectively for evening chronotypes, appears to have few costs and substantial benefits.

Keywords: school start times, sleep, circadian, illness, academic performance, adolescence, circadian social science

INTRODUCTION

Despite the well-established natural shifts to later wake and sleep times that occur in adolescence, most schools retain early start times. Currently school starting times are not adjusted for the shift to later wake and sleep times that occur naturally in adolescence. This mismatch between adolescent biology and the conventional practice of starting school early leads to a systematic reduction in the amount of time available for sleep to teenagers and consequently chronic sleep deficiency. The Centers for Disease Control and Prevention and the American Academy of Pediatrics (2014) have

stated that early school starts are associated with increased health risks of obesity, depression, and drug use as well as poorer academic performance (Owens et al., 2014; Wheaton et al., 2015). Their recommendation that middle and high schools should open no earlier than 8:30 a.m. is now supported by the American Medical Association (2016). Almost all studies to date, while scheduling school starting times later than before, have retained a starting time at or earlier than 9:00 a.m. (Kirkby et al., 2011). One study in New Zealand did report improved sleep in 17 to 18-year-old students when the starting time was moved from 9:00 to 10:30 a.m., as compared to younger controls who remained at 9:00 a.m. (Borlase et al., 2013). A recent study of optimal times for cognitive performance for students aged 18–19 concluded that much later times were optimal, specifically after 11 a.m. or 12 noon (Evans et al., 2017).

Evidence for increased health risks associated with early school starting times is both substantial and demonstrated through a variety of research methodologies (Hansen et al., 2005; Millman, 2005; Basch et al., 2014; de Souza and Hidalgo, 2014). The underlying biological drivers are also well established. Adolescents need anywhere from 8 to 10 h of sleep per night for full health and academic performance depending on age and inter-individual differences, yet most get far less (Iglowstein et al., 2003; Foster et al., 2013; Hirshkowitz et al., 2015). Biological changes in the timing of the 24-h circadian clock during adolescence delay the onset of wake and sleep times, and this shift does not reverse until early adulthood (Roenneberg et al., 2004). Additionally, homeostatic regulation of pressure to sleep builds more slowly, taking a longer time to reach the critical threshold required to initiate sleep (Carskadon, 2011). Adolescent sleep restriction is clearly linked to early school starts as on non-school days adolescents have wake times two or more hours later (Roenneberg et al., 2007), a finding seemingly not substantially affected by cultural factors (Gradisar et al., 2011; Foster et al., 2013).

Later school starting times provide benefits to adolescent sleep, health and learning (Curcio et al., 2006; Carskadon, 2011; Lufi et al., 2011). There is substantial evidence that later starting times benefit academic outcomes even in the early stages of puberty and this positive impact continues into late adolescence (Carrell et al., 2011; Edwards, 2012; Meltzer et al., 2014). Later starting times are associated with improved sleep that continues into the years following implementation (Borlase et al., 2013; Wahlstrom et al., 2014). Later starting times also reduce the rate of student driving accidents (Danner and Phillips, 2008), and lower reports of depression (Kirkby et al., 2011).

The principle that school start times for adolescents should be later than currently the norm in American schools (about 8:00 a.m.) is now widely accepted. Research is now needed into synchronizing school starting times more closely with adolescent biology, taking into account the increasing impact of circadian rhythm changes in adolescence (Shekleton et al., 2013; Kelley et al., 2015). While the evidence for starting middle and high schools at 8:30 a.m. or later is positive, this recommendation is based on evidence limited to studies where school times are shifted to no later than 9:00 a.m., or often earlier, leaving unanswered the question of how late school starting times should

be. To explore the impact of a much later school start time, we examined the impact of a 10 a.m. starting time (representing a 1:10 h delay from an 8:50 a.m. starting time), on rates of illness of 13 to 16-year-old students and academic performance of 14 to 16-year-old students.

PARTICIPANTS AND METHODS

Determining the Start Time

In light of the then-existing evidence on start times, the lead investigator of the present study (PK) designed and directed a field experimental to implement and assess the efficacy of a much later school start time. The design is summarized in **Table 1**, columns 1, 2, and 4 (see section Results). The school was a mixed sex, state-funded school's for 13 to 18-year-old students. At the start of the study, the school student performance was considerably below the national mean. The school was in an urban area of 0.7 million, in a region of England where achievement was lower than national average.

The school had no additional funding, selective entry or other educational interventions during the duration of the study. The school operated within The Innovation Trust, a charitable trust to improve schools in partnership with leading scientists (see section Acknowledgments). Working with this group of experts, a later start time was considered with a focus on determining a starting time for adolescents aged 13–18. Initial testing of student academic performance at 10:00 and 14:00, voluntary chronotype surveys and body fat levels of students led to the conclusion there should be a core time period for all students between 11:00 and 15:00, with additional learning time before (for morning/average chronotypes) and after (for evening chronotypes) (Sussman, 2009). This appeared too radical a change and, combined with the then current National Health Service Trust judgement that schools could not legally hold any mental health data on students 16 or under including sleep and body fat data, led to limiting the age range of 13–16 and using only existing national measures of health and academic performance. Therefore analysis of the relevant sleep and circadian neuroscience research indicating 10:00–10:30 a.m. starting times would be appropriate at age 16 (Kelley et al., 2015). This process led to the final choice of 10:00 a.m. as the school start time, a 1 h 10 min delay compared to the traditional 8:50 a.m. start time then in use.

The 2-year process for changing start times was ethically approved by both The Innovation Trust, a charitable educational research trust linked to the school, and the school's Governing Body on the basis that students would not be subject to additional surveys, tests or measures other than a later starting time. A legally-required consultation with parents, students, teachers and others was conducted. This raised a number of concerns including transport to school. Unlike American schools, urban schools in England generally do not provide transport for all students, and the existing transport infrastructure made this a relatively minor issue. The school already supported parental work hours and transport by opening 1 h early and 1 h after school (this included sports and club opportunities). Therefore, with minor changes, all the times in the school were simply moved an hour and 10 min later. Although

TABLE 1 | Effect of experimental treatment (10 a.m. school start time) on absences due to illness.

