Dec 2019 (Revised Mar 2020) No.450

Searching for Answers: The Impact of Student Access to Wikipedia

Laura Derksen, Catherine Michaud Leclerc and Pedro CL Souza

WORKING PAPER SERIES

This paper was revised in March 2020. It was originally circulated in December 2019.

Centre for Competitive Advantage in the Global Economy

Department of Economics

Searching for Answers: The Impact of Student Access to Wikipedia

Laura Derksen, Catherine Michaud Leclerc and Pedro CL Souza*

March 23, 2020

Abstract

Can schools use the internet to promote reading and learning? We provided Wikipedia access to randomly-selected students in Malawian boarding secondary schools. Students used Wikipedia broadly and intensively, and found it trustworthy, including for information about news and safe sex. We find a 0.10σ impact on English exam scores, and a higher impact among low achievers (0.20σ). Students used Wikipedia to study Biology, and exam scores increased for low achievers (0.14σ). Our results show that a source of engaging and accessible online material can encourage reading and affect educational outcomes.

JEL Codes: I21, I28, O15

*Derksen: University of Toronto (email: laura.derksen@utoronto.ca); Michaud Leclerc: University of Toronto (email: c.michaudleclerc@mail.utoronto.ca); Souza: University of Warwick (email: pedro.souza@warwick.ac.uk). We thank Natalie Bau, Jonas Hjort, Clément Imbert, Jason Kerwin, Anita McGahan, Uros Petronijevic, Steve Pischke, Roland Rathelot, seminar participants at Columbia, Notre Dame, PUC-Rio, UMN, SFU, the University of Toronto, AFE 2019 and CEA 2019 for excellent comments. Fatema Bhayani, Tarun Chopra, Jessica Gallant, Leandro Pongeluppe and Tanvi Shetty provided outstanding research assistance. We thank Abdallah Chilungo and the field team for their dedication. Errors and omissions remain our own. This project was funded by a SSHRC Insight Development grant, and pre-registered in the American Economic Association RCT Registry, number AEARCTR-0003824.

1 Introduction

In the developing world, school books are often in short supply, yet programs that simply provide reading material often have no impact on literacy or academic performance.[1] If reading material is not at the right level or does not cater to student interests, students are unlikely to read it or learn from it. Encouraging students to read usually requires additional incentives, or teacher training and engagement.[2] In order to be compelling, useful and accessible on its own, reading material must satisfy the demands of heterogeneous students, and be relevant across contexts.

As the internet expands worldwide, information technology offers a potential solution. The internet hosts reading material on almost every topic, at every level of difficulty. Young people in particular are enthusiastic internet users; in Africa, young people aged 15 to 24 use the internet at twice the rate of the general population.[3] Yet, internet in schools presents challenges of its own. While information on the internet is plentiful, it varies in its accuracy, trustworthiness and complexity (MacMillan and MacKenzie, 2012; Allcott and Gentzkow, 2017; Lazer et al., 2018). Moreover, students often prefer games, videos and social media to learning. In fact there is evidence that full internet access does not improve academic performance.[4]

In this paper, we show that the internet has a place in schools, and can be introduced in a way that promotes reading and learning. We provide students with an online experience restricted to Wikipedia, a vast yet accessible

[1] See for example Glewwe, Kremer and Moulin (2009), Borkum, He and Linden (2012), Sabarwal, Evans and Marshak (2014), and Knauer et al. (2020).

[2] Successful reading interventions that involve teachers include Machin and McNally (2008), Abeberese, Kumler and Linden (2014), Lucas et al. (2014), and Kerwin and Thornton (2018). He, Linden and MacLeod (2008) show that involving teachers improves the efficacy of an English learning intervention compared to adult supervision.

[3] Source: International Telecommunications Unit `https://www.itu.int/en/ITU-D/Statistics/Pages/stat/`, accessed on May 13, 2019.

[4] See Goolsbee and Guryan (2006), Vigdor, Ladd and Martinez (2014), Faber, Sanchis-Guarner and Weinhardt (2015), and Malamud et al. (2019).

source of accurate reading material.[5,6] This preserves one of the most exciting aspects of the internet: detailed, up-to-date information on almost any topic. Restricted access to online information might compel students to spend time reading, while avoiding other online distractions that do not involve reading. Wikipedia is also simple to use and understand, even for students of heterogeneous ability. Many articles have versions written in both standard and simple English. Students can easily search for information they need and click to learn more about concepts they do not understand, including concepts related to their studies.

We provide Malawian secondary school students with access to Wikipedia to answer three research questions. First, how do students use this new online resource? Do they find Wikipedia engaging, accessible, useful, and trustworthy? Second, Wikipedia gives students access to reading material on a vast range of topics. Are students compelled to spend time reading, and how does this affect English language ability? Third, what is the impact on academic performance in Biology, an important subject for which study materials are crucial? Biology is the most popular subject, and is important for career aspirations, as many secondary students in Malawi go on to a career in health care.[7]

We conducted a randomized experiment in government boarding schools in Malawi, a country with rapidly improving internet infrastructure, but where students have limited internet experience, no internet access, and few study resources. This setting allows us to isolate both treatment and control students from the broader internet. Students were allowed to use Wikipedia inside a

[5]Rather than restricting to a single information source, existing work measured the impact of full scale internet access on education (Bulman and Fairlie, 2016; Yanguas, 2018; Malamud, 2019), political and economic behavior (Bailard, 2012; Miner, 2015; Campante, Durante and Sobbrio, 2018; Chen and Yang, 2019) and development (Galperin and Viecens, 2017; Hjort and Poulsen, 2019). Randomized experiments specifically involving Wikipedia focused primarily on the decision to contribute to a public good (Chen et al., 2018; Hinnosaar, 2019).

[6]There is evidence that Wikipedia is mostly accurate, though incomplete. See Giles (2005), Rosenzweig (2006), Heilman et al. (2011), and Mesgari et al. (2015).

[7]We pre-registered final (term 3) English and Biology scores as our two primary outcomes (AEA RCT Registry number AEARCTR-0003824). English and Biology are core courses and are most often named as a favorite subject at baseline, and these subjects have the highest rate of exam completion. English is an official language of Malawi, and most courses are taught in English.

digital library. It was open evenings and weekends during one school year, and access was restricted to treated students. This design limits potential spillovers on English language skills and Biology exam scores. Students did not have any other internet access during term time.

Students found the online material engaging, as evidenced by their frequent and broad use of Wikipedia. They spent, on average, one hour and twenty minutes per week online. Rather than relying on aggregate usage statistics, we observe individual browsing histories, which allows us to characterize demand for specific topics at the level of an individual. Each student browsed, on average, more than 800 different pages across a range of topics.

Students came to use and trust Wikipedia, particularly for topics which are important, prone to misinformation and often absent from school books, such as world news and safe sex. We find spikes in activity in the week surrounding world news events that occurred during the experiment. We also show that students with access to Wikipedia are able to find news information that control group students cannot. Young people are generally curious about sex, and accurate information is vital (Dupas, 2011; Derksen and van Oosterhout, 2018; Kerwin, 2018). We find that students spent 5 percent of their browsing time on topics related to sex and sexuality. Students sought information on both news and sex and sexuality independently, without prompts or incentives.[8] At endline, treatment students were more likely than control students to report a preference for Wikipedia over other internet sources for both online news and information on safe sex, and were more likely to view it as easier to use and more trustworthy. Indeed, a majority of treatment students preferred Wikipedia along these dimensions.

Students used Wikipedia intensively for general interest reading, and we find a positive impact on English final exam scores. We find a significant improvement on average (0.10σ) and for low achievers in particular (0.20σ). We do not find any impact on high achievers. Students in the treatment group spent more than one hour per week reading articles in English, primarily on

[8]In fact, Chen and Yang (2019) show that even when provided with an internet VPN, university students in China do not search for international news unless incentivized. Our results suggest that interest in world news may be different outside of a censored regime.

topics that were not directly related to the school syllabus. This should not be viewed as a harmful distraction, as we can rule out even small negative effects across most subjects. Instead, we find a positive treatment effect on English exam scores.

By linking search terms to the school syllabus, we show that students find Wikipedia to be a useful study resource, especially for Biology. In other contexts, survey data suggests that students see value in Wikipedia as a study tool (Lim, 2009; Head and Eisenberg, 2010). Here we observe student browsing choices directly. We did not incentivize or pressure students to use Wikipedia for school, yet the average student did spend 22 percent of their time on pages related to the school syllabus. They spent more than twice as much time on Biology-related pages as on any other school subject.

This translates to an improvement in study time productivity and Biology exam scores for low achievers. We find a positive but insignificant impact on Biology exam scores (0.06σ), and a significant impact for low achievers (0.14σ). We again find no impact on high achievers. Low achievers did not spend more time on syllabus-related pages than high achievers, and neither low nor high achievers changed their total study time in response to the intervention. This implies an increase in study time productivity for low achievers. Indeed, we find that most treatment students preferred Wikipedia to their Biology textbooks, and were able to find academic information that their control group peers could not. Students perceived Wikipedia to be easy to use and understand; this would explain larger gains for students who were struggling at baseline. We do not, however, find any treatment effect on student education or career goals, which suggests that the effect on Biology exam scores is driven by study inputs and not by a change in aspirations.

This paper contributes to a small but emerging literature on education interventions in secondary school. Most education interventions to date target primary or middle school students,[9] and learning gaps in secondary school merit attention. While secondary school attendance is rising, completion rates

[9]See Banerjee et al. (2013) for a review of studies which focus on post-primary school students. Barrera-Osorio and Linden (2009) study an ICT intervention which did include secondary schools.

are low in Malawi and across sub-Saharan Africa.[10] Yet, returns to secondary school are high (Ozier, 2015). Secondary school is a necessary step towards postsecondary education, and a career in policy, education or health care. Finally, the effect of providing study material to secondary schools is likely to be different from the effect observed in primary schools, due to the advanced subject matter, and the fact that students are not illiterate. In this paper we show that reading material can in fact be useful to secondary students with a base level of literacy, for an advanced subject such as Biology.

We also contribute to the literature on reading interventions, and an expanding literature on the use of information technology and computer-aided learning in the classroom, especially programs that "teach at the right level." The fact that Wikipedia alone helps to close the gap between low and high achievers is remarkable, because interventions that simply provide reading material or internet access are usually ineffective.[11] On the other hand, computer-aided learning programs have shown promise (Barrow, Markman and Rouse, 2009; Carrillo, Onofa and Ponce, 2011; Mo et al., 2013; Beg et al., 2019), particularly for low achievers (Banerjee et al., 2007; Linden, 2008; Muralidharan, Singh and Ganimian, 2019). Muralidharan, Singh and Ganimian (2019) highlight a potential mechanism: even in heterogenous classrooms, a computer program can adapt to a student's ability. However, such programs are context specific and often involve teachers and administrators. Wikipedia offers an alternative to this tailored approach which is still appropriate for students of heterogeneous ability. It provides students with an ever expanding set of information, relevant across many contexts. It allows each student to

[10] In Malawi, 26 percent of women and 36 percent of men have at least some secondary education, however, less than half of those who start go on to graduate, see Malawi DHS 2015-16 (National Statistical Office/Malawi and ICF, 2017). According to Barro and Lee (2013), in 2010, 27 percent of individuals in sub-Saharan Africa aged 15 and over had completed some secondary education.

[11] For reading interventions, see Glewwe, Kremer and Moulin (2009), Borkum, He and Linden (2012), Sabarwal, Evans and Marshak (2014), and Knauer et al. (2020). Evidence suggesting that computers, software and internet access have a negative or no effect on school performance includes Malamud and Pop-Eleches (2011), Fairlie and Robinson (2013), Beuermann et al. (2015), and Cristia et al. (2017). See Bulman and Fairlie (2016) for a more detailed review of the literature. Furthermore, see Glewwe and Muralidharan (2016) for a review in the context of developing countries.

search for the specific information they need, written in accessible language. Indeed, Wikipedia does appear to have an impact primarily among students with lower baseline ability.

Wikipedia appears to be a useful substitute for English books and Biology textbooks, and is an accessible, cost-effective and up-to-date alternative for schools operating in low resource settings. Books are expensive to ship, necessarily limited in scope, and become out of date. Internet-enabled tablets and phones are available locally, and internet infrastructure is in place. We estimate that our intervention, as implemented, costs $4 USD per student per month. Internet and technology costs are decreasing over time, and if implemented in entire schools, the intervention might benefit from additional economies of scale. This is clearly more cost-effective than programs that provide reading material to primary schools, with no impact. It is also more cost-effective than many computer-aided learning programs. It is, however, less cost-effective than some of the most impactful primary school interventions, especially those that improve the quality of instruction. It is difficult to compare our intervention to other potential impacts in secondary school, as evidence is limited (Banerjee et al., 2013), but we might expect smaller returns in secondary school due to higher baseline ability levels.

The paper proceeds as follows. Section 2 describes the setting, the experimental design, the intervention, and our data sources. In Section 3, we explore student use of Wikipedia and the digital library. In Section 4, we investigate outcomes related to student beliefs, abilities and preferences in the context of Wikipedia and other online information sources. Section 5 presents our results on student academic performance. We conclude in Section 6 by discussing mechanisms, policy implications, and future research.

2 The Intervention: Wikipedia in Schools

2.1 Wikipedia

Wikipedia is an online encyclopedia, providing up-to-date reading material on a wide range of topics. It is the largest and most visited reference site on

the internet. It is a source of collaborative, accurate, open source information.[12] Content is created through open collaboration, and its accuracy on scientific topics is comparable to an offline encyclopedia (Giles, 2005). However, Wikipedia is frequently updated, and offers far more informational content than an offline encyclopedia, in terms of breadth, depth, and relevance.

Wikipedia is a high quality and accessible resource for students. Information is easy to find and understand. Articles exist in English and Simple English[13] (among many other languages), and Wiktionary serves as a companion dictionary. Wikipedia has a page for every topic on the typical secondary school syllabus, and often provides more detail than a textbook. For example, the English page for photosynthesis (a topic from secondary school Biology) has over 7000 words and several diagrams, and students can easily click links to similarly detailed pages on related concepts. There is also a Wikipedia page for photosynthesis in Simple English, with less detail, but with simple explanations, such as "Photosynthesis is the process by which plants and other things make food."

2.2 Setting and Sample

Malawi is a country in southern Africa with a GDP of less than $400 USD per capita, yet internet infrastructure is present throughout the country.[14] In 2006, 93 percent of the Malawian population lived in an area with access to a mobile network.[15] This surpasses the network coverage in neighboring Zambia and Mozambique (both at around 40 percent), and is comparable to the much richer South Africa (see Table A1).

[12]Source: Wikipedia, https://en.wikipedia.org/wiki/Wikipedia, accessed on May 23rd 2019. Wikipedia is free and owned by Wikimedia, a non-profit organization with no advertising.

[13]Simple English is a language defined by Wikipedia, which uses simpler words and shorter sentences than English Wikipedia. As of 2019, Simple English Wikipedia has more than 150,000 pages.

[14]According to the World Bank, GDP per capita in 2017 was $339 USD. This is well below the Sub-Saharan Africa and world average of $1,575 and $10,749, respectively. Current USD values.

[15]See Buys et al. (2009). 2G networks are largely accessible in rural areas, and 3G and 4G networks are available in towns and cities. See Batzilis et al. (2010) for a detailed description and analysis of the mobile network in Malawi.

Though internet infrastructure exists, access to the internet is unaffordable for most Malawians. 54 percent of Malawian households have a mobile phone (DHS, 2015-16),[16] but most of these phones do not have internet capabilities. Moreover, 1GB of internet costs the average Malawian 18 percent of their monthly income (see Table A1). This income share is larger than in Mozambique or Zambia, where incomes are higher. It is therefore not surprising that internet use in Malawi, while on the rise, has lagged behind other countries. In 2007, less than 1 percent of Malawians had regular internet access (see Table A1). In 2015, this rose to approximately 12 percent (DHS, 2015-16).

Malawi is on the verge of internet adoption, yet Malawian schools have limited resources and no internet access, making this a unique and appropriate setting for our study. The presence of internet infrastructure makes internet in schools feasible. Yet, most of the population, including youth, have limited internet experience. School libraries are small and contain textbooks in limited quantities. Most students do not have access to online resources. Mobile phones are usually prohibited. While some schools do have computer labs, they are typically offline.

At the same time, secondary school is challenging and completion is rare. Only 10 percent of women and 17 percent of men complete secondary school (DHS, 2015-16). Courses are taught in English, and require adequate language skills. The courses are difficult, and study materials are likely to be important.[17] In the fourth and final year, students take a national examination which determines university admission. Among those who sit their final exams, more than one third fail.[18]

Our experiment took place in four government boarding schools which serve students of mixed socioeconomic status. Each school has approximately five hundred students spread over four forms (grade levels). Government boarding schools are common in Malawi and across sub-Saharan Africa. They

[16] See National Statistical Office/Malawi and ICF (2017) for Malawi DHS, 2015-16.

[17] The core subjects are English, Biology, Chichewa (the local language) and Mathematics. Other subjects including Chemistry, Geography, History, Life Skills, Physics, and Social Studies are offered depending on the school, form (grade level) and interests.

[18] The 2018 pass rate for the Malawi Secondary Certificate of Education (MSCE) was 63 percent (https://maneb.edu.mw).

are more academically competitive than government day schools and most private schools (de Hoop, 2010). However, even in these schools, many students do struggle academically. In particular, one quarter of students had an English exam score below 50/100 in the year before the intervention. While government boarding schools attract good students, fees are not exorbitant.[19] Indeed, according to our baseline survey, many students at our sample schools are of lower socioeconomic status: 42 percent do not have electricity at home, and 45 percent do not have running water. One third of students have at least one parent who did not complete primary school.

Boarding schools provide a controlled environment; students have no access to the internet outside of our intervention, allowing us to cleanly limit internet use to Wikipedia. At the time of the intervention, the school grounds had consistent 3G or 4G network coverage. However, students were not allowed to access the internet or use phones, even outside of class time, and being caught with a phone at school was grounds for suspension. Students sleep in dormitories, and are not permitted to leave the school grounds. In particular, they do not go home during the term, so those who do have home internet access cannot use it.[20]

2.3 Experimental Design

In each boarding school, we set up a digital library where students could access Wikipedia outside of class time. The digital library was open most of one school year: from November 2017 to June 2018. It was open for four hours after school and eight hours on Saturday and Sunday. Each digital library was equipped with 12 internet-enabled Android devices. The devices were shared among 69 to 82 students in each school. We used password-protected software to restrict the devices to Wikipedia and Wiktionary.[21] We put links to English Wikipedia, Simple English Wikipedia and Wiktionary on the main login page.

[19] Admission is based on a national primary school exam. The school fees in our schools range from 75 to 165 USD per term, with many students on bursaries or scholarships.

[20] Students are sent home for two to four weeks between terms.

[21] We used the software *Kioware* to prevent students from accessing other webpages or applications. Students did not manage to exit the software or access other applications on the devices.

Inside the digital library, students could browse Wikipedia privately and anonymously. The digital library was supervised by our research staff, referred to as digital librarians. To log into a device, each student used a personal, unique and anonymous username and password.[22] The librarian did not monitor the content browsed by students. Students used the devices on their own (not in pairs or groups), and were not permitted to leave the digital library with a device. Students were allowed to take notes, and many did.

In October, 2017, we introduced the project to students, conducted a baseline survey, and collected baseline exam scores.[23] Our team of eight enumerators surveyed every student in Forms 2, 3 and 4. In total, we interviewed 1,508 students to collect information on their background, past internet use, time use, career and life aspirations, interests, and social networks.

After completing the baseline survey, we randomly assigned students to a treatment group or to a larger control group. The randomization assigned one fifth of students, a total of 301, to the treatment group. The remaining 1,207 students formed the control group. A sparse treatment ratio was chosen to limit spillovers, and jealousy, between students.[24] This also limits the potential for teachers to adapt their lesson plans in tandem, as most of their students do not have access to Wikipedia. A subset of students in the control group (299 students out of 1,207) was randomly assigned to a supplementary survey sample. This subsample would be surveyed more extensively for the construction of some secondary outcomes.

We randomized at the student level, and stratified on four key variables: school, form, exam scores and internet experience.[25] The bin for exam scores is defined as above or below the median score (within the school and form). We used the average of English and Biology exam scores. These are our two

[22] No one, including the research team, would be able to link a specific student to their browsing history.

[23] We introduced the project to students and teachers at each school, one form at a time, and all received the same information. See Supplementary Materials in the appendix for a detailed description of the classroom introduction.

[24] We also hoped this might reduce feelings of unfairness or disappointment, as a large majority of students found themselves in the control group. At endline, 79 percent of control students and 91 percent of treatment students felt the program was fair.

[25] We used a computer to randomize using the Stata command *randtreat*, seeded with the date of the randomization (2910).

primary outcomes; we have data for both English and Biology scores for 95 percent of students at baseline. We constructed a separate bin for students with missing exam score data. Internet experience is defined as whether the student has ever used the internet. There are 51 stratification bins. Panel A of Table 1 shows that our randomization is balanced across baseline variables (Table A2 in the appendix shows balance across stratification variables).

After the randomization took place, we publicly announced the names of the students in the treatment group, and held a mandatory induction session in the digital library.[26] During the induction, the students obtained an anonymous username which would be linked to their browsing history. The first letter of the username identifies coarsened student characteristics. Students with similar characteristics attended the same induction, and drew their usernames from the same envelope. This made it clear that browsing data obtained by the researchers would not be linked to a particular student.

Treatment students were invited to visit the digital library during opening hours, and sign in with the digital librarian to use a device within the digital library. If all devices were in use, they would join the waitlist or come back later. If there were students waiting, usage was restricted to approximately 30 minutes. Only students in the treatment group used the digital library, and the librarians used student photos to verify identities.[27] This restriction limits the scope for any spillovers to the control group that would rely on direct access to devices, Wikipedia or the internet. Teachers did not have access to the devices.

[26]The digital librarians explained the digital library and its rules. They also showed students how to access Wikipedia, and allowed the students to practice for fifteen minutes. Students were told that breaking the rules would result in suspension or removal of access. See Supplementary Materials in the appendix for a detailed description of the induction and digital library rules.

[27]Every week, a field team leader would visit each digital library to spot check the identities of the students and verify that no student in the control group was given access to the digital library. We also conducted spot checks, comparing student signatures to the baseline survey. We did not encounter a case where a control student gained access to the digital library.

3 How Students Used Wikipedia

In this section, we describe in detail how students in the treatment group used Wikipedia. Our browsing data is rich and granular, which allows us to provide a detailed analysis of browsing behavior, beyond a description of basic usage statistics. We explore how students use a new online information source, what types of information they value, and the tradeoff they face between general interests and academic subjects. Browsing behavior gives us a window into student interests and demand for information, which we will explore further in Section 4 using survey data. Understanding browsing behavior will also be key to interpreting results on academic performance in Section 5.

3.1 Browsing Data

Browsing data was recorded by software on our Wikipedia devices, and contains the complete sequence of pages visited by a particular student (linked to an anonymous username), a timestamp, and the time spent on each page. Although the browsing data does not identify any individual student, each username is linked to coarsened student characteristics.

Most students made frequent use of the digital library, and every student in the treatment group visited at least once. The average student visited the digital library on 33 days during the school year and each visit lasted 52 minutes.[28] This is approximately one hour and twenty minutes per week for each student, or 29 hours over the course of the year. Each student visited an average of 878 unique pages, and spent about two and a half minutes per page. 99.9 percent of pages visited were in English, and nearly 7 percent were in Simple English.

In Panel A of Figure 1 we present the distribution of browsing hours across students. The distribution is skewed to the right. While the average student spends 29 hours in the digital library, some students spent more than 150 hours browsing Wikipedia, over more than 100 visits. The time spent in the

[28]The digital library was open for 20-22 weeks, from November 2017 to June 2018, excluding Christmas and Easter vacations. We consider any browsing time within the same day to constitute one visit. Figure A1 shows browsing over time.

digital library is similarly distributed across low and high achievers (Panel B of Figure 1). Here, we define a low achiever to be a student whose average exam score (English and Biology) at baseline is below the median. This suggests that the intervention was accessible even to students with weaker language skills.

3.2 Topic Classifications

We use the Wikipedia category tree to classify pages according to topics in order to shed light on student interests and search behavior.[29] Panel A in Figure 2 presents the 24 most common Wikipedia Browsing topics according to time spent. The typical student spread their browsing time across several different topics (see Panel A of Figure A2 for detail on within-student variation in topics). The most popular topic is "People", with an average of four hours per student. This topic includes politicians, musicians, athletes, and other individuals of interest. Many popular topics including "Life", "Academic disciplines", "Arts", and "Nature" overlap with school subjects. We will identify school-related pages using a narrow classification in Section 3.5.

3.3 News and World Events

In this section, we ask whether students use Wikipedia to learn about the news. Indeed, the popularity of "People" pages may indicate interest in individuals at the center of a news story. Other popular news sources, such as social media and online news sites, are often biased and sometimes inaccurate (Chung, Nam and Stefanone, 2012). By comparison, news articles on Wikipedia are often impartial and accurate (Lih, 2004). If provided with this type of fact-based resource, will young people use it to read about world events?

We examine student browsing in the time leading up to or immediately following the event. Students might learn about news events from Wikipedia itself (as Wikipedia's main page has a section on news), from teachers, or

[29] See the Appendix for additional detail on the classification procedure.

during term breaks. We use Wikipedia's comprehensive list of 54 major world events that happened after the start of the intervention and prior to the start of the endline surveys (November 2nd, 2017 to May 9th, 2018).[30]

We observe a clear spike in student browsing activity during the week the event occurred (Panel A of Figure 3). The average student spent 1.7 minutes browsing these news stories, aggregated over 54 events. While few students read about any particular event, most students searched for at least one. This greatly underestimates total interest in the news, as most news events, and particularly news stories from Africa and Malawi, are not included among Wikipedia's top 54 stories. The spike in browsing emerges for both African and non-African events (Panels B and C of Figure 3). Students spent 15 times longer on news events taking place in Africa (Panel B of Figure 3).

3.4 Sex and Sexuality

Sex and sexuality are important topics for young people, and while teenagers are often curious, the information they obtain is not always accurate. Misinformation has serious consequences. It can lead to unwanted pregnancy, inappropriate behavior, and HIV infection. Wikipedia contains detailed, accurate, and up-to-date information on human reproduction, sexuality and sexual health.

We find that the average student spent 1.5 hours, or approximately five percent of their time on pages related to sex and sexuality, as determined by the Wikipedia category for "Human reproduction", which includes subcategories for "Human Sexuality" and "Sexual Health". They visited 48 different pages on this topic. Comparing this to Figure 2, we see that sex and sexuality would rank among the top ten general interest topics. The page for "Sexual Intercourse" is the most popular page within this topic and across all Wikipedia pages.

[30]Source: https://en.wikipedia.org/wiki/2017 and https://en.wikipedia.org/wiki/2018. For a particular event, we generate a list of narrow keywords, and include Wikipedia pages that contain any keyword in the title. The full list of keywords is available in Supplementary Materials in the appendix.

3.5 The School Syllabus

While Wikipedia has the potential to impact student learning in various direct and indirect ways, here we focus on whether students use Wikipedia to study their school subjects directly. Wikipedia has content on every academic subject, and might replace textbooks, which are often in short supply. The findings of this section will inform our later discussion of results on academic performance and student time use.

By manually mapping the Malawian secondary school syllabus to specific Wikipedia pages and narrow subcategories, we can show that students do use Wikipedia as a study tool.[31] Students face a tradeoff between browsing general interest pages and syllabus pages, and on average, students allocate 22 percent of their browsing time to pages directly related to the syllabus. The average student spent 6.3 hours on pages related to the school syllabus, with some students spending as many as 20 hours on school subjects (Panel A of Figure 4). Comparing this to Figure 2, we see that students spent more time on school subjects than on any general interest topic. High achievers spend more time on the syllabus than low achievers (7.5 versus 5.3 hours). We will discuss these patterns further in relation to the intervention's impact on academic performance in Section 5.

We expected Wikipedia to be useful, and used, for Biology, and students indeed browsed Biology pages more than any other subject (2.7 hours on average, Panel B of Figure 2). This was followed by other science subjects (Physics and Chemistry, one hour each), humanities (Social Science, Geography, History, Life Skills and Agriculture, thirty minutes to an hour), and, finally, English and Mathematics (below thirty minutes each). The average student spread their study time across five different school subjects (see Panel B of Figure A2).