Year and school start time	Study design [1] (Target school's GCSE course lasts 2 years)				Absences per student due to illness			Effect of 10 a.m. start time experiment (after adjusting for yearly changes in the national rate of absences)			
	Field Experiment (treatment = 10 a.m. start)	Natural experiment (drop 10 a.m. start)	Years subject to 10 a.m. start treatment	National baseline (mean)	Target school (mean)	Difference: Target school - national baseline (cols 6-5)	Target school		Partial effects		Total effect
							as percent of national baseline	Change in number of absences (from col 7)	Change in number of absences (from col 7)	Change, as a percent of national baseline	
[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]	[9]	[10]	[11]	[12]
Year 0 8:50 a.m. (baseline)	Field control group	-	No 10 a.m. treatment	11.5	15.4 (<i>sd</i> = 7.037) (<i>n</i> = 496)	3.9 <i>t</i> = 12.34 [2] <i>p</i> < 0.0005	34 % higher	-	-	-	-
Year 1 10:00 a.m.	Field treatment group #1	Nat. exp. control group #1	1 year without 10 a.m. treatment and 1 year with it	10.7	11.3 (<i>sd</i> = 7.037) (<i>n</i> = 538)	0.6 <i>t</i> = 1.88 <i>p</i> = 0.061, ns 95% CI = (10.67, 11.93)	6 % higher	Year 1 to Year 0: -3.3 # down <i>t</i> = -9.36, <i>p</i> < 0.0005 95% CI = (-4.960, -3.240) (Cohen's <i>d</i> = 0.58)	-31 % down	-	-
Year 2 10:00 a.m.	Field treatment group #2	Nat. exp. control group #2	2 years with 10 a.m. treatment	8.8	7.9 (<i>sd</i> = 7.037) (<i>n</i> = 541)	-0.9 <i>t</i> = -2.97 <i>p</i> = 0.003 95% CI = (7.31, 8.50)	-10 % lower	Year 2 to Year 1: -1.5 # down <i>t</i> = -7.94, <i>p</i> < 0.0005 95% CI = (-4.24, -2.56) (Cohen's <i>d</i> = 0.48)	-17 % down	Year 2 to Year 0: -4.8 # down <i>t</i> = -17.14, <i>p</i> < 0.0005 95% CI = (-8.36, -6.64) (Cohen's <i>d</i> = 1.07)	-55 % down
Year 3 8:50 a.m.	-	Nat. exp. treatment group	1 year with 10 a.m. treatment and 1 year without it	9.3	11.2 (<i>sd</i> = 7.037) (<i>n</i> = 474)	1.9 <i>t</i> = 5.88 <i>p</i> < 0.0005 95% CI = (10.57, 11.84)	20 % higher	-	-	Year 3 to Year 2: 2.8 # up <i>t</i> = 7.45, <i>p</i> < 0.0005 95% CI = (2.43, 4.17) (Cohen's <i>d</i> = 0.47)	30 % up

Key results (columns 11 and 12) highlighted in red boldface.

[1] The field experiment was designed and directed by one of the authors (PK). The natural experiment was the result of a change in control of the target school to local government that reinstated the early start in that was used in all local schools.

[2] Raw data are not publicly available and standard deviations are not published. We use here an estimate of the standard deviation based on the grouped percentage distribution of absences. Assuming uniform distributions within intervals gives a conservative (wider) IQR. Then assuming absences are approximately normally distributed, we take IQR/1.35 = estimated standard deviation. Here that is 7.037.

parents supported the changes, there was resistance from the external local educational administrators, a phenomenon noted in other studies (Wahlstrom et al., 2014). The concerns raised during the consultation, while often valid, were addressed as possible, but considered secondary to the *raison d'être* for schools—to help healthy children realize their full academic potential.

After approval of the change, a 6-week trial run with a 10:00 a.m. start was conducted in 2010 after all annual examinations were completed and students taking examinations had left, but before the start of the new academic year. The trial moved all related scheduled events to match the 10:00 a.m. start including transport to and from school, family arrangements, and school clubs and activities. There were also some additional and unexpected benefits reported: The travel times avoided rush hour and were considered safer; and some staff could take their children to their primary school in the morning.

After 2 years with the 10 a.m. start time, a natural experiment was created by a change in local education administrators that reinstated the early start as used in all local schools. Consequently, the school starting time for 13 to 16-year olds was changed back to 8:50 a.m. due to these policy changes, not part of the study design; however we took advantage of this natural experiment.

Intervention and Data Collection

This observational study used a before-after-before (A-B-A) design. School start times changed as follows over 4 years: Year 0 had an 8:50 a.m. start; Years 1 and 2 had a 10:00 a.m. start; and Year 3 had an 8:50 a.m. start. Over these four academic years (September 2010–August 2014) national data and school data on illness and performance of 2,049 students aged 13–16 years were collected.

English education data are collected nationally for each school and these data for all schools are published by the UK Office of National Statistics (2016). Two variables were used to assess the impact of the time change: Student absence due to illness, and overall student academic performance. English schools are required to distinguish absence (not in school for any reason) from absence due to illness (which accounts for ~60% of all absences, varying by year). Absence due to illness is recorded for each student for every morning and afternoon session, and so the measure of illness utilized here was the number of school sessions lost through illness per student per year. This measure gives a more precise measure of illness than raw absences rates (Department for Education, 2016).

At that time, the government's key measure of student academic performance was the General Certificate of Secondary Education (GCSE) examinations taken at the end of compulsory schooling at age 16. Good academic progress at 16 was defined as achieving five or more GCSE grades of C or above in English, Mathematics and at least three other subjects. The measure of school performance was the percentage of students achieving good academic progress.