3.6 Discussion of Student Browsing Patterns

As we examine student browsing patterns, the following stylized facts emerge. First, the intervention was effective in encouraging students to read. The aver-

[31] See the Appendix for more details about this mapping.

age student used the new resource intensively, and spent one hour and twenty minutes per week reading articles on Wikipedia. Second, individual students have broad interests: they visited a multitude of pages on a variety of topics, mostly not related to their studies. Third, students showed an interest in using Wikipedia to learn about important topics such as world events and sex and sexuality. Finally, by matching the Wikipedia pages to the school syllabus, we find that approximately a fifth of their time was spent on pages directly related to their school subjects. Students appear to find Wikipedia useful as a study tool, especially for Biology.

4 Information and the Internet

By giving students an intensive introduction to Wikipedia, we might affect how students engage with online information. First, students might become more comfortable with information technology, and learn to use it to quickly find accurate information. Second, students might learn that Wikipedia is an easy to use and reliable source of information, and develop a preference for it over other online sources.

4.1 Data and Empirical Strategy

We conducted two endline surveys. Endline Survey A took place between May and June, 2018. It had two versions: a short version that was administered to all students in Forms 2, 3, and 4, and a longer version that was administered to students in the treatment group and to the subsample of control students who were randomly selected for supplementary surveys. Endline Survey B was a survey administered to treatment students and to the subsample of control students. Endline B took place after Endline A, in June and July, 2018.

We have a low rate of attrition for both Endline Surveys A and B (Panel B of Table 1). The attrition rate for Endline A is 5 percent in both the treatment group and the subsample control group. There is significantly higher attrition in the full control group (8 percent), and we therefore include Lee (2009) bounds when interpreting the results on time use and participation in Section

5.3. The attrition rate for Endline B is 8 percent, with no differential attrition.

To investigate the impact of the intervention on survey outcomes, we estimate the following equation:

$$y_i = \beta \text{Treatment}_i + \sigma_s + \varepsilon_i. \quad (1)$$

Here, y_i is a survey outcome measure for student i at endline. Treatment_i is a indicator for treatment status. ε_i is a mean-zero error term. To estimate our standard errors consistently, we also include a fixed effect for the stratification bin, σ_s, where s is the stratification bin for student i.[32] We report heteroskedasticity-robust standard errors, as well as randomization inference p-values.[33]

We use ordinary least squares to estimate the treatment effect β. Because treatment status Treatment_i is randomly assigned, we expect the error term to be mean-independent of treatment status, $\mathbb{E}(\varepsilon_i|\text{Treatment}_i) = 0$. Therefore, in the absence of spillovers, the OLS estimate $\hat{\beta}$ is unbiased.

For some outcomes, we also interact the treatment variable with an indicator for past internet use (as reported at baseline), as past use may shape students' comparisons of Wikipedia to the broader internet.

For the outcomes examined in this section, any spillovers are likely to be positive, from treatment students to other treatment students or from treatment students to control students. Then, for all specifications, $\hat{\beta}$ is an underestimate of the true effect of a school-level intervention.

4.2 Ability to Find Information

First, we show that treatment students are more likely than control students to understand what Wikipedia is and how it can be used. In Endline A, we ask students whether they believe it is possible to find information about world events on Wikipedia, to find student exam scores, to watch movies, and to

[32] This is necessary to produce consistent standard errors (Bruhn and McKenzie, 2009).

[33] We randomize at the individual level, and therefore do not report cluster-robust standard errors (Abadie et al., 2017). Randomization inference p-values are based on 10,000 replications.

communicate with friends. We also ask students to identify several logos for internet applications, and note whether they correctly name the Wikipedia logo. We construct a summary index[34] based on correct answers to these five questions. Column 1 of Table 2 shows that relative to the control group, treatment students have a significantly better understanding of Wikipedia (0.88σ).

The intervention helped students learn how to use an internet-enabled device to find information quickly and easily. During the Endline B survey, the enumerator handed the student an internet-enabled device equipped with several internet applications including both Wikipedia and Google. The student was asked to find the number of stars in the Milky Way. Most treatment students (58 percent) are able find the correct answer within 2 minutes (Column 2 of Table 2). Only 39 percent of control students succeed at this task.

Finally, we show that students with access to the digital library have an advantage over their peers when it comes to finding information about both the news and about academic subjects, which suggests that the digital library may be useful as a study resource, over and above the resources provided by the school. We used a small experiment to capture a student's ability to find information at school. In Endline A, each student was given two quiz questions: a news question and an academic question.[35] These questions were different for every student. We incentivized students to find the correct answer in time for the next endline survey: students were told that two weeks later, during Endline B, they would be given a prize for each correct answer. The digital library was open between the two endline surveys. Students in the treatment group are 9 percentage points more likely to find the answer to the news question (Column 3 of Table 2). They are also 11 percentage points more likely to correctly answer the academic question (Column 4 of Table 2). This is perhaps more surprising, as all students had access to the school library, their

[34]We are guided by Anderson (2008) in our construction of a summary index. For each outcome in the index, we standardize by subtracting the mean and dividing by the control group standard deviation. For each student, the index variable is a weighted mean of these standardized outcomes. The weights are determined by the inverse of the covariance matrix for the standardized outcomes. These procedure gives less weight to highly correlated outcomes and outcomes with many missing values.

[35]The student drew each question from a hat, and kept the slip of paper. See Supplementary Materials in the appendix for a list of sample questions.

notes and their teachers.

4.3 Preference for Wikipedia over the Internet

Early access to Wikipedia might affect the way that young people search for online information and trust its accuracy. Here, we ask whether students prefer Wikipedia to the wider internet, for which types of information, and why.

Using survey results from Endline B, we show that students in the treatment group are more likely to prefer Wikipedia for information about safe sex and the news (Columns 1 and 2 of Table 3). A majority of treatment group students prefer Wikipedia to the internet for these two topics. Indeed, we saw in Section 3 that students did often search for both of these topics. This preference for Wikipedia appears for both students with and without past internet experience. However, when we consider news information, the Wikipedia preference is smaller and insignificant for students with past internet experience. This suggests that the intervention may not shift students away from online news sources altogether.

We also show that treatment students are more likely to actually choose Wikipedia over other internet information sources. As previously mentioned, during Endline B we asked students to find the number of stars in the Milky Way, using a device equipped with both Wikipedia and Google Chrome, as well as other applications. Treatment group students were twice as likely to use Wikipedia for this task (Column 3).

We find that treatment group students are more likely to find information on Wikipedia trustworthy, easy to understand and easy to find as compared to information on the internet (Columns 4 to 6). Again, these effects appear for both students with and without past internet experience. Within the treatment group, a large majority prefer Wikipedia to the internet along each of the three dimensions.

4.4 Discussion of Student Information Seeking

Students were able to effectively use Wikipedia to find information, including academic information that their peers could not find using school resources. It also appears that intensive exposure to Wikipedia generated a strong preference for Wikipedia over other online information sources. This is true for sensitive and controversial topics such as world news and safe sex. These topics are interesting to students yet seldom discussed in school books, and a lack of accurate information has serious consequences. In addition to finding Wikipedia easy to use, treatment group students are keenly aware that it is a relatively accurate and trustworthy information source.

5 Academic Performance

In this section we investigate the impact of Wikipedia access on academic performance, as well as student time use, class participation, and aspirations. Wikipedia might improve English language skills, and English exam scores, by offering compelling and accessible reading material. It might be used directly as a study tool in place of, or in support of, textbooks, notes and teachers. We saw in Section 3 that many students use Wikipedia as a study tool, especially for Biology. Wikipedia content might inspire students to higher aspirations, or shape student interests. There is also the potential for a negative impact, if Wikipedia acts primarily as a form of entertainment or distraction. Given the share of browsing time devoted to non-syllabus topics (Section 3.5), this is a potential concern.

5.1 Data and Outcomes

Our primary outcomes are English and Biology exam scores in the final term.[36] We selected these two subjects as primary outcomes for several reasons. If Wikipedia serves as a literacy intervention, English language skills should im-

[36] We pre-registered term 3 English and Biology scores as our two primary outcomes (AEA RCT Registry number AEARCTR-0003824).

prove over time and impact English exam scores.[37] Biology exams require students to absorb a large amount of information, and Biology students are likely to benefit from additional study materials. At baseline, English and Biology are the most popular subjects in our sample, as measured by enrollment and stated preference. Biology is especially important for students' career prospects. At baseline, a majority of students aspired to become doctors, nurses, or other health care professionals. Many of the students who pass their final exams do go on to college programs in nursing, medicine, or other health specialties. This interest in Biology reflects career prospects for Malawian secondary school graduates more generally; the schools in our sample do not have any particular focus on Biology or health care.

We also measure impacts on other academic subjects including Mathematics, Chichewa (a local language), science subjects and the humanities. We did not expect Wikipedia access to have a positive impact on these subjects. Math and science subjects are primarily skill based. While Wikipedia does have a version in Chichewa, it hosts only a few hundred pages, none of which were visited by our students. The humanities are unpopular with students at baseline, as measured by enrollment and stated preference. We therefore did not include any of these subjects as primary outcomes. However, we might expect a negative impact if students shift time away from those subjects towards English, Biology, or online distraction.

To measure academic performance, we use administrative data on school exam scores. We collected exam scores for all subjects in all three terms, as well as end-of-year scores for the year before the intervention began.

For each core subject (English, Biology, Math and Chichewa), we construct a separate outcome variable y_{it} representing the standardized exam score in that subject, for student i in term t.[38] Other subjects are offered as electives, or are offered only in certain schools and forms. We combine similar subjects us-

[37] In our setting, English exam scores measure English language ability. We include a sample English exam in Supplementary Materials in the appendix.

[38] These values are standardized within term, form and school by subtracting the mean and dividing by the control group standard deviation. We then subtract the overall control group mean. Form 4 students do not receive school exam scores at the end of term 3, and for these students we instead collected data on national exam scores.

ing an index measure that assigns weight to non-missing values at the student level (again following Anderson, 2008). We construct an outcome for science subjects (Physics and Chemistry) and a separate outcome for subjects which we loosely define as humanities (Social Science, Geography, History, Life Skills and Agriculture). We again standardize each subject before constructing the summary index measure y_{it}.

Administrative data is missing for a few exam scores, as some students drop out or miss an exam. We are missing data for approximately 7 percent of students (Panel B of Table 1).

5.2 Empirical Strategy and Main Results

We estimate the effects of the intervention on exam scores for each subject in the final term.

$$y_i = \beta \text{Treatment}_i + \delta \left(y_{i0} \times \text{Data}_{i0} \right) + \delta_0 \, \text{MissingData}_{i0} + \sigma_s + \varepsilon_i \qquad (2)$$

Here, y_i is the measure of academic performance for student i in term 3. Treatment_i is an indicator for treatment status. ε_i is a mean-zero error term. To improve precision, we control for the baseline measure of the outcome, y_{i0}, taken from term 3 of the previous school year. We dummy out missing baseline scores: Data_{i0} and MissingData_{i0} are indicators for whether or not we have baseline data y_{i0} for student i.[39] We include a fixed effect for the stratification bin. We report robust standard errors, as well as randomization inference p-values based on 10,000 replications. Our parameter of interest is the average treatment effect β.

Because treatment status Treatment_i is randomly assigned at the student level, we expect the error term to be mean-independent of treatment status, $\mathbb{E}(\varepsilon_i | \text{Treatment}_i) = 0$. Therefore, in the absence of spillovers, the OLS estimate of β is unbiased.

[39] We do not use *pdslasso* to select covariates as we are missing baseline exam score data for 4 percent of students. Baseline exam score is likely to be an important covariate, and dropping these observations will reduce statistical power. We instead use indicator variables for missing baseline data.

We also estimate heterogenous treatment effects by baseline achievement, interacting the treatment variable with an indicator for high achievement (above median average score in English and Biology) at baseline.

Spillovers are possible in our setting, from treatment students to other treatment students or to control students. While information is likely to be shared between treatment and control students, any impact on English language skills or Biology exam scores is likely to be small without direct access to the reading material. We hypothesize small, positive spillovers, which would imply that our estimates are a slight underestimate of the impact of a school-level intervention. In the appendix (Table A3), we provide some evidence that this is indeed the case, using a specification that controls for treated study friends.[40] A different type of spillover may operate through teacher behavior, however, given the sparse treatment, and standardized syllabus and exams, there was little opportunity for teachers to adapt to the intervention. We discuss this further in Section 6.1.

We find a significant impact on English exam scores, overall (0.10σ) and for low achievers (0.20σ), and a significant impact on Biology scores for low achievers (0.14σ, see Table 4). We find no significant impact on high achievers in either subject.

We do not find any impact, positive or negative, on other school subjects (see Table 5). Average treatment effects are between -0.03σ and 0.04σ for Mathematics, and science and humanities subjects. We cannot rule out a small negative impact on Chichewa; while insignificant, the estimate is -0.07σ. We do not find any significant impact on low or high achievers for any of these subjects.

[40]The spillover effect specification in Table A3 in the appendix contains controls for the number of named study friends at baseline, treated study friends and treated study friends interacted with being a control student. It is difficult to fully capture spillovers using a baseline networks, and doing so in our case introduces noise. In fact, our friendship networks are endogenous to the treatment itself, a finding which will be explored in depth in future research. We choose to rely on study friend networks because Malawian schools assign students to "study circles" at the beginning of the school year, and so such friendship networks are less responsive to the intervention.

5.3 Time Use, Class Participation and Aspirations

We next examine student time use across different activities, to determine whether treatment students substituted away from study time to spend time in the digital library. We collected time use data from all students in Endline A, while the digital library was still in operation. We asked students to recall their time spent on specific activities, day by day, for the three days preceding the survey. We then classify time use as studying, recreation or sleep.[41] Study time includes time the students spent studying in the digital library, but not other browsing time. We use equation (1) to estimate the impact of the intervention on study time, recreation time, and sleep. Because Endline A was subject to differential attrition in the full control group, we also report Lee (2009) bounds.

It appears that students did not take time away from their studies to visit the digital library, and did not cut back on sleep (Columns 1 and 4 of Table 6).[42] Rather, the digital library crowded out time spent hanging out with friends, playing sports, and attending religious activities (Column 3 of Table 6). Low and high achievers reallocated their time in a similar way (Column 2 of Table 6).

Access to Wikipedia might affect class participation by increasing student confidence, motivation, or interest. In Endline A, we asked each student to report the number of times they raised their hand in each class, day by day, over three days. We then take the average number of times they raised their hand over the three days. We estimate equation (1).

We do not find evidence for a change in class participation. On average, a student raises their hand three times per school day. There is no significant difference between treatment and control students (Column 5 of Table 6).