In addition, a national system of value-added analysis of individual students and school potential achievement in GCSE grades and good academic progress was produced by the Fisher

Family Trust¹ for all students and schools in England. As the prior ability of student cohorts can vary year-to-year, so can the effectiveness of schools, and thus a second analysis of performance based on these value-added predictions was undertaken (Visscher and Coe, 2003; Koedel et al., 2015). The predicted results were based on the students' previous achievement in national tests. Schools exceeding predicted percentage good academic progress were deemed to have added value, given as a percentage good academic progress above the predicted outcomes (with negative results for schools below predictions). For example, if the students achieved an average of 4% higher good academic progress, it would have a value-added score of +4%. These two national measures of actual and predicted school performance were used in our analyses.

Analyses

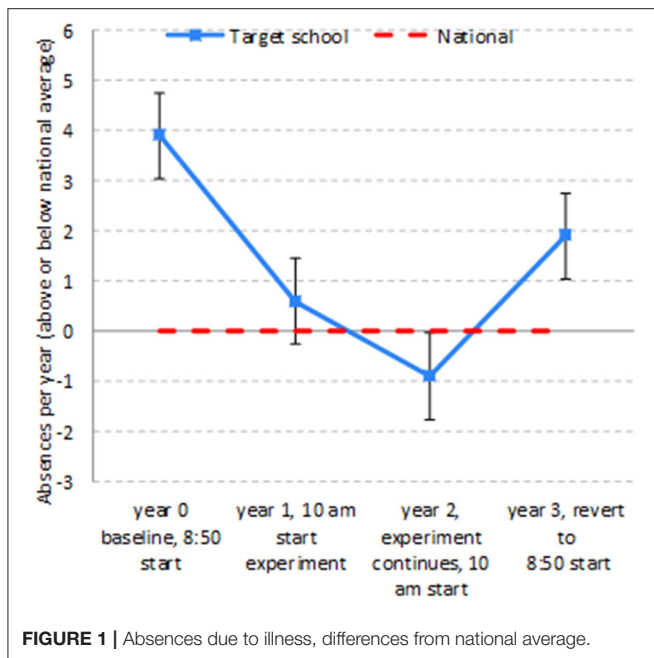
National data were used to analyse illness and performance. National data were used where available to estimate standard deviations. In addition, a data set of 2,880 similar schools was constructed where "similar" was defined as state-funded school with 13 to 16-year-old students in cohorts >100, and reporting examination data for the period 2010–2014. The data sets used are school-level data published by the UK's Office of National Statistics, and other sources as indicated (UK Office of National Statistics, 2016). The achievement data have limitations due to policy and publication practices. Policy limitations include a decision to remove schools becoming academies (similar to Charter schools) from national data, changes in examinations, and in the data reported during the 4 years.

Analysis of illness and performance data during this period led to reporting both raw scores and scores adjusted to account for national variation. For national trends in illness, performance, *T*-tests were used to assess significance, and Cohen's *d* and *h* for effect size, taking into account Hattie's critique of educational measures to determine impact of interventions (Hattie, 2008). Hattie's critique of educational research based on synthesizing over 800 meta-analyses on raising achievement raised questions about educational data analyses and significance. Specifically, he argued that more sophisticated statistical techniques, large numbers of subjects usually possible within education research, and other factors meant effect sizes should be >0.5 in individual studies. Hattie lists 138 effect sizes found in meta-analyses of the significant educational interventions of the last 20 years, showing the highest 20 interventions had effect sizes ranging from 0.61 to 1.44. Hattie's concerns about educational research and effect sizes are shared by others (Snow, 2015; Churches et al., 2017). For illustrative purposes, the cost/benefits of educational reforms in England and New York to increase educational performance was also analyzed.

RESULTS

Rates of absence due to illness of students aged 13–16 were lower with the 10 a.m. start time (see **Table 1** and **Figure 1**).

¹Fisher Family Trust (2016). Available online at: www.fft.org.uk (Accessed December 2016).



In Year 0, for the baseline cohort of students aged 13–16 which had experienced 8:50 start times for both years of their course in the target school, the mean number of absences due to illness was 15.4, compared to 11.5 nationally (Year 0 “Field Control Group” in **Table 1**, compare column 6 for the target school to column 5 for the national benchmark). That is a difference of 3.9 absences which is statistically significant at $p < 0.0005$ (column 7). Moreover, it is large: The rate in the target school was 34% higher than in the national benchmark (column 8).

The situation changed sharply for Year 1 (“Field Treatment group #1” in **Table 1**), the first year with the experimental 10 a.m. start treatment. The mean number of absences in the target school dropped from 15.4 to 11.3, a drop of 4.1 absences (column 6, compare Year 0 and Year 1). The gap between the target school absence rate and the national benchmark shrank by 3.3 absences to a mere 0.6 absences (Year 1, column 7). This shrinkage amount to a 31% decline in the gap between the target school and the national benchmark (Year 1, column 10), a decline that is statistically significant at $p < 0.0005$. Cohen’s d for the change is 0.58, making the effect size in the medium-to-large range. This reduced absence rate left the target school a mere 0.6 absences above the national benchmark, a difference so small that it is not statistically significant ($p > 0.05$, column 7).

The crux of the field experiment is Year 2, when the 10 a.m. start was operating for its second full year (the target school’s GCSE courses last 2 years with examinations throughout). Absences dropped further in the target school, from a mean of 11.3 to a mean of 7.9 (column 6), a drop of 3.4 absences. This brought the target school’s absence rate significantly below the national benchmark (column 7). This shift from slightly above the national benchmark in Year 1 to significantly below the national benchmark in Year 2 has a Cohen’s d of 0.48, a medium effect size.

The most crucial comparison is between Year 2 when the 10 a.m. start was operating for its second full year and Year 0 in which students had been on the 8:50 start for at least 2 years. The mean number of absences fell by fully 7.5 (Column 6, Year 2 vs. Year 0). The gap between the target school rate and the national benchmark shrank by 55%, amounting to a shrinkage of 4.8 absences. This change is statistically significant at $p < 0.0005$. The Cohen’s d for the change is 1.07, a large effect size.

Thus, the field experiment demonstrates in multiple ways that the experimental treatment is associated with a decline in absences that is large and statistically significant not only relative to the control group but also relative to the national benchmark. The pattern is evident in **Figure 1**.

Next comes the natural experiment when the start time in the target school returned to 8:50 in Year 3 due to a change in policy which shifted control of the target school to local education authorities who imposed a uniform early start time on all comparable schools. The mean number of absences rebounded to 11.2, well above the level in the previous year (column 6). This was 20% above the national benchmark, a clear contrast with the previous year’s 10% below the benchmark (column 8). Adjusting for yearly changes in the national rate of absences, the gap between the target school and the nation inflated to 2.8 absences a year, a 30% increase (columns 11 and 12). This change is statistically significant at $p < 0.0005$ and has a Cohen’s d of 0.47, a medium effect size.

All in all, for absences the field experiment demonstrates that the change to a 10 a.m. start reduced absences due to illness by over 50 percent compared to national rates, a large and statistically significant decline. The natural experiment demonstrates that reverting to the 8:50 start induced, already in its first year, a medium sized, statistically significant increase in absences. These data are consistent with a dose-dependent response to a 10:00 start.

In the target school’s Year 0 (“Field Control Group” in **Table 2**) the baseline cohort of students which had experienced 8:50 start times for both years of their course, the percentage who successfully completed the GCSE examinations was 34, compared to 56.2 nationally (columns 5 and 6). That is a vast difference of 22 percentage points, statistically significant at $p < 0.0005$ (column 7). This difference amounts to 40% of the national baseline (column 8), marking the target school as having a very low success rate. In value added terms, the actual percentage of successful students was 5 percentage points lower than the FFT estimate for the school based on the cohort’s past performance (column 9), showing a negative value-added value. All in all, in Year 0 in the target school, the performance picture was grim.

The situation changed sharply for Year 1, the first year with the experimental 10 a.m. start treatment. The actual percent of academically successful students shot up by 19 percentage points to 53 (column 6, compare Year 0 and Year 1). The gap between the target school and in national benchmark in Year 1 is so small that it is not statistically significant (Year 1, column 7). Between Year 0 and Year 1, the gap between the target school success rate and the national benchmark shrank from 22 to 5 (Year 1, column 7), a shrinkage of 17 percentage points that is statistically

TABLE 2 | Effect of experimental treatment (10 a.m. school start time) on academic success.

Year and school start time	Study design [1] (Target school's GCSE course lasts 2 years)			Actual percent academically successful (C or above in 5 or more GCSE subjects)			Value added by school (FFT estimate [3])			Actual change (raw, unadjusted)			Change in value added	
	Field experiment (add 10 a.m. start) [2]	Natural experiment (drop 10 a.m. start) [3]	In target School	Difference from national (cols 6-5) [7]	Difference as % of national (cols 7/5) [8]	Predicted (based on cohort's past performance) [9]	Actual-predicted (cols 6-9)	Number	Change as percent of national (% of col 5) [12]	Number	Change as percent of national (% of col 5)	Number	Change as percent of national (% of col 5)	
Year 0 8:50 a.m. (baseline)	Field control group	No 10 a.m. treatment	34 (n = 169)	-22 worse p < 0.0005	-40 % worse	39	-5 worse	-	-	-	-	-	-	
Year 1 10:00 a.m.	Field treatment group #1	Nat. exp. control group #1	53 (n = 166)	-5 worse p = 0.182, ns	-9 % worse	49	4 better	19 # up	33 % up	9 # up	15 % up	9 # up	15 % up	
Year 2 10:00 a.m.	Field treatment group #2	Nat. exp. control group #2	52 (n = 164)	-7 worse p = 0.081, ns	-11 % worse	45	7 better	-1 # down	-2 % down	12 # up	20 % up	12 # up	20 % up	
Year 3 8:50 a.m.	-	Nat. exp. treatment group	51 (n = 179)	-9 worse p = 0.014	-15 % worse	48	3 better	-1 # down	-2 % down	-4 # down	-7 % down	-4 # down	-7 % down	

Success is defined, following government convention, as achieving a grade of C or above in five or more subjects in the national General Certificate of Secondary Education (GCSE) examinations. Key results (columns 13 and 14) highlighted in red boldface.

[1] The field experiment was designed and directed by one of the authors (PK). The natural experiment was the result of a change in control of the target school to local government that reinstated the early start in that was used in all local schools.

[2] Raw data are not publicly available and standard deviations are not published. We use here an estimate of the standard deviation based on the grouped percentage distribution of absences. Assuming uniform distributions within intervals gives a conservative (wider) IQR. Then assuming absences are approximately normally distributed, we take IQR/1.35 = estimated standard deviation. Here that is 7.037.

[3] Fisher Family Trust Value-added estimates, widely used in British education (<https://fft.org.uk>).

significant at $p < 0.0005$. The value added by the school has also risen: The actual value added by the school is 4 percentage points more than the FFT prediction, up 9 percentage points from Year 0. This amounts to 15% of the national baseline and is statistically significant at $p < 0.0005$.