Finally, Wikipedia might affect student aspirations, by helping students plan for a career or introducing new role models. We ask students, at baseline

[41] We compute average daily study time by summing time spent studying alone and time spent studying with others. To construct a measure of time spent on recreational activities, we sum the time spent hanging out with friends, in school clubs, religious activities, sports activities and any other activities. Finally, we asked students the time at which they woke up and went to bed, and compute average awake time over the previous three days.

[42] Time spent on general browsing (but not studying) in the digital library is an omitted category.

and in Endline Survey B, which career they hope to have in the future. In Endline B, we also ask students to name the college they will most likely attend, as well as their dream college.

We define an indicator variable for a change in career choice between baseline and endline.[43] We use equation (1) to estimate the impact of the intervention on the likelihood of a change in career aspirations.

The intervention does not appear to cause students to change their career aspirations (Column 6 of Table 6). At endline, treatment students and control students choose similar types of careers, with most aspiring to health care positions. In Figure A3 we present the career aspirations of treatment and control students at endline. There are no clear systematic differences. In Panels A and B of Figure A4, we present most likely and dream colleges reported by treatment and control students, and again see no systematic differences. We note that our pool of students had high aspirations at baseline, suggesting limited scope for an increase in self-reported aspirations. At baseline, one third of students hoped to become a doctor, specialist doctor or surgeon.

6 Discussion and Conclusion

We find that providing students with access to Wikipedia affects students along three key dimensions. First, it affects their perceptions of online information. Students are more likely to prefer Wikipedia to the broader internet for its usability and trustworthiness, even for important controversial topics like world news and safe sex. Second, students use Wikipedia intensively, and read articles, in English, on a broad range of topics of general interest. This access to wide ranging reading material, during a full school year, leads to positive gains in English exam scores, especially for low-achieving students. Finally, students use Wikipedia as a study tool for Biology, and prefer it to their textbooks. This has a significant impact on exam scores for low-achieving students, whose study time becomes more productive with access to Wikipedia.

[43] The outcome variable is coded as equal to one if the individual reported any career choice change between baseline and endline surveys. This can arise due to change in career as well as a change in precision (for example, "doctor" in the baseline to "neurologist" in the endline)

6.1 Mechanisms

Though students spent more than an hour per week in the digital library, and spent most of that time on topics unrelated to the school syllabus, Wikipedia did not have a negative impact on academic performance. Using 95 percent confidence intervals, we can rule out negative effects for English scores, and effects below -0.03σ for Biology. We also find no impact on scores in Mathematics, science subjects, or humanities subjects, with point estimates between -0.03σ and 0.04σ. The impact on Chichewa is insignificant at -0.07σ. We cannot rule out small substitution effects from Chichewa, as students shift their attention away from that subject towards subjects taught in English.

Rather, we find a positive impact on English exam scores, which leads us to view student browsing behavior in a different light. English exams are a good test of English language ability; they include multiple choice questions that test student understanding of words, sentences, and grammar, and essay questions. If Wikipedia serves as a literacy intervention, it matters less whether students choose to read about academic topics. In fact, we posit that Wikipedia is effective as a literacy tool precisely because it gives students access to reading material on any topic they choose.

The impacts on both English and Biology exam scores are larger for low achievers, nearly half of whom had a failing score at baseline. Low achievers with access to Wikipedia score 0.20σ higher in English and 0.14σ higher in Biology than their counterparts without Wikipedia access. In the final term, the English score gap between low and high achievers is closed by one fifth due to Wikipedia access. Because government boarding schools are academically competitive, a low-achieving student in one of the study schools may in fact better represent the typical Malawian secondary school student. For this reason, impacts among low achievers are particularly relevant.

We do not find any impact on high achievers for any subject. It appears that for highly literate students, access to Wikipedia serves as equal part distraction and input to academic performance, with a net effect of zero.

Neither low nor high achievers increased their study time in response to the intervention, yet Biology scores improved for low achievers. Study

time must have become more productive, in particular for low achievers. If Wikipedia is easier to use and understand than standard textbooks, this would explain a rapid increase in study time productivity, especially among students who are struggling. In Endline B, we asked students to rank Wikipedia against their textbooks for each subject. We find that 56 percent of treatment students preferred Wikipedia for Biology. This is consistent with the focus on Biology we saw in the browsing data. Students spent at least twice as much time on Biology as on any other subject (Figure 2). This is also consistent with the small experiment we conducted in Section 4.2, showing that students with Wikipedia access were able to find academic information that control students were not (Column 4 of Table 2). Finally, in Section 4.3, we saw that students find and understand information on Wikipedia easily. Taken together, these results indicate that Wikipedia serves as a useful and accessible study tool for Biology.

It is plausible that Biology exam scores are subject to positive spillovers from treatment students to both treatment and control students, if students shared the information they learned on Wikipedia. In this case, the effect sizes we estimate understate the true effect of an intervention at scale. English exam scores are less likely to be subject to spillovers. If these gains represent an improvement in English language ability, they are likely due to direct exposure to reading material. Any spillovers appear to be small and positive for both Biology and English exam scores. Controlling for baseline study friends increases the average treatment effect and effect for low achievers (see the appendix).

Because we randomized access to Wikipedia at the student level, our ability to measure general equilibrium effects is limited. For example, if all teachers and students had access to Wikipedia, teachers might be able to incorporate it into their lesson plans. On the one hand, teachers follow a strict syllabus from the Ministry of Education with little room for adaptation. On the other hand, involving teachers typically improves the efficacy of literacy and other primary school interventions. There is less evidence in secondary schools, and this could be the subject of future research.

6.2 Cost-Effectiveness, Policy Implications and Future Research

Providing access to Wikipedia is cost-effective as a substitute for other types of reading materials, and as a literacy intervention in general. We estimate that our intervention, as implemented, costs less than $4 USD per student per month, or $28 USD per 0.1σ of improvement in English scores. This includes the cost of project management, digital library staff, internet-enabled devices and internet data packages. In many developing countries, Internet.org provides access to Wikipedia for free.[44] Providing access to Wikipedia through Internet.org would reduce the intervention cost to less than $3 USD per student per month. This is more cost-effective than programs that provide reading material to primary schools, as most have no impact. It is also more cost-effective than many computer-aided learning programs.[45] Our cost-effectiveness is similar to many primary school interventions that increase the teacher-student ratio, provide incentives for teacher incentives, or provide remedial lessons, but lower than programs that provide performance incentives to teachers (McEwan, 2015). There are some reasons to expect smaller returns in secondary school, as subject matter increases in difficulty, and students are starting from a higher level of baseline literacy.

This study has important policy implications for educators. Where textbooks are in short supply, Wikipedia might serve as a useful and inexpensive substitute. It is common for universities and workplaces to restrict access to certain websites, and secondary schools might do the same. Wikipedia may in fact be easier to use and understand than classic textbooks, especially for students who are struggling. For students with lower literacy levels, Wikipedia, with both English and Simple English options, is a low-cost and effective literacy intervention. Not only is the reading material simple and informative, it engages student interest. Students are excited to use Wikipedia, and choose

[44] Internet.org is a partnership between social media and telecommunications firms that provides free access to selected Internet services in poor countries.

[45] For example, Muralidharan, Singh and Ganimian (2019) show that Mindspark, a computer-aided learning platform for primary school students, generates a language score impact at a cost of $39 per 0.1σ. However, they do also find a significant impact on math scores, suggesting that overall their program might be considered highly cost effective. Indeed, they find that this is more cost effective than default public spending in India.

to spend a great deal of time reading. This translates to real gains in English language ability.

Providing Wikipedia to students serves as an appropriate introduction to online information, and might affect the way young people use the internet more broadly. After graduation, many of the students in this study will have access to the internet on a regular basis. In future research, it will be important to measure the long run effects of this intervention on internet use and the ability to find accurate and trustworthy information online.

References

Abadie, Alberto, Susan Athey, Guido W. Imbens, and Jeffrey Wooldridge. 2017. "When Should You Adjust Standard Errors for Clustering?" National Bureau of Economic Research Working Paper 24003.

Abeberese, Ama Baafra, Todd J. Kumler, and Leigh L. Linden. 2014. "Improving Reading Skills by Encouraging Children to Read in School: A Randomized Evaluation of the Sa Aklat Sisikat Reading Program in the Philippines." *Journal of Human Resources*, 49(3): 611–633.

Allcott, Hunt, and Matthew Gentzkow. 2017. "Social Media and Fake News in the 2016 Election." *Journal of Economic Perspectives*, 31(2): 211–36.

Anderson, Michael L. 2008. "Multiple Inference and Gender Differences in the Effects of Early Intervention: A Reevaluation of the Abecedarian, Perry Preschool, and Early Training Projects." *Journal of the American Statistical Association*, 103(484): 1481–1495.

Bailard, Catie Snow. 2012. "A Field Experiment on the Internet's Effect in an African Election: Savvier Citizens, Disaffected Voters, or Both?" *Journal of Communication*, 62(2): 330–344.

Banerjee, Abhijit, Paul Glewwe, Shawn Powers, and Melanie Wasserman. 2013. "Expanding Access and Increasing Student Learning in Post-Primary

Education in Developing Countries: A Review of the Evidence." Abdul Latif Jameel Poverty Action Lab

Banerjee, Abhijit V., Shawn Cole, Esther Duflo, and Leigh L. Linden. 2007. "Remedying Education: Evidence from Two Randomized Experiments in India." *The Quarterly Journal of Economics*, 122(3): 1235–1264.

Barrera-Osorio, Felipe, and Leigh L. Linden. 2009. "The Use and Misuse of Computers in Education: Evidence from a Randomized Experiment in Colombia." The World Bank 4836, The World Bank.

Barro, Robert J., and Jong Wha Lee. 2013. "A New Data Set of Educational Attainment in the World, 1950–2010." *Journal of Development Economics*, 104: 184–198.

Barrow, Lisa, Lisa Markman, and Cecilia Elena Rouse. 2009. "Technology's Edge: The Educational Benefits of Computer-Aided Instruction." *American Economic Journal: Economic Policy*, 1(1): 52–74.

Batzilis, Dimitrios, Taryn Dinkelman, Emily Oster, Rebecca Thornton, and Deric Zanera. 2010. "New Cellular Networks in Malawi: Correlates of Service Rollout and Network Performance." National Bureau of Economic Research Working Paper 16616.

Beg, Sabrin, Adrienne Lucas, Waqas Halim, and Umar Saif. 2019. "Beyond the Basics: Improving Post-Primary Content Delivery through Classroom Technology." National Bureau of Economic Research Working Paper 25704.

Beuermann, Diether W., Julian Cristia, Santiago Cueto, Ofer Malamud, and Yyannu Cruz-Aguayo. 2015. "One Laptop per Child at Home: Short-Term Impacts from a Randomized Experiment in Peru." *American Economic Journal: Applied Economics*, 7(2): 53–80.

Borkum, Evan, Fang He, and Leigh L. Linden. 2012. "The Effects of School Libraries on Language Skills: Evidence from a Randomized Controlled Trial in India." National Bureau of Economic Research Working Paper 18183.

Bruhn, Miriam, and David McKenzie. 2009. "In Pursuit of Balance: Randomization in Practice in Development Field Experiments." *American Economic Journal: Applied Economics*, 1(4): 200–232.

Bulman, George, and Robert W. Fairlie. 2016. "Technology and Education: Computers, Software, and the Internet." In *Handbook of the Economics of Education*. Vol. 5, 239–280. Elsevier.

Buys, Piet, Susmita Dasgupta, Timothy S. Thomas, and David Wheeler. 2009. "Determinants of a Digital Divide in Sub-Saharan Africa: A Spatial Econometric Analysis of Cell Phone Coverage." *World Development*, 37(9): 1494–1505.

Campante, Filipe, Ruben Durante, and Francesco Sobbrio. 2018. "Politics 2.0: The Multifaceted Effect of Broadband Internet on Political Participation." *Journal of the European Economic Association*, 16(4): 1094–1136.

Carrillo, Paul E., Mercedes Onofa, and Juan Ponce. 2011. "Information Technology and Student Achievement: Evidence from a Randomized Experiment in Ecuador." Inter-American Development Bank Working Paper IDB-WP-223, Washington, DC.

Chen, Yan, Rosta Farzan, Robert E. Kraut, Iman YeckehZaare, and Ark Fangzhou Zhang. 2018. "Motivating Contributions to Public Information Goods: A Field Experiment on Wikipedia." *Working Paper*.

Chen, Yuyu, and David Y. Yang. 2019. "The Impact of Media Censorship: 1984 or Brave New World?" *American Economic Review*, 109(6): 2294–2332.

Chung, Chung Joo, Yoonjae Nam, and Michael A. Stefanone. 2012. "Exploring Online News Credibility: The Relative Influence of Traditional and Technological Factors." *Journal of Computer-Mediated Communication*, 17(2): 171–186.

Cristia, Julian, Pablo Ibarrarán, Santiago Cueto, Ana Santiago, and Eugenio Severín. 2017. "Technology and Child Development: Evidence from the One

Laptop per Child Program." *American Economic Journal: Applied Economics*, 9(3): 295–320.

de Hoop, Jabobus. 2010. "Selective Secondary Education and School Participation in Sub-Saharan Africa: Evidence from Malawi." *Tinbergen Institute Discussion Papers Series 10-041/2*.

Derksen, Laura, and Joep van Oosterhout. 2018. "Love in the Time of HIV: Testing as a Signal of Risk." *Working Paper*.

Dupas, Pascaline. 2011. "Do Teenagers Respond to HIV Risk Information? Evidence from a Field Experiment in Kenya." *American Economic Journal: Applied Economics*, 3(1): 1–34.

Faber, Benjamin, Rosa Sanchis-Guarner, and Felix Weinhardt. 2015. "ICT and Education: Evidence from Student Home Addresses." National Bureau of Economic Research Working Paper 21306.

Fairlie, Robert W., and Jonathan Robinson. 2013. "Experimental Evidence on the Effects of Home Computers on Academic Achievement Among Schoolchildren." *American Economic Journal: Applied Economics*, 5(3): 211–40.

Galperin, Hernan, and M. Fernanda Viecens. 2017. "Connected for Development? Theory and Evidence about the Impact of Internet Technologies on Poverty Alleviation." *Development Policy Review*, 35(3): 315–336.

Giles, Jim. 2005. "Internet Encyclopaedias Go Head to Head." *Nature*.

Glewwe, Paul, and Karthik Muralidharan. 2016. "Improving Education Outcomes in Developing Countries: Evidence, Knowledge Gaps, and Policy Implications." In *Handbook of the Economics of Education*. Vol. 5, 653–743. Elsevier.

Glewwe, Paul, Michael Kremer, and Sylvie Moulin. 2009. "Many Children Left Behind? Textbooks and Test Scores in Kenya." *American Economic Journal: Applied Economics*, 1(1): 112–35.

Goolsbee, Austan, and Jonathan Guryan. 2006. "The Impact of Internet Subsidies in Public Schools." *The Review of Economics and Statistics*, 88(2): 336–347.

Head, Alison J., and Michael B. Eisenberg. 2010. "How Today's College Students Use Wikipedia for Course-Related Research." *First Monday*, 15(3).

He, Fang, Leigh L. Linden, and Margaret MacLeod. 2008. "How to Teach English in India: Testing the Relative Productivity of Instruction Methods within the Pratham English Language Education Program." Abdul Latif Jameel Poverty Action Lab Working Paper.

Heilman, James M, Eckhard Kemmann, Michael Bonert, Anwesh Chatterjee, Brent Ragar, Graham M. Beards, David J. Iberri, Matthew Harvey, Brendan Thomas, Wouter Stomp, et al. 2011. "Wikipedia: A Key Tool for Global Public Health Promotion." *Journal of medical Internet research*, 13(1): e14.

Hinnosaar, Marit. 2019. "Gender Inequality in New Media: Evidence from Wikipedia." *Journal of Economic Behavior & Organization*, 163: 262–276.

Hjort, Jonas, and Jonas Poulsen. 2019. "The Arrival of Fast Internet and Employment in Africa." *American Economic Review*, 109(3): 1032–1079.

Kerwin, Jason. 2018. "Scared Straight or Scared to Death? The Effect of Risk Beliefs on Risky Behaviors." *Working Paper*.

Kerwin, Jason, and Rebecca L. Thornton. 2018. "Making the Grade: The Sensitivity of Education Program Effectiveness to Input Choices and Outcome Measures." Social Science Research Network SSRN Scholarly Paper ID 3002723, Rochester, NY.

Knauer, Heather A., Pamela Jakiela, Owen Ozier, Frances Aboud, and Lia C. H. Fernald. 2020. "Enhancing Young Children's Language Acquisition through Parent–Child Book-Sharing: A Randomized Trial in Rural Kenya." *Early Childhood Research Quarterly*, 50: 179–190.

Lazer, David MJ, Matthew A. Baum, Yochai Benkler, Adam J. Berinsky, Kelly M. Greenhill, Filippo Menczer, Miriam J. Metzger, Brendan Nyhan, Gordon Pennycook, David Rothschild, et al. 2018. "The Science of Fake News." *Science*, 359(6380): 1094–1096.

Lee, David S. 2009. "Training, Wages, and Sample Selection: Estimating Sharp Bounds on Treatment Effects." *The Review of Economic Studies*, 76(3): 1071–1102.

Lih, Andrew. 2004. "Wikipedia as Participatory Journalism: Reliable Sources? Metrics for Evaluating Collaborative Media as a News Resource." *Nature*, 3(1): 1–31.

Lim, Sook. 2009. "How and Why do College Students Use Wikipedia?" *Journal of the American Society for Information Science and Technology*, 60(11): 2189–2202.

Linden, Leigh L. 2008. "Complement or Substitute?: The Effect of Technology on Student Achievement in India." Abdul Latif Jameel Poverty Action Lab (J-PAL) Working Paper.

Lucas, Adrienne M., Patrick J. McEwan, Moses Ngware, and Moses Oketch. 2014. "Improving Early-Grade Literacy in East Africa: Experimental Evidence from Kenya and Uganda." *Journal of Policy Analysis and Management*, 33(4): 950–976.

Machin, Stephen, and Sandra McNally. 2008. "The Literacy Hour." *Journal of Public Economics*, 92(5): 1441–1462.

MacMillan, Margy, and Allison MacKenzie. 2012. "Strategies for Integrating Information Literacy and Academic Literacy: Helping Undergraduate Students Make the Most of Scholarly Articles." *Library Management*, 33(8/9): 525–535.

Malamud, Ofer. 2019. "The Effect of Home Computers and the Internet on Children's Human Capital Development." *DICE Report*, 17(2): 34–40.

Malamud, Ofer, and Cristian Pop-Eleches. 2011. "Home Computer Use and the Development of Human Capital." *The Quarterly Journal of Economics*, 126(2): 987–1027.

Malamud, Ofer, Santiago Cueto, Julian Cristia, and Diether W. Beuermann. 2019. "Do Children Benefit from Internet Access? Experimental Evidence from Peru." *Journal of Development Economics*, 138: 41–56.

McEwan, Patrick J. 2015. "Improving Learning in Primary Schools of Developing Countries: A Meta-Analysis of Randomized Experiments." *Review of Educational Research*, 85(3): 353–394.

Mesgari, Mostafa, Chitu Okoli, Mohamad Mehdi, Finn Arup Nielsen, and Arto Lanamaki. 2015. ""The Sum of All Human Knowledge": A Systematic Review of Scholarly Research on the Content of Wikipedia." *Journal of the Association for Information Science and Technology*, 66(2): 219–245.

Miguel, Edward, and Michael Kremer. 2004. "Worms: Identifying Impacts on Education and Health in the Presence of Treatment Externalities." *Econometrica*, 72(1): 159–217.

Miner, Luke. 2015. "The Unintended Consequences of Internet Diffusion: Evidence from Malaysia." *Journal of Public Economics*, 132: 66–78.

Mo, Di, Johan Swinnen, Linxiu Zhang, Hongmei Yi, Qinghe Qu, Matthew Boswell, and Scott Rozelle. 2013. "Can One Laptop per Child Reduce the Digital Divide and Educational Gap? Evidence from a Randomized Experiment in Migrant Schools in Beijing." *World Development*, 46.

Muralidharan, Karthik, Abhijeet Singh, and Alejandro J. Ganimian. 2019. "Disrupting Education? Experimental Evidence on Technology-Aided Instruction in India." *American Economic Review*, 109(4): 1426–1460.

National Statistical Office/Malawi, and ICF. 2017. "Malawi Demographic and Health Survey 2015-16." National Statistical Office and ICF, Zomba, Malawi.

Ozier, Owen. 2015. "The Impact of Secondary Schooling in Kenya: A Regression Discontinuity Analysis." The World Bank Policy Research Working Papers 7384, The World Bank.

Rosenzweig, Roy. 2006. "Can History be Open Source? Wikipedia and the Future of the Past." *The Journal of American History*, 93(1): 117–146.

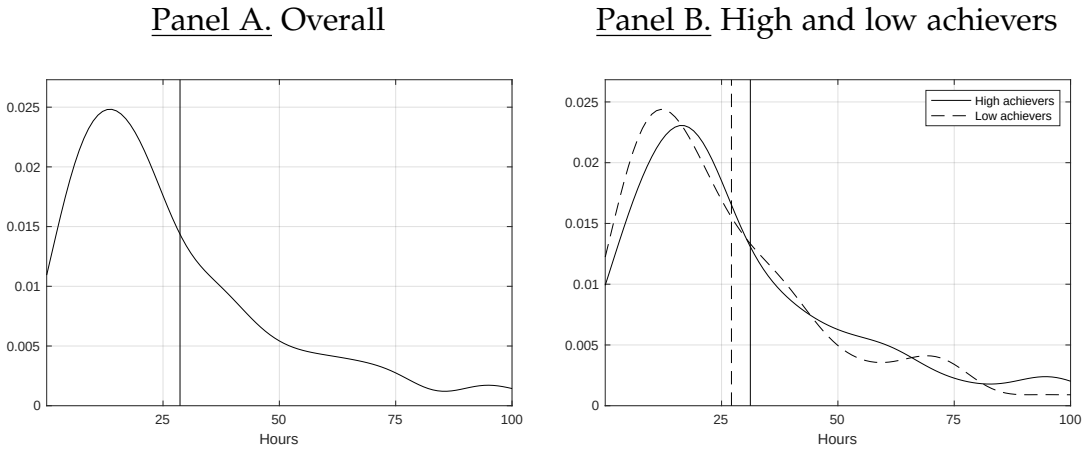
Sabarwal, Shwetlena, David K. Evans, and Anastasia Marshak. 2014. "The Permanent Input Hypothesis: The Case of Textbooks and (No) Student Learning in Sierra Leone." The World Bank Policy Research Working Papers, The World Bank.

Vigdor, Jacob L., Helen F. Ladd, and Erika Martinez. 2014. "Scaling the Digital Divide: Home Computer Technology and Student Achievement." *Economic Inquiry*, 52(3): 1103–1119.

Yanguas, Maria Lucia. 2018. "Technology and Educational Choices: Evidence from a One-Laptop-per-Child Program." UCLA CCPR Population Working Paper.

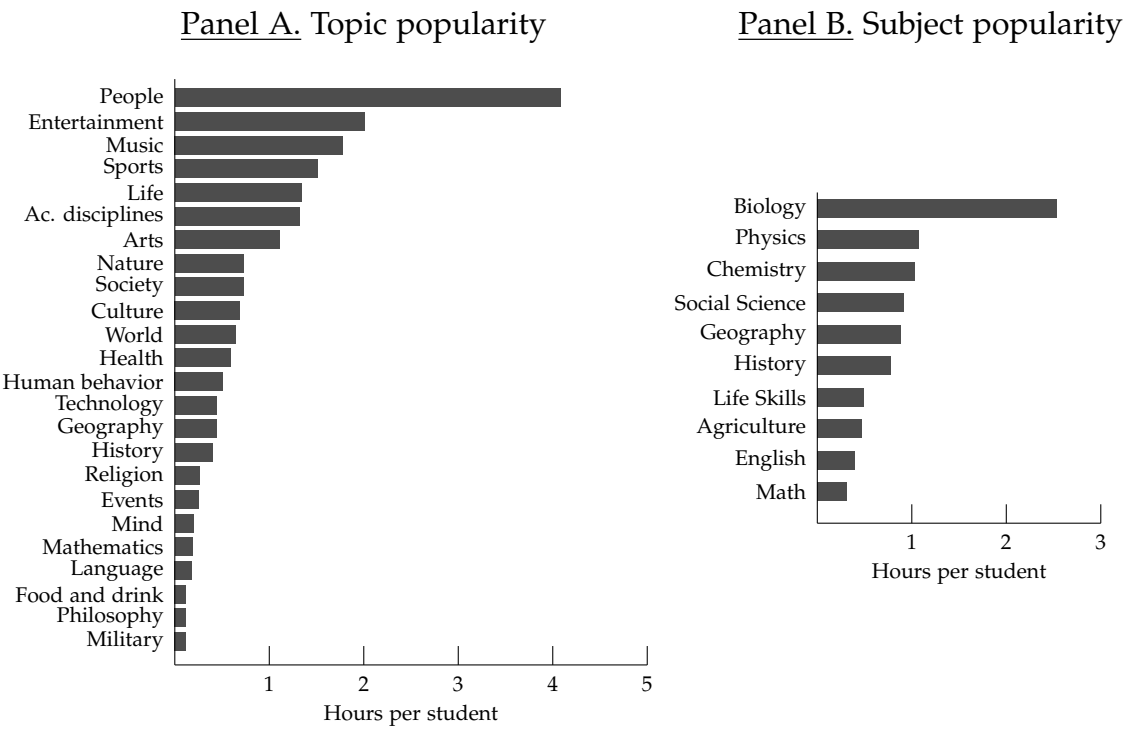
Figures

Figure 1: Histogram of Hours Spent Browsing Wikipedia



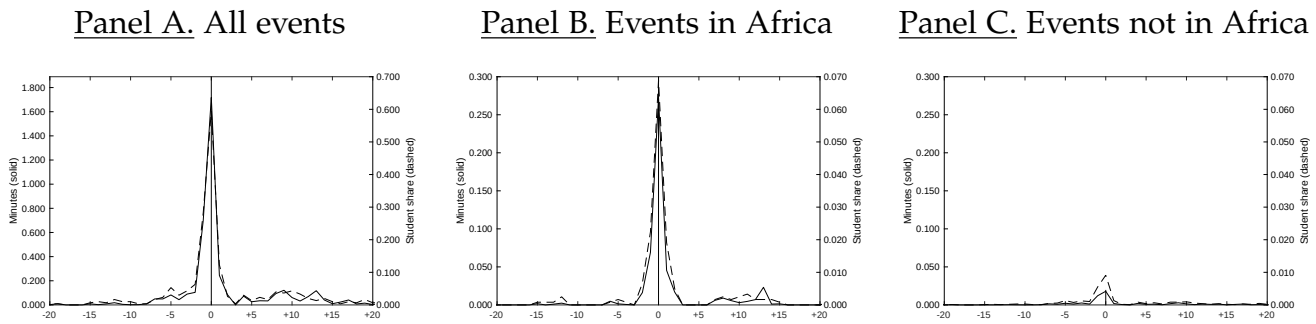
Notes: Density of browsing hours, across treatment students only, aggregated over the course of one academic year. Panel A: Vertical line is the average hours spent browsing of 28.6 hours. Panel B: 31.2 for high achievers and 27.1 for low achievers. The digital library was open for 20-22 weeks, from November 2017 to June 2018, excluding Christmas and Easter vacations.

Figure 2: Hours Spent Browsing Wikipedia by Topic and School Subject



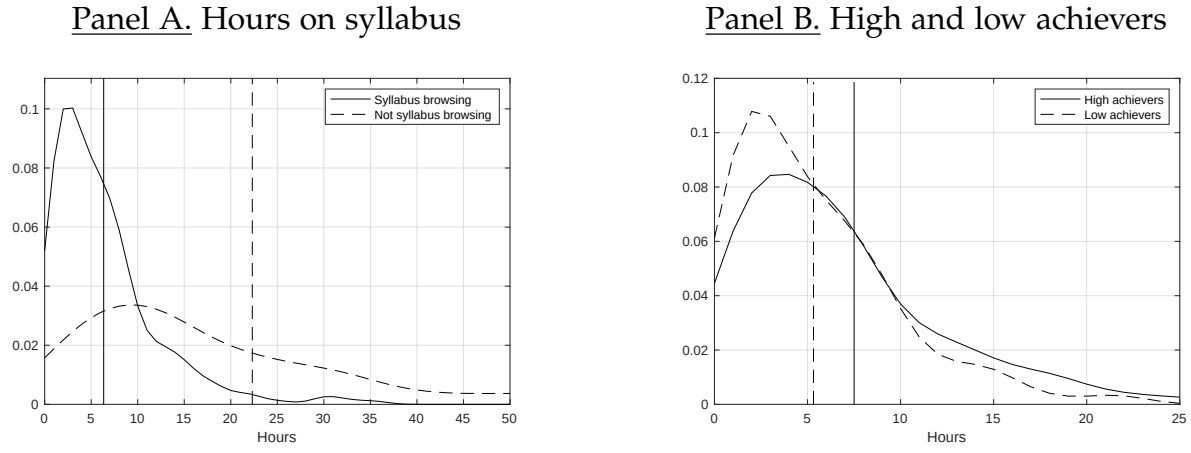
Notes: Panel A: Browsing hours per topic, per student, aggregated over the course of one academic year. See Appendix for details on topic classification. The topics Business, Concepts, Crime, Economy, Education, Energy, Government, Humanities, Knowledge, Law, Objects, Organizations, Politics, Science, and Universe are excluded from the figure and are less than 0.12 hours. Panel B: Browsing hours per school subject, per student, aggregated over the course of one academic year.