The heart of the field experiment is the comparison of Year 2 vs. Year 0. Here again there is extensive evidence for the positive impact of the 10 a.m. start on student success. Student success in Year 2 is 52% compared to 34% for Year 0 (column 6), a gain which is significant at $p = 0.001$ and of medium size, with a Cohen's H of 0.37. Between Year 0 and Year 2, the gap between the target school success rate and the national benchmark shrank from 22 to 7 (Year 2, column 7). The actual student success in Year 2 exceeds the FFT prediction by 7 percentage points (column 10). Indeed, there is a 12-percentage point gain in the value added by the school, a gain that is statistically significant at $p < 0.0005$ (column 13) and amounts to 20% of the national benchmark.

Thus, the field experiment demonstrates in multiple ways that the experimental treatment is associated with a statistically significant and substantial gain in performance not only relative to the control group but also relative to the national benchmark.

Then, in the natural experiment in Year 3, the school starting time reverted from 10 a.m. to 8:50 a.m. The student success rate fell slightly to 51 (Year 3, column 6) in comparison to a national benchmark of 60, giving the target school a deficit of 9 percentage points which is statistically significant at $p = 0.014$ (column 7). The value added by the school has shrunk by 4 percentage points which is 7% of the national baseline (column 14).

Thus, even though these students had the first year of their 2-year course with the 10 a.m. start, reverting to the 8:50 a.m. start for their second year is associated with a reduction in success relative to the national baseline and with a reduction in the value added by the school.

The whole pattern for the combined results of the field experiment and the natural experiment is clear in **Figure 2**.

In Year 0, student success in the target school was actually 5 percentage points below what would have been expected in light of the students' prior performance (FFT). Then the field experiment began. The introduction of the 10 a.m. start halfway through the Year 1 cohort's course lifted the school's value added to 9 percentage points above the expected success percent. There was a further small rise to a value added of 12 percentage points in Year 2 for the cohort which had experienced the 10 a.m. start for both years of their course. The field experiment then ended. The natural experiment then reverted the start time to 8:50, so students in the Year 3 cohort had the first year of their course with the 10 a.m. start and the second year of their course with the 8:50 start. Value added was still positive for this cohort, but had fallen by 4 percentage points compared to the previous cohort which had 2 years of 10 a.m. starts.

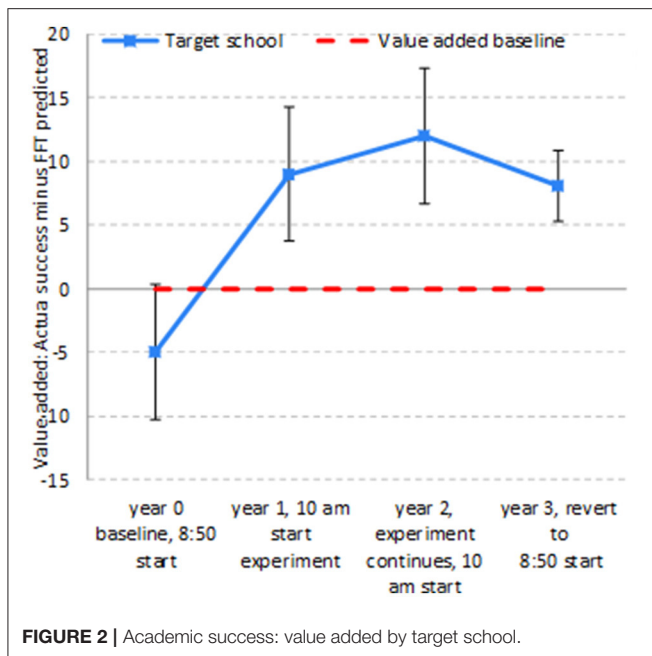
DISCUSSION

Based on this study, moving school start times later so that they are better aligned with adolescent sleep and chronotype

patterns is practical and beneficial. Following a change to a 10:00 a.m. school start time, rates of absence due to illness in students aged 13–16 reduced, and academic performance of students aged 14–16 significantly improved. When the school start time was returned to 8:50 a.m., these benefits were abolished. These findings suggest that the general policy recommendation to start high schools no earlier than 08:30 a.m., while helpful, should not be taken as justification to exclude consideration of much later starting times. As the 8:50 a.m. would already be considered appropriate using the 8:30 a.m. recommendation, this study shows further improvements can be made when starting times are much later, in this case 10:00 a.m.

The absence due to illness rate data are consistent with a direct benefit of later school starting times on student health. The differentiation in recording absence and absence due to illness in England offers a more precise measure of illness and a large national data set. Following initiation of the 10:00 a.m. start, illnesses decreased in the second year by over 50%. With the return of an 8:50 a.m. these benefits were reversed, with illness increasing by of 30%. Academic performance of students aged 14–16 also improved with a 10:00 a.m. start. Having 2 years of later start times was particularly beneficial; showing a 12-percentage point gain in the value added that is statistically significant at $p < 0.0005$ and amounts to 20% of the national benchmark. There is no reason to believe that these outcomes reflected motivation changes. The pupils were studying for their final national exams, which determine their eligibility for continued study and ultimately college or university, or their competitiveness in the jobs market. These were not study-related tests but the real, once-in-lifetime exams that have a major influence on the children's futures and thus would be equally motivated. While there are several possible explanations for the poor performance of students in similar low socioeconomic status areas, the possible impact of sleep restriction linked to early school starts has rarely been considered.

One of the objections raised to changing school start times is that the change is impractical and cannot overcome other barriers, such as bus timetables or sports program scheduling. While the English legal framework makes changing to much later start times a formal process any school can undertake (and therefore more practical than in some other countries), moving all of this school's schedules later produced no practical difficulties. By choosing a 10 a.m. start time, the school aimed to provide benefits for the largest proportion of children possible, given the inter-individual range in phase delays that children experience means some substantially delayed children would still be waking too early in their circadian cycle (Iglowstein et al., 2003; Kelley et al., 2015). Even later start times might address this issue, but 10 a.m. was considered a reasonable compromise between maximizing biological benefits for most children while remaining practical. The national US recommendation that middle and high schools should start after 8:30 a.m. is a clearly justified positive step, although the evidence in this study suggests a much later start of 10 a.m.—even when replacing a school start later than the 8:30 a.m. recommendation—brings additional benefits. A recent study of



university students aged 18–19 found that later starting times (after 11 a.m. or 12 noon) were optimal and much later than an 8:30 a.m. start recommended for High School students of 17–18. The study also found that using a fixed time for all students would disadvantage one or more chronotypes, and evening chronotypes in particular if starts were early (Kelley et al., 2015; Evans et al., 2017). Thus, even with the benefits accrued using a 10 a.m. wake time for all students, this approach does not address the wider variation of wake times in different chronotypes.