Figure 3: Wikipedia Browsing for News about World Events in 2017-18



Notes: Panel A: Average browsing minutes by student on Wikipedia pages related to all worldwide events on the left hand side axis (solid line). Share of students that visited the pages associated to at least one event on the right hand axis (dashed lines). Panels B and C: Average amount of browsing minutes by student and event on Wikipedia pages related to all worldwide events on the left hand side axis (solid lines). Share of students that visited the pages associated to a given event on the right hand axis (dashed lines). All events from November 2nd 2017 to May 9th 2018 as reported in https://en.wikipedia.org/wiki/2017 and https://en.wikipedia.org/wiki/2018 were considered, for the 20 weeks before and after they occurred. Total of 64 events, out of which 10 did not have a specific Wikipedia page associated to them and were eliminated from the analysis. Week of the event is set at zero. Negative (positive) numbers on the x-axis are weeks before (after) the event.

Figure 4: Hours Spent Browsing Pages Related to the School Syllabus



Notes: Density of browsing hours, across treatment students only, aggregated over the course of one academic year for school subjects syllabus-related pages. Panel A: Hours on syllabus- and non syllabus-related Wikipedia pages. Vertical lines are the average hours spent browsing (6.3 on syllabus and 22.3 on non syllabus-related). Panel B: High (low) achievers defined as above (below) median exam scores at the baseline. Vertical lines are the average hours spent browsing (7.5 for high achievers and 5.3 for low achievers).

Tables

Table 1: Balance Table and Attrition in Endline Surveys and Exam Scores

	(1) Treatment	(2) Control (subsample)	(3) p-value	(4) Control (full)	(5) p-value
Panel A. Balance, non-stratification variables					
Average exam score in Biology and English	55.530	55.491	.973	55.555	.977
	(13.427)	(13.844)		(13.746)	
Age	15.973	16.060	.577	16.033	.635
	(1.971)	(1.845)		(1.869)	
Female	.452	.433	.641	.423	.361
	(.499)	(.496)		(.494)	
District of origin	.605	.574	.444	.575	.348
	(.490)	(.495)		(.495)	
Mother's education	.746	.698	.224	.718	.258
	(.436)	(.460)		(.450)	
Father's education	.849	.852	.918	.856	.775
	(.359)	(.356)		(.351)	
Household has electricity	.611	.557	.179	.576	.262
	(.488)	(.498)		(.494)	
Household has mobile phone	.870	.849	.451	.866	.852
	(.336)	(.359)		(.340)	
Panel B. Attrition					
Endline A	.047	.050	.653	.076	.027
	(.211)	(.219)		(.265)	
Endline B	.083	.084	.933	—	—
	(.276)	(.278)			
Exam Scores (English)	.060	.050	.680	.065	.736
	(.238)	(.219)		(.246)	
Exam Scores (Biology)	.063	.054	.700	.069	.715
	(.244)	(.226)		(.253)	
Number of students	301	298		1,207	

Notes: Panel A: Balance table across the treatment (N=301), subsample of control (N=298) and full sample of control (N=1,207) groups. Columns (3) and (5) show the p-value of the difference between treatment and subsample of control, and treatment and full sample of control groups, respectively. District of origin equals 1 if the district where the student is from is the same district as the school district. Mother's and father's education is equal to one if she or he has completed primary education. Standard errors in parenthesis. Panel B: Differential attrition between treatment and control groups. Regression of attrition indicator in endline surveys A, B, and Biology and English grades on the treatment status with strata bins fixed effects

Table 2: Understanding of Wikipedia and Ability to Find Information

	(1)	(2)	(3)	(4)
		Ability to find information		
	Understands what Wikipedia is (index)	Milky way phone test	News quiz	Academic quiz
Treatment	.877***	.183***	.089**	.108**
	(.074)	(.039)	(.042)	(.041)
	p = .000	p = .000	p = .032	p = .010
Units	s.d. of controls	Binary	Binary	Binary
Mean of dependent variable in control	.000	.392	.513	.567
Strata FE	yes	yes	yes	yes
Number of students	549	548	535	538

Notes: Effects on the student understanding about Wikipedia. "Understands what Wikipedia is" refers to the index calculated over correct answers to the following questions: "Can you find information about world news events on Wikipedia?", "Can you find the MSCE results for students from your school on Wikipedia?", "Can you watch movies on Wikipedia?", "Can you communicate with friends on Wikipedia?" and if the Wikipedia app was recognized among seven other apps (not prompted). "Milky way phone test" refers to the test whereby students were asked "How many stars are there in the Milky Way?" and were allowed to consult internet during the survey to find the answer. Registered as a binary outcome if the student was correct within two minutes of search. "News" and "academic quiz" record if the student correctly answers to each quiz question. Questions were student-specific and correct answers were incentivized. The sample is students in the treatment group and in the subsample of the control group with supplementary surveys. Stratification by school, form, above median achievement and past internet use. Robust standard errors in parenthesis. * p<0.10, ** p<0.05, *** p<0.01. Randomization inference p-values based on 10,000 replications in brackets and denoted as "p = ".

Table 3: Is Wikipedia Information Better than the Internet?

	(1)	(2)	(3)	(4)	(5)	(6)
	Wikipedia is better for			How is Wikipedia better		
	Safe sex info	News info	Milky way phone test (opened Wikipedia)	Trustwor-thy	Easy to under-stand	Easy to find
Panel A. Overall effects						
Treatment	.185***	.155***	.253***	.262***	.333***	.247***
	(.043)	(.043)	(.038)	(.039)	(.038)	(.038)
	p = .000	p = .000	p = .000	p = .000	p = .000	p = .000
Panel B. Heterogeneous treatment effects						
Treatment x No past internet use	.213***	.242***	.394***	.309***	.391***	.380***
	(.059)	(.059)	(.050)	(.057)	(.055)	(.054)
	p = .000	p = .000	p = .000	p = .000	p = .000	p = .000
Treatment x Past internet use	.160***	.071	.117***	.218***	.277***	.118**
	(.060)	(.059)	(.053)	(.054)	(.050)	(.051)
	p = .000	p = .223	p = .032	p = .000	p = .000	p = .000
Mean of dependent variable in control	.457	.377	.212	.436	.495	.542
Strata FE	yes	yes	yes	yes	yes	yes
Number of students	536	548	549	549	549	549

Notes: Columns (1) to (3): "Safe sex" refers to the question "What is the best place to find information about safe sex?" Students were asked to rank the following six options: a teacher, books in the school, Wikipedia, internet (other sites), another student, a family member. Coded as one if Wikipedia was ranked higher than internet (other sites). "News" refers to the question "What is the best place to find information about news events?" where, again, students ranked options. "Milky way phone test (opened Wikipedia)"is a binary variable equal to one if student opened the Wikipedia app during the Milky Way phone test. Columns (4) to (6): Multiple choice answers to the question "How is Wikipedia better than other sites on the internet?" Outcomes in column (4) is equal to one if option "Information on Wikipedia is more trustworthy" was chosen. Columns (5) if "It is easier to understand information on Wikipedia"was chosen. Columns (6) if "It is easier to find information on Wikipedia"was chosen. Other alternatives were "There is more information on Wikipedia", "There are more things to do on Wikipedia", and "Don't know". "No past internet use" if student did not relate having used internet at the baseline. The sample is students in the treatment group and in the subsample of the control group with supplementary surveys. Randomization stratified by school, form, above median achievement and past internet use. Robust standard errors in parenthesis. * p<0.10, ** p<0.05, *** p<0.01. Randomization inference p-values based on 10,000 replications denoted as "p = ".

Table 4: Treatment Effects on Exam Scores, Primary Outcomes

	(1)	(2)
	English	Biology
Panel A. Overall effects		
Treatment	.103**	.063
	(.050)	(.047)
	p = .045	p = .194
Panel B. Heterogeneous treatment effects		
Treatment x low achiever	.195***	.143**
	(.076)	(.067)
	p = .001	p = .042
Treatment x high achiever	.003	-.025
	(.062)	(.064)
	p = .966	p = .720
Mean of dependent variable in control	.000	.000
Strata FE	yes	yes
Number of students	1,412	1,406

Notes: Regression results of exam scores from term 3. We include a control for baseline exam score with an indicator for missing baseline score. Randomization stratified by school, form, above median achievement and past internet use. Robust standard errors in parenthesis. * p<0.10, ** p<0.05, *** p<0.01. Randomization inference p-values based on 10,000 replications denoted as "p = ".

Table 5: Treatment Effects on Exam Scores, Other School Subjects

	(1) Science	(2) Human.	(3) Math	(4) Chichewa
Panel A. Overall effects				
Treatment	-.029	-.001	.042	-.066
	(.047)	(.050)	(.044)	(.057)
	p = .520	p = .988	p = .331	p = .239
Panel B. Heterogeneous treatment effects				
Treatment x low achiever	.016	.100	.105*	-.071
	(.070)	(.070)	(.062)	(.079)
	p = .803	p = .142	p = .084	p = .386
Treatment x high achiever	-.076	-.112	-.022	-.060
	(.062)	(.070)	(.061)	(.083)
	p = .213	p = .061	p = .734	p = .438
Mean of dependent variable in control	.000	.000	.000	.000
Strata FE	yes	yes	yes	yes
Number of students	1,370	1,396	1,376	1,406

Notes: Regression results of exam scores from term 3. Science is a summary index of Chemistry and Physics. Humanities is a summary index of Agriculture, Geography, History, Life Skills and Social Science. We include a control for baseline exam score with an indicator for missing baseline score. Randomization stratified by school, form, above median achievement and past internet use. Robust standard errors in parenthesis. * p<0.10, ** p<0.05, *** p<0.01. Randomization inference p-values based on 10,000 replications denoted as "p = ".

Table 6: Time Use, Participation in Class and Career Goals

	(1)	(2)	(3)	(4)	(5)	(6)
	Time Use (hours per day)				Participation in class (per day)	Career goal change
	Study	Study	Recreational	Awake		
Treatment	-.029		-.286***	.007	-.025	.031
	(.071)		(.078)	(.086)	(.159)	(.042)
	p = .718		p = .005	p = .931	p = .874	p = .564
Treatment x low achiever		-.038				
		(.098)				
		p = .744				
Treatment x high achiever		-.019				
		(.104)				
		p = .867				
Mean of dependent variable in control	1.937	1.937	.967	17.032	2.891	.549
Strata FE	yes	yes	yes	yes	yes	yes
Day-of-the-week FE	yes	yes	yes	yes	yes	no
Number of students	1,402	1,402	1,396	1,398	1,402	549

Notes: Treatment effects on time use and participation in class. "Study", "Recreational" and "Awake" refers to the time spent on study, recreational activities and not sleeping, respectively, and averaged over the three days prior to the interview. Study time is the sum of the answers to the questions "How much time did you study alone?" and "How much time did you study with others?". Recreational time is the sum of the answers to the questions "How much time did you hang out with friends?", "(...) in a school club?", "(...) in religious activities?", "(...) sports activities?" and "(...) other activities?". Awake time is the duration between waking up and going to sleep at night. Time Use specifications (1)-(4) include baseline controls. Lee (2009) bounds are conducted in Columns (1), (3) and (4) to assess the robustness with respect to differential attrition. For recreational time, the bounds are [-.361, -.222] and both are statistically significant at the 5% level; for study time, the bounds are [-.149,.020]; awake time, [-.104,.090]. Participation in class counts the number of times that students responded that they raised their hands in class to ask a question, also averaged over the three days prior to the survey. Column (6) shows change in career goal between baseline and endline surveys. Outcome variable considers both changes in career category and precision in the career choice, e.g. "doctor" to "surgeon" is considered a change. Robust standard errors in parenthesis. * p<0.10, ** p<0.05, *** p<0.01. Randomization inference p-values based on 10,000 replications in brackets and denoted as "p = ".

Searching for Answers: The Impact of Student Access to Wikipedia

Laura Derksen, Catherine Michaud Leclerc and Pedro CL Souza

Appendices
For Online Publication

Appendix: Wikipedia Category Tree Classifications

We use the Wikipedia category tree to classify pages into broad yet meaningful topics. Wikipedia has a user-generated and user-maintained category tree. The tree has 39 top-level categories which we adopt as topic classifications. Each top-level category branches into one or more subcategories which, in turn, may contain both pages and narrower subcategories. We trace each page visited by a student to one top-level category.[46]

We manually map the Malawian secondary school syllabus to specific Wikipedia pages and narrow Wikipedia subcategories from the category tree described in 3.2. For example, the subcategory for "Circulatory System"

[46]The full list of top-level categories can be found at https://en.wikipedia.org/wiki/Category:Main_topic_classifications. For more information on the tree structure, see https://en.wikipedia.org/wiki/Wikipedia:Categorization#Topic_categories. A Wikipedia page typically belongs to more than one narrow subcategory. For example, the page on Barack Obama is associated to over 40 subcategories such as "Presidents of the United States", "University of Chicago Law School faculty" and "Grammy Award winners". By following different paths through the Wikipedia category tree, we might categorize it under more than one top-level category. In order to generate a unique topic classification, we consider every path in the Wikipedia category tree that reaches the top of the tree in at most six steps. We then select the top-level category that appears most often at the top of these paths. For example, the topic we assign to Barack Obama's Wikipedia page is "People".

matches a topic in the Biology syllabus, and we include it in our list of syllabus subcategories. We do not include broad categories such as "Biology" or "History". If a Wikipedia page exactly matches a topic for a particular school subject, or belongs to a syllabus subcategory, we classify it as directly related to that subject syllabus.

Appendix: Spillovers

While control students did not gain direct access to Wikipedia or the internet during the experiment, our primary outcomes may still be subject to spillovers. For example, a student's language skills or Biology knowledge may improve if they study with students who have benefited from Wikipedia access.

To test for spillovers on academic performance, we use baseline social network data. At baseline, we ask every student to name the friends they study with. We say there is a study link between two students if both students name each other. While social networks change over the course of a school year, study friends are in part determined by school-level decisions such as classroom, dormitory, and formal study groups. These formal study groups are assigned by teachers at the start of the school year, and meet regularly. Therefore, study friends are more likely to remain constant over the school year. Study friends are also most likely to benefit from spillovers that impact academic performance.

Similar to Miguel and Kremer (2004), we estimate equation 2 and the heterogeneous treatment effects by baseline achievement from Section 5, adding controls for the number of study friends total and the number of study friends in the treatment group. We also interact own treatment status with the number of treated study friends, as spillovers may exist only between treatment and control group students.

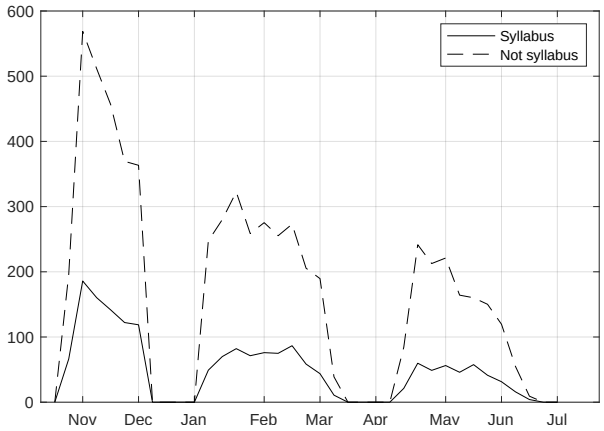
We find positive, insignificant spillovers from the treatment group to the control group (see Table A3, which reports results for term 3 English and Biology). In this specification, our estimated effect sizes are slightly larger (0.18σ

and 0.15σ for English and Biology, respectively), but have larger standard errors. Effects for low achievers are also larger, and remain significant.

This specification may not capture all spillovers, as spillovers may exist beyond study friends at baseline. However, it suggests that spillovers are likely to be positive from treatment to control students, and that our estimates slightly underestimate the true impact of the intervention.

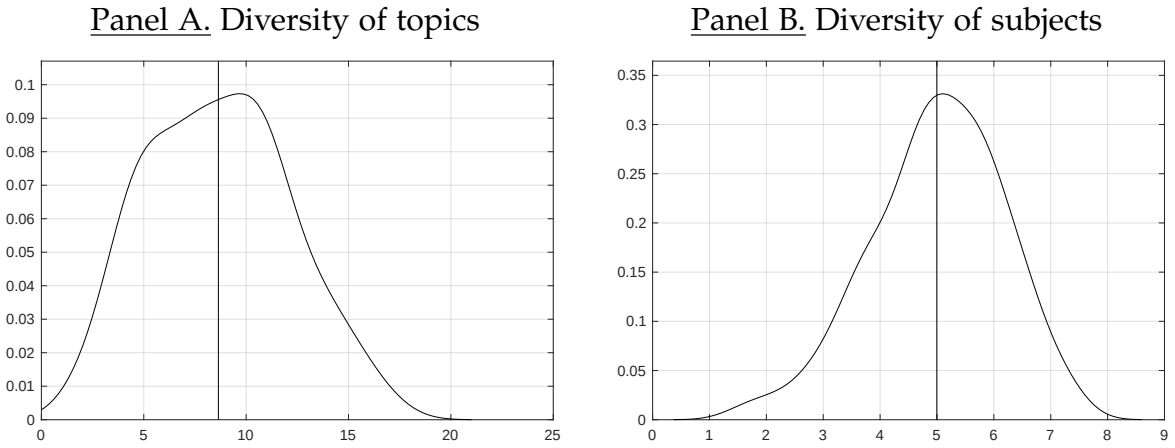
Appendix: Figures and Tables

Figure A1: Weekly Hours Spent on Browsing in Total and on School Subjects Over Time



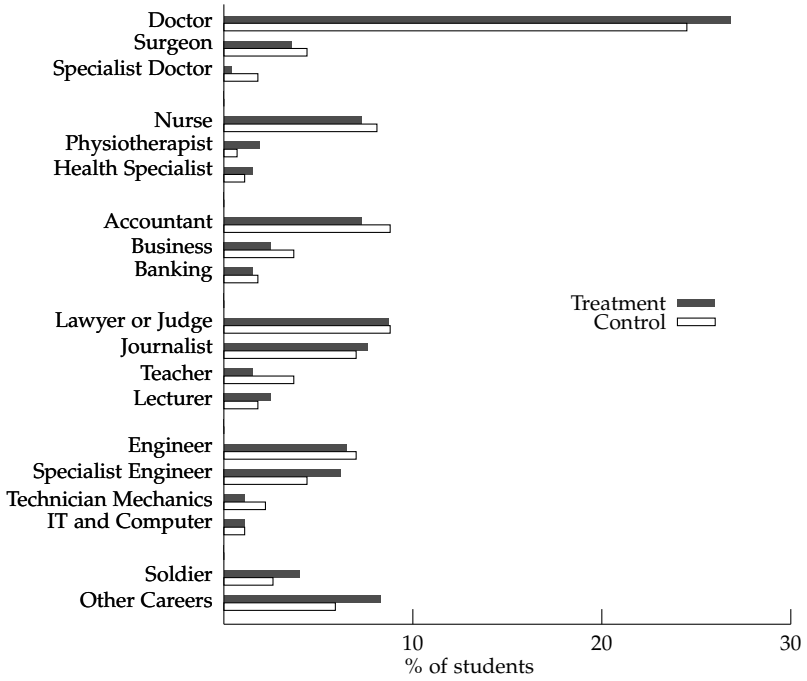
Notes: Weekly browsing hours on syllabus- and non syllabus-related Wikipedia pages. The digital library was open for 20-22 weeks, from November 2017 to June 2018, excluding Christmas and Easter vacations.

Figure A2: Diversity of browsing pattern across students



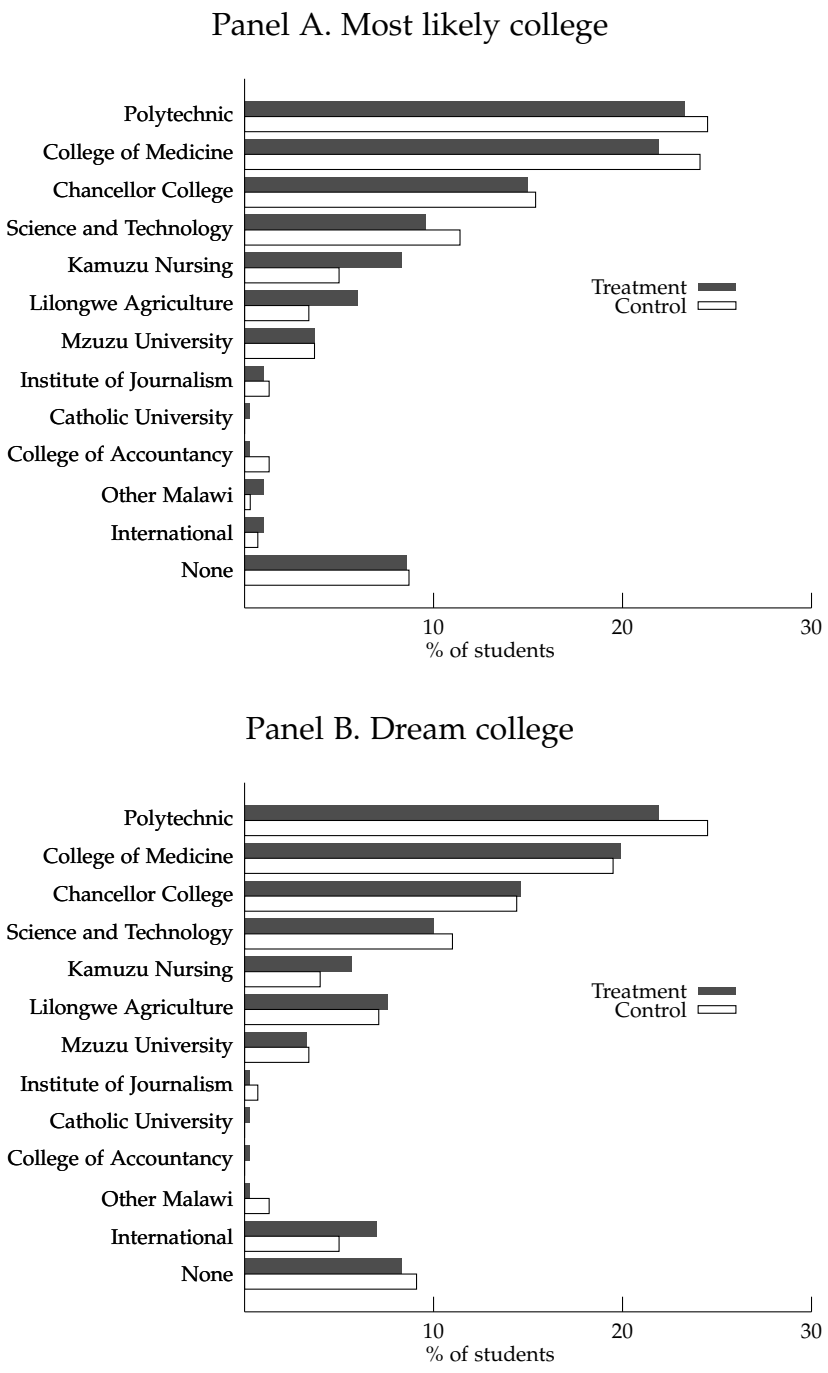
Notes: Panel A: Diversity of browsing topics across students. Adaptation of the Herfindahl index, computed as $d_i = 1/\sum_j s_{ij}^2$ where s_{ij} is the share of time that student i spends in topic j, throughout the duration of the experiment. Larger numbers represent broader diversity of topics. Dashed line is the average (8.64 topics). Panel B: Diversity of browsing hours per school subject across students, aggregated over the course of one academic year. Similar adaptation of the Herfindahl index. Dashed line is the average (5.00 subjects). The digital library was open for 20-22 weeks, from November 2017 to June 2018, excluding Christmas and Easter vacations.

Figure A3: Career Plans at Endline



Notes: Frequency of career choices, as registered in the endline survey, across 301 treated and 298 subsample control students.

Figure A4: College Choice at Endline



Notes: Frequency of most likely and dream college in Panels A and B, respectively, as registered in the endline survey, across 301 treated and 298 subsample control students.

Table A1: Mobile Phone and Internet Use

	% population within network coverage	Mobile subscriptions per 100 inh.		Internet bundle price as % of income		% population with internet use		GDP per capita
	2006	2014	2017	2007	2017	2007	2017	2017
Malawi	93.1	7.6	41.7	45.2	18.0	1.0	13.8	$339
Zambia	44.9	20.7	78.6	15.2	12.6	4.9	27.9	$1,332
Mozambique	42.1	13.9	40.0	13.3	7.2	.9	20.8	$4412
South Africa	99.8	84.8	156.0	1.3	1.2	8.1	56.2	$6,340
LDC	–	15.1	68.6	21.4	14.8	1.9	17.8	$1,093
Developing	–	39.1	99.0	9.0	6.3	11.8	42.3	$5,229
Developed	–	102.0	127.0	.9	.8	59.1	79.5	$42,346
World	–	50.6	103.6	6.5	4.6	20.5	48.6	$10,749

Notes: "% population with GMS coverage" from Buys et al. (2009). Remaining data, excluding GDP per capita, from the International Telecommunications Unit. Classification of "Least Developed Country" (LDC), "developing" and "developed countries" also drawn from the International Telecommunications Unit. "Internet bundle price as % of income" is the proportion of the average national income to purchase 1GB of a data bundle, monthly. GDP per capita obtained from the World Bank. Definition of LDC in the last column uses the United Nations' classification. Average GDP per capita of developing (developed) countries approximated by the average GDP per capita of middle (upper) income countries. Income in current US$.

Table A2: Balance Table on Stratification Variables

	(1) Treatment	(2) Control (subsample)	(3) p-value	(4) Control (full)	(5) p-value
School 1	.272	.262	.768	.261	.689
	(.446)	(.440)		(.439)	
School 2	.262	.289	.475	.280	.537
	(.441)	(.454)		(.449)	
School 3	.236	.232	.900	.229	.792
	(.425)	(.423)		(.420)	
School 4	.229	.218	.745	.230	.968
	(.421)	(.414)		(.421)	
Form 2	.342	.346	.929	.342	.999
	(.475)	(.476)		(.475)	
Form 3	.332	.329	.930	.328	.892
	(.472)	(.471)		(.470)	
Form 4	.326	.326	.998	.330	.891
	(.469)	(.469)		(.470)	
Above median Bio. and Eng. exam scores	.468	.473	.908	.472	.906
	(.500)	(.500)		(.499)	
Past internet use	.502	.500	.968	.505	.908
	(.501)	(.501)		(.500)	
Number of students	301	298		1,207	

Notes: Balance table across the treatment (N=301), subsample of control (N=298) and full sample of control (N=1,207) groups. Columns (3) and (5) show the p-value of the difference between treatment and subsample of control, and treatment and full sample of control groups, respectively. "Above median Biology and English exam scores" computed on the previous school year. "Past internet use" is a dummy variable which indicates if the student had any exposure to internet prior to the experiment. Standard errors in parenthesis.

Table A3: Treatment Effects on Exam Scores with Spillover Controls

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Biology				English			
Treatment	.062	.148			.102**	.180		
	(.047)	(.091)			(.050)	(.098)		
Treatment x low achiever			.149**	.214**			.196***	.253**
			(.066)	(.100)			(.075)	(.108)
Treatment x high achiever			-.035	-.040			-.001	.063
			(.064)	(.104)			(.062)	(.110)
Treated Study Friends	-.010	-.029	-.010	-.026	-.004	-.021	-.004	-.018
	(.013)	(.021)	(.013)	(.021)	(.014)	(.023)	(.014)	(.023)
Study Friends	.016***	.016***	.016***	.016***	.006	.006	.006	.006
	(.004)	(.004)	(.004)	(.004)	(.005)	(.005)	(.005)	(.005)
Control x Treated Study Friends		.023		.019		.021		.016
		(.021)		(.021)		(.023)		(.023)
Mean of dependent variable in control	.000	.000	.000	.000	.000	.000	.000	.000
Strata FE	yes	yes	yes	yes	yes	yes	yes	yes
Number of students	1,406	1,406	1,406	1,406	1,412	1,412	1,412	1,412

Notes: Regression results of test scores for terms 3. Columns (1) and (4) are the spillover specifications corresponding to Equation (5). Explanatory variables are the treatment indicator, number of treated study friends and number of study friends, all interacted with the term dummy, in addition to an indicator for missing baseline score. Columns (2) and (6) add control for the number of treated study friends interacted with control student indicator. Columns (3) and (7) are the spillover specification corresponding to Equation (6). Explanatory variables are the treatment indicator interacted with high and low achiever indicator, number of treated study friends and number of study friends, all interacted with the term dummy, in addition to an indicator for missing baseline score. Columns (4) and (8) add control for the number of treated study friends interacted with control student indicator. Study friend network collected at the baseline. A link is considered to be present if students nominate each other during the three-day recall based on the question "With whom did you study with [yesterday]?"Ṙandomization stratified by school, form, above median achievement and past internet use. Robust standard errors in parenthesis. * p<0.10, ** p<0.05, *** p<0.01.