The limitations in this study include small sample sizes in some measures, a focus mainly on illness, and an inability to measure students' sleep duration. Although other studies have shown sleep improvements following less substantial interventions, some have tracked sleep improvements over 4 years and found that improvements persist (Borlase et al., 2013; Wahlstrom et al., 2014). The English location and measures of academic performance are difficult to contextualize with previous research using U.S. state-wide testing and graduation data. The nationally available data on illness on a school-by-school basis, focus on students in mid-adolescence and the starting time of 10:00 a.m. are, we believe, unique in this research field. A larger, more detailed study is needed of school starts after 9:00 a.m., including those that are later than 10:00 a.m. There is preliminary evidence that such changes can have benefits, particularly for older students (Carrell et al., 2011).

The school-level improvements in performance in this study, if more widely replicated, should be examined from economic and educational perspectives. For example, expenditure to reduce the English attainment gap between rich and poor students reached more than one billion pounds between 2010 and 2015, and yet had relatively little impact compared to the gains made following the 10 a.m. school start time change in

this study. The change to smaller schools in key U.S. cities such as New York City have been thoroughly researched after additional expenditure of over a billion dollars to improve performance (Schneider et al., 2007; Stiefel et al., 2015). Other educational policy changes including creating new kinds of school such as Charter, Academy, and Science, Technology, Engineering and Math (STEM) schools, increasing the duration of school times, curriculum and test changes, or No Child Left Behind also have very high expenditure but minimal impact and little scientific rigor in evaluating their effects. In contrast, changing to later start times is a very cost-effective intervention to raise educational standards with substantial scientific backing (Jacob and Rockoff, 2011; Snow, 2015; Hafner et al., 2017).

The key finding in this study is reduced absences due to illness by over 50% compared to national rates (Table 1, columns 12 and 13). This huge change is both practically important and highly significant statistically ($p < 0.0005$ and Cohen's d for the change is 1.07, a large effect size). The broader impact of later starts on specific aspects of adolescent health, such as sleep duration and quality, mental health, and social development were not assessed, although other studies have shown potential impacts (de Souza and Hidalgo, 2014; Meltzer et al., 2014; Minges and Redeker, 2016). Additional research into much later starts should measure both actual sleep patterns and optimal performance times for individual students. The most important area for further research may be the impact of later starts on areas of social behavior development and mental health. For example, the daily sleep loss of two or more hours per day imposed by early school starts (which cannot be recovered with 10 or more extra hours of sleep at weekends), may put those with a genetic predisposition to a mental illness at greater risk given that direct links between sleep of <6 h and gene expression have been established (Möller-Levet et al., 2013). Sleep deprivation is also associated with adolescents being less perceptive readers of human emotions (van der Helm et al., 2010; Guadagni et al., 2014), during a period of greater sensitivity to sociocultural signals (Blakemore and Mills, 2014) and related brain developments in adolescence. These interrelated factors of significant sleep deprivation, genetic predisposition, the high prevalence of the onset of mental illness during adolescence for a range of disorders (Schmitt et al., 2014) and less a perceptive reading of sociocultural signals, may impact on levels of mental illness and emotional disorders in adolescence (Wulff et al., 2010, 2012).

Using a research-based approach to determine a school starting time for 13 to 16-year-old students led to the implementation of a 10 a.m. school starting time. This later starting time had a substantial benefit for rates of illness and academic performance. A research-based approach to school starting times is clearly replicable in different contexts, cultures, and countries. More importantly, a post-9:00 a.m. school start strategy is one with few costs and many potential benefits which start to accrue, quite literally, overnight. Application of sleep research in this way demonstrates the powerful impact on society and individuals of making evidence-based policy changes.

ETHICS STATEMENT

The Innovation Trust and School Governing Body approved the change in starting times in compliance with English Educational Law. Parents and students were consulted during the change as legally required.

AUTHOR CONTRIBUTIONS

PK, SL, and JK: study design; JK and ME: data analysis; PK, SL, JK, and ME: writing and critical review of manuscript.