Searching for Answers: The Impact of Student Access to Wikipedia

Laura Derksen, Catherine Michaud Leclerc and Pedro CL Souza

Supplementary Materials For Online Publication

Examples from English Examinations

In this section, we provide examples of the multiple choice questions and composition questions that students have to take for their English classes. The questions were provided by the school administration.

Figure 1: Sample of Multiple Choice Questions: English Examinations

4. We travelled_____train from Salima to Lilongwe
   A. on
   B. in
   C. through
   D. by
5. Everyone was surprised_____the bad weather.
   A. with
   B. by
   C. at
   D. for
6. The storm had_____when we started our journey.
   A. died away
   B. died out
   C. died off
   D. died down

Sample of Composition Question: English Examinations

" – Answer one question only
– Spend the first 10 minutes reading the questions and planning your answers to the question chosen. Planning may include writing rough notes. Cross out your rough notes before you hand in your Answer Book.
– Marks will be awarded for layout, language, content and creativity. Candidates will be penalized for committing mechanical errors and writing answers that are short and /or off-point.
– You are expected to write between 350 and 500 words for the question you have chosen.

EITHER
1. Write an original short story entitled 'The Imposter' (40 marks)

2. Imagine that the area where you live was flooded. The floods destroyed homes and property. Write a report to the District Commissioner informing him or her of the disaster. (40 marks)"

Classroom Introduction to the Project

- We are working for the University of Toronto in Canada for a research project
- The research project will take place in this school for the entire school year in which some students in Forms 2-4 will have access to a digital library with phones with access to an online encyclopedia.
- First, we would like to survey every student in Forms 2-4.
- The survey is not too long – about 10-20 minutes
- After the survey is finished, we are going to select some students for the mobile phone program
- The students are going to be selected RANDOMLY – it is not the best students. Every student in Forms 2-4 has the chance to be selected. [Make sure this is extremely clear]
- We cannot select every student, only a few students will be selected
- During the year, those selected students will be able to take part in a digital library program
- A digital library will be set up in [classroom]
- There will be a number of mobile phones with access to an online encyclopedia
- Students taking part in the program will be able to search online for information about their studies and other information [see examples below]

- If you are not selected for the program, you are free to ask selected students to search for something or to explain what they have learned
- If you are selected, you are free to take part or to refuse, you are not obligated to use the digital library
- We will continue to ask some students to answer short surveys throughout the year – these will include some selected students and other students
- Any questions?
- Looking forward to seeing you again when we will be conducting the survey

About Wikipedia

A lot of information can be accessed on Wikipedia. This includes information about academics, health, politics, world news, sports and entertainment.

For example, suppose your Biology teacher says that next week you will start the topic of photosynthesis. If you search Wikipedia, you will find a detailed explanation of the process of photosynthesis, with equations and illustrations.

I will give you another example. Suppose you did not understand the different types of soil you discussed in agriculture class. You can use Wikipedia to find out more about the topic, including the definition of soil and the various types of soil. Wikipedia includes information about soil fertility, soil formation and the different functions of the soil.

In Wikipedia, you can find information about almost any topic from your studies. For example, you can find information about chemical reactions and the periodic table in chemistry, matter in physics and volcanos in geography. You can even review different rules you learn in mathematics such as the rules for exponents. You can find information about local and international authors.

If you are thinking about what you want to do after secondary school, you can search Universities in Malawi and you will find a list of all colleges, public and private universities in Malawi. You can even look into the careers you may be interested in pursuing.

As we said, there is information just about everything on Wikipedia. If you want to know more about menstruation, birth control or pregnancy, you will find it in Wikipedia. Wikipedia talks also about different diseases such as malaria, Ebola and HIV. You can find information about the causes, symptoms and prevention.

You can find information about local politics and international news. You can find information about sports stars like Lionel Messi, and celebrities like Jay-Z or Nicki Minaj.

We think this project will help you a lot with your studies. Even if you are not selected for the program, you can ask your friends to search for information on a topic from class. If you are selected, you can share what you learn with your classmates.

Digital library Induction

Instructions for Digital Librarians

- The induction should be done in small groups – enough so that each student can use one phone. Only for selected students
- Explain the digital library itself; Opening hours
- Explain Wikipedia. What it is, what kind of information you can find
- Explain privacy. Anonymous, you are free to search anything.
- You can only use Wikipedia. Everything else is blocked
- Practice together. Give several examples of things to search for (e.g. photosynthesis and Malawi).
- Show how to solve common problems. How to get back to search page (home three dots OR icon). Show what happens if they try to click on external links or restart the phone

How to Use the Digital Library

- There are 12 phones in the digital library
- Sign in with the librarian
- If all phones are in use, join the waiting list or come back later
- If there is a waiting list, students are restricted to 30 minutes (35-40 minutes when the network is not good)
- Use the phone within the library
- Do no try to tamper with the phones
- There are no backup phones so if one breaks or goes missing there will be fewer phones to use
- Privacy. Your searches are anonymous – no one can see what you personally searched for (not the researchers, not the field team, not the teachers). This is a very important point – make sure they students can explain it back.

Digital Library Rules

1. Only selected students can use the digital library
2. When you arrive, sign in with the digital librarian.
3. The phones should be used one by one (not in pairs)
4. Take care not to damage or tamper with the phone
5. Do not try to access other websites than Wikipedia
6. When you are done, return the phone to the digital librarian
7. Do not hand the phone to any other student

If you break the rules you will be suspended or removed from the program

Table 1: Classification of News Events in 2017-18 (I/II)

Date	Description	Eliminated	Not specific	No Searches	Africa	Geopolitics	Keywords
11/2/17	New species of Orangutan	no			no	no	Orangutan
11/3/17	ISIL defeated in Syria	no			no	yes	Islamic State
11/5/17	Appleby scandal	yes	no	yes	no	no	
11/5/17	Sutherland Springs shooting	no			no	no	Sutherland Springs
11/12/17	Earthquake in Iran and Iraq	no			no	no	2017 Iran-Iraq earthquake
11/15/17	Robert Mugabe arrested	no			yes	yes	Robert Mugabe
11/15/17	da Vinci auction	no			no	no	Leonardo da Vinci
11/15/17	ARA San Juan missing	no			no	no	ARA San Juan
11/20/17	Oumuamua asteroid detected	no			no	no	Oumuamua
11/22/17	Mladic found guilty	no			no	yes	Mladic
11/24/17	Mosque attack in Egypt	no			yes	no	2017 Sinai Mosque Attack
12/5/17	Russia banned from Winter Olympics	no			no	no	Doping in Russia, Russia at the Olympics
12/6/17	US recognizes Jerusalem as Israeli capital	no			no	yes	Consulate General of the United States Jerusalem, United States recognition of Jerusalem as Israeli capital, Status of Jerusalem
12/9/17	Iraq liberated from ISIS	no			no	yes	ISIS
12/14/17	Disney acquires 21st Century Fox	no			no	no	Proposed acquisition of 21st Century Fox by Disney, The Walt Disney Company
12/22/17	UN imposes sanctions to North Korea	no			no	yes	North Korea
12/24/17	Guatemala recognizes Jerusalem as Israeli capital	yes	yes		no	yes	
1/13/18	Killing of Mehsud in Pakistan	yes	no	yes	no	yes	
1/20/18	Turkey invades northern Syria	yes	no	yes	no	yes	
1/20/18	US Federal government shutdown	no			no	yes	US government shutdown
1/24/18	China announces cloning of monkeys	no			no	no	Zhong Zhong and Hua Hua
1/31/18	Total lunar eclipse	no			no	no	January 2018 lunar eclipse, Lunar eclipse
2/6/18	Falcon Heavy launch	no			no	no	Falcon Heavy, SpaceX
2/9/18	Winter Olympics starts	no			no	no	Winter Olympics
2/10/18	First female archbishop nominated	yes	yes		no	no	
2/11/18	Saratov Airlines crash in Russia	no			no	no	Saratov
2/14/18	Jacob Zuma resigns	no			yes	yes	Jacob Zuma
2/14/18	Majory school shooting	no			no	no	Stoneman Douglas
2/18/18	Iran Aseman Airlines crash	no			no	no	Iran Aseman
3/4/18	Skripal poisoning	no			no	yes	Skripal
3/6/18	Russian Air Force crash	yes	yes		no	no	
3/9/18	Winter paralympics start	no			no	no	Winter Paralympics
3/9/18	Trump accepts meeting with Kim Jong-un	yes	yes		no	yes	
3/11/18	Jinping named President for Life in China	no			no	yes	President for Life, Xi Jinping

Notes: All major newsworthy events as reported in `https://en.wikipedia.org/wiki/2017` and `https://en.wikipedia.org/wiki/2018` during the experiment. "Eliminated" marks the events that were not considered either because no specific Wikipedia page could be associated (column "not specific") or because we found no evidence of browsing (column "no searches"). "Africa"and "Geopolitics" refer to whether the events were considered to be African or of geopolitical nature. Elements in column "keywords" are used to match the Wikipedia pages with terms associated to them.

Table 2: Classification of News Events in 2017-18 (II/II)

Date	Description	Eliminated	Not specific	No Searches	Africa	Geopolitics	Keywords
3/12/18	US-Bangla Airlines crash in Nepal	yes	no	yes	no	no	
3/14/18	School walkout in response to shootings in the US	yes	yes		no	yes	
3/18/18	Putin re-elected president	no			no	yes	Vladimir Putin, Russian Elections 2018
3/19/18	White rhino declared extinct	no			no	no	White rhinoceros
3/23/18	Carcassone terrorist attack	yes	no	yes	no	no	
3/24/18	Demonstrations against gun violence	yes	no	yes	no	yes	
3/25/18	Quantas launches Perth-London flight	yes	yes		no	no	
3/25/18	Kemerovo fire	yes	no	yes	no	no	
3/26/18	Russian diplomats expelled in the wake of Skripal poisoning	yes	yes		no	yes	
3/28/18	Kim Jong-un meets Xi Jinping	no			no	yes	Kim-Xi meetings
3/28/18	Fire in Valencia, Venezuela	yes	no	yes	no	no	
4/4/18	Commonwealth games start	no			no	no	Commonwealth games
4/5/18	Lula arrested	yes	yes		no	yes	
4/6/18	Humboldt Broncos crash	yes	no	yes	no	no	
4/8/18	Sarin attack in Douma, Syria	no			no	yes	Douma, Use of chemical weapons in the Syrian Civil War
4/11/18	Algerian Air Force crash	no			yes	no	2018 Algerian Air Force Il-76 crash
4/14/18	Syrian bases bombed by US	no			no	yes	American-led intervention in the Syrian Civil War
4/18/18	Nicaragua protests	no			no	yes	2018 Nicaraguan protests
4/18/18	Movie theathers open in Saudi Arabia	yes	yes		no	no	
4/18/18	NASA TESS satellite launched	yes	no	yes	no	no	
4/19/18	Diaz-Canel sworn President of Cuba	yes	no	yes	no	yes	
4/19/18	Swaziland changes name to Eswatini	yes	no	yes	yes	yes	
4/23/18	Toronto van attack	no			no	no	Toronto Attack, Toronto van attack
4/27/18	Kim Jong-un meets Moon Jae-in in the DMZ	yes	yes		no	yes	
5/3/18	ETA announces dissolution	yes	no	yes	no	yes	
5/3/18	Volcano Puna erupts	yes	no	yes	no	no	
5/5/18	Insight probe launched	no			no	no	InSight
5/8/18	Trump withdrawals from Iranian nuclear agreement	no			no	yes	Joint Comprehensive Plan of Action, Negotiations leading to the Joint Comprehensive Plan of Action
5/8/18	Eurovision contest starts	no			no	no	Eurovision
5/9/18	Pakaran Harapan coallition wins majority in Malaysia	yes	no	yes	no	yes	
5/16/18	Agong pardons Ibrahim in Malaysia	yes	yes		no	yes	
5/18/18	Cubana airline crash	no			no	no	Cubana de Aviacion Flight 972
5/19/18	Prince Harry and Meghan Markle wedding	no			no	no	Prince Harry, Meghan Markle
5/20/18	Venezuelan elections	yes	no	yes	no	yes	
5/24/18	Punggye-ri nuclear test site destroyed in North Korea	yes	no	yes	no	yes	
5/25/18	EU data protection regulation goes into effect	no			no	no	General Data Protection Regulation
5/25/18	Abortion referendum in Ireland	no			no	yes	Abortion in the Republic of Ireland

Notes: as above.

Academic Questions – Sample

Biology Questions
A spirochaete is a type of...
Which of the following bacteria is gram-negative?
Which of the following bacteria is gram-positive?
How do fungi acquire their food?
Penicillin is derived from penicillum, a type of
Cholera is a
Which of the following is an example of an endocrine gland?
Which of the following is both an endocrine and an exocrine gland?
Where is insulin produced?
History Questions
World War I began in which year?
Adolf Hitler was born in which country?
John F. Kennedy was assassinated in
Who fought in the war of 1812?
Which general famously stated "I shall return"?
American involvement in the Korean War took place in which decade?
The Battle of Hastings in 1066 was fought in which country?
The Magna Carta was published by the King of which country?
Who first successfully developed the printing press?
Geography Questions
Which of the following cities is the capital of Argentina?
Which ocean lies on the east coast of the United States?
How many Great Lakes are there in the United States/Canada?
Which is the world's highest mountain?
Which is the longest river in the World?
Which is the biggest desert in the World?
Which of this cities is not in Europe?
Which of the following cities is the capital of Netherlands?
Which of this is the largest city in Africa?
What is the capital of Turkey?