REFERENCES

- American Medical Association (2016). *News Release*. Available online at: <https://www.ama-assn.org/ama-supports-delayed-school-start-times-improve-adolescent-wellness> (Accessed November 2017).
- American Academy of Pediatrics (2014). Adolescent Sleep Working Group: school start times for adolescents. *Pediatrics* 134, 642–649. doi: 10.1542/peds.2014-1697
- Basch, C. E., Basch, C. H., Ruggles, K. V., and Rajan, S. (2014). Prevalence of sleep duration on an average school night among 4 nationally representative successive samples of American high school students, 2007–2013. *Prev. Chronic Dis.* 11:E216. doi: 10.5888/pcd11.140383
- Blakemore, S. J., and Mills, K. L. (2014). Is adolescence a sensitive period for sociocultural processing?. *Annu. Rev. Psychol.* 65, 187–207. doi: 10.1146/annurev-psych-010213-115202
- Borlase, B. J., Gander, P. H., and Gibson, R. H. (2013). Effects of school start times and technology use on teenagers' sleep: 1999–2008. *Sleep Biol. Rhythms* 11, 46–54. doi: 10.1111/sbr.12003
- Carrell, S. E., Maghikian, T., and West, J. E. (2011). A's from Zzzz's? The causal effect of school start time on academic achievement of adolescents. *Am. Econ. J.* 3, 62–81. doi: 10.1257/pol.3.3.62
- Carskadon, M. A. (2011). Sleep's effects on cognition and learning in adolescence. *Prog. Brain Res.* 190, 137–143. doi: 10.1016/B978-0-444-53817-8.00008-6
- Churches, R., Dommert, E., and Devonshire, I. (2017). *Neuroscience for Teachers: Applying Research Evidence from Brain Science*. London: Crown House Publishing.
- Curcio, G., Ferrara, M., and De Gennaro, L. (2006). Sleep loss, learning capacity and academic performance. *Sleep Med. Rev.* 10, 323–337. doi: 10.1016/j.smrv.2005.11.001
- Danner, F., and Phillips, B. (2008). Adolescent sleep, school start times, and teen motor vehicle crashes. *J. Clin. Sleep Med.* 4, 533–535.
- de Souza, C. M., and Hidalgo, M. P. (2014). Midpoint of sleep on school days is associated with depression among adolescents. *Chronobiol. Int.* 31, 199–205. doi: 10.3109/07420528.2013.838575
- Department for Education (2016). *Pupil Absence in Schools in England: Autumn 2015 and Spring 2016*. London, UK: Department of Education, SFR 51/2016.
- Edwards, F. (2012). Early to rise? The effect of daily start times on academic performance. *Econ. Educ. Rev.* 31, 970–983. doi: 10.1016/j.econedurev.2012.07.006
- Evans, M. D. R., Kelley, P., and Kelley, J. (2017). Identifying the best times for cognitive functioning using new methods: matching university times to undergraduate chronotypes. *Front. Hum. Neurosci.* 11:188. doi: 10.3389/fnhum.2017.00188
- Foster, R. G., Peirson, S. N., Wulff, K., Winnebeck, E., Vetter, C., and Roenneberg, T. (2013). Sleep and circadian rhythm disruption in social jetlag and mental illness. *Prog. Mol. Biol. Transl. Sci.* 119, 325–346. doi: 10.1016/B978-0-12-396971-2.00011-7
- Gradisar, M., Gardner, G., and Dohnt, H. (2011). Recent worldwide sleep patterns and problems during adolescence: a review and meta-analysis of

FUNDING

This work was funded by Science + Technology in Learning (Award number 1/2017).

ACKNOWLEDGMENTS

The authors would like to thank Professors David Reynolds, Russell Foster, and Daniel Muijs for their contribution to the thinking behind changing the start time based on sleep research, and David Reynolds contributed in its implementation as Chair of The Innovation Trust.

- age, region, and sleep. *Sleep Med.* 12, 110–118. doi: 10.1016/j.sleep.2010.11.008
- Guadagni, V., Burles, F., Ferrara, M., and Iaria, G. (2014). The effects of sleep deprivation on emotional empathy. *J. Sleep Res.* 23, 657–663. doi: 10.1111/jsr.12192
- Hafner, M., Stepanek, M., and Troxel, W. M. (2017). *Later School Start Times in the U. S.: and An Economic Analysis*. Santa Monica, CA: RAND Corporation. Available online at: https://www.rand.org/pubs/research_reports/RR2109.html
- Hansen, M., Janssen, I., Schiff, A., Zee, P. C., and Dubocovich, M. L. (2005). The impact of school daily schedule on adolescent sleep. *Pediatrics* 115, 1555–1561. doi: 10.1542/peds.2004-1649
- Hattie, J. (2008). *Visible learning: A synthesis of Over 800 Meta-Analyses Relating to Achievement*. London: Routledge.
- Hirshkowitz, M., Whiton, K., Albert, S. M., Alessi, C., Bruni, O., DonCarlos, L., et al. (2015). National Sleep Foundation's updated sleep duration recommendations: Final report. *Sleep Health* 1, 233–243. doi: 10.1016/j.sleh.2015.10.004
- Iglowstein, I., Jenni, O. G., Molinari, L., and Largo, R. H. (2003). Sleep duration from infancy to adolescence: reference values and generational trends. *Pediatrics* 111, 302–307.
- Jacob, B. A., and Rockoff, J. E. (2011). *Organizing Schools to Improve Student Achievement: Start Times, Grade Configurations, and Teacher Assignments*. Washington, DC: Brookings/Hamilton Project Discussion Paper 2011–08.
- Kelley, P., Lockley, S. W., Foster, R. G., and Kelley, J. (2015). Synchronizing education to adolescent biology: 'let teens sleep, start school later'. *Learn. Media Technol.* 40, 210–226. doi: 10.1080/17439884.2014.942666
- Kirkyby, M., Maggi, S., and D'Angiulli, A. (2011). School start times and the sleep-wake cycle of adolescents: a review and critical evaluation of available evidence. *Educ. Res.* 40, 56–61. doi: 10.3102/0013189X11402323
- Koedel, C., Mihaly, K., and Rockoff, J. E. (2015). Value-added modeling: a review. *Econ. Educ. Rev.* 47, 180–195. doi: 10.1016/j.econedurev.2015.01.006
- Lufi, D., Tzischinsky, O., and Hadar, S. (2011). Delaying school starting time by one hour: some effects on attention levels in adolescents. *J. Clin. Sleep Med.* 7, 137–143.
- Meltzer, L. J., Shaheed, K., and Ambler, D. (2014). Start later, sleep later: school start times and adolescent sleep in homeschool versus public/private school students. *Behav. Sleep Med.* 14:140. doi: 10.1080/15402002.2014.963584
- Millman, R. P. (2005). Excessive sleepiness in adolescents and young adults: causes, consequences, and treatment strategies. *Pediatrics* 115, 1774–1786. doi: 10.1542/peds.2005-0772
- Minges, K. E., and Redeker, N. S. (2016). Delayed school start times and adolescent sleep: a systematic review of the experimental evidence. *Sleep Med. Rev.* 28, 82–91. doi: 10.1016/j.smrv.2015.06.002
- Möller-Levet, C. S., Archer, S. N., Bucca, G., Laing, E. E., Slak, A., Kabiljo, R., et al. (2013). Effects of insufficient sleep on circadian rhythmicity and expression amplitude of the human blood transcriptome. *Proc. Natl. Soc. Sci. U.S.A.* 110, E1132–E1141. doi: 10.1073/pnas.1217154110
- Owens, J., Drobni, D., Baylor, A., and Lewin, D. (2014). School start time change: an in-depth examination of school districts in the United States. *Mind Brain Educ.* 8, 182–213. doi: 10.1111/mbe.12059

- Roenneberg, T., Kuehnle, T., Juda, M., Kantermann, T., Allebrandt, K., Gordijn, M., et al. (2007). Epidemiology of the human circadian clock. *Sleep Med. Rev.* 11, 429–438. doi: 10.1016/j.smrv.2007.07.005
- Roenneberg, T., Kuehnle, T., Pramstaller, P. P., Ricken, J., Havel, M., Guth, A., et al. (2004). A marker for the end of adolescence. *Curr. Biol.* 14, R1038–R1039. doi: 10.1016/j.cub.2004.11.039
- Schmitt, J. E., Neale, M. C., Fassassi, B., Perez, J., Lenroot, R. K., Wells, E. M., et al. (2014). Dynamic role of genetics on cortical patterning during childhood and adolescence. *Proc. Natl. Acad. Sci. U.S.A.* 111, 6774–6779. doi: 10.1073/pnas.1311630111
- Schneider, B., Wyse, A. E., and Keesler, V. (2007). “Is small really better? Testing some assumptions about high school size,” in *Brooking Papers on Education Evaluation Policy*, eds T. Loveless and F. M. Hess (Washington, DC: Brookings Institution), 15–47.
- Shekleton, J. A., Rajaratnam, S. M., Gooley, J. J., Van Reen, E., Czeisler, C. A., and Lockley, S. W. (2013). Improved neurobehavioral performance during the wake maintenance zone. *J. Clin. Sleep Med.* 9:353. doi: 10.5664/jcsm.2588
- Snow, C. E. (2015). 2014 Wallace foundation distinguished lecture: rigor and realism: doing educational science in the real world. *Educ. Res.* 44, 460–466. doi: 10.3102/0013189X15619166
- Stiefel, L., Schwartz, A. E., and Wiswall, M. (2015). Does small high school reform lift urban districts? Evidence from New York City. *Educ. Res.* 44, 161–172. doi: 10.3102/0013189X15579187
- Sussman, A. (2009). *The British Broadcasting Corporation (BBC), Horizon Series 2008-2009, Episode 11, The Secret Life of Your Bodyclock.*
- UK Office of National Statistics (2016). *GCSE Data Releases 2010–2015.* Available online at: <https://www.ons.gov.uk/> (Accessed November 2017).
- van der Helm, E., Gujar, N., and Walker, M. P. (2010). Sleep deprivation impairs the accurate recognition of human emotions. *Sleep* 33, 335–342. doi: 10.1093/sleep/33.3.335
- Visscher, A. J., and Coe, R. (2003). School performance feedback systems: conceptualisation, analysis, and reflection. *Sch. Effect. Sch. Improv.* 14, 321–349. doi: 10.1076/sesi.14.3.321.15842
- Wahlstrom, K., Dretzke, B., Gordon, M., Peterson, K., Edwards, K., and Gdula, J. (2014). *Examining the Impact of Later School Start Times on the Health and Academic Performance of High School Students: A Multi-Site Study.* Carei; St. Paul, MN: University of Minnesota.
- Wheaton, A. G., Ferro, G. A., and Croft, J. B. (2015). School start times for middle school and high school students- United States, 2011–12 School Year. *MMWR Morb. Mortal. Wkly. Rep.* 64, 809–813. doi: 10.15585/mmwr.mm6430a1
- Wulff, K., Dijk, D. J., Middleton, B., Foster, R. G., and Joyce, E. M. (2012). Sleep and circadian rhythm disruption in schizophrenia. *Brit. J. Psychiatry* 200, 308–316. doi: 10.1192/bjp.bp.111.096321
- Wulff, K., Gatti, S., Wettstein, J. G., and Foster, R. G. (2010). Sleep and circadian rhythm disruption in psychiatric and neurodegenerative disease. *Nat. Rev. Neurosci.* 11, 589–599. doi: 10.1038/nrn2868

Conflict of Interest Statement: SL has had a number of commercial interests in the last 12 months (2016–17). None are directly related to the research or topic reported in this paper but, in the interests of full disclosure, are outlined below. SL has received consulting fees from the Atlanta Falcons, Atlanta Hawks, BHP Billiton and Slingshot Insights; has current consulting contracts with Akili Interactive; Consumer Sleep Solutions; Delos Living LLC; Environmental Light Sciences LLC; Headwaters Inc.; Hints Performance AG; Light Cognitive; Mental Workout; OpTerra Energy Services Inc.; Pegasus Capital Advisors LP; PlanLED; and Wyle Integrated Science and Engineering; has received unrestricted equipment gifts from Biological Illuminations LLC, Bionetics Corporation and F. Lux Software LLC; royalties from Oxford University Press; and has served as a paid expert in legal proceedings related to light, sleep and health. He holds a patent through Harvard University and Brigham and Women’s Hospital for “Systems and methods for determining and/or controlling sleep quality.” He is a Program Leader for the CRC for Alertness, Safety and Productivity, Australia.

The other authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2017 Kelley, Lockley, Kelley and Evans. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.