ELSEVIER
www.elsevier.com/locate/worlddev

CrossMark

World Development Vol. 64, pp. 766–780, 2014

http://dx.doi.org/10.1016/j.worlddev.2014.07.015

Text Messaging and its Impacts on the Health and Education of the Poor: Evidence from a Field Experiment in Rural China

DI MO [a,b,c], RENFU LUO [c], CHENGFANG LIU [c], HUIPING ZHANG [d], LINXIU ZHANG [c], ALEXIS MEDINA [a] and SCOTT ROZELLE [a,*]

[a] *Stanford University, USA*
[b] *LICOS Centre for Insitutions and Economic Performance, University of Leuven, Belgium*
[c] *Center for Chinese Agricultural Policy, Chinese Academy of Sciences, China*
[d] *Ningxia University, China*

Summary. — There is little evidence showing whether health information transmitted via text messages can change health and educational outcomes. We conducted a randomized field experiment involving 900 primary students in rural China to study whether a health education campaign conducted via text message could affect caregiver knowledge or student outcomes. When caregivers received both weekly health messages and monthly quiz questions (testing retention of the information conveyed in the weekly messages), caregiver knowledge improved and students experienced gains in both health and academic performance. When caregivers received weekly health messages only, there was no impact on caregiver knowledge or student outcomes.

Key words — health education, text messaging, caregiver knowledge, student health, academic performance

1. INTRODUCTION

Despite China's achievement in poverty reduction in recent decades (Huang & Rozelle, 2008), millions of children in poor areas of rural China still suffer from undernutrition, a condition that can adversely affect their future development (Zhang *et al.*, 2013). Studies using government data have reported high anemia rates among children in rural areas in China (Chen *et al.*, 2005). Recent research by Luo *et al.* (2011), conducted among 12,768 students aged 8–12 years, shows that the prevalence of anemia across six provinces in rural China ranges from 25% to 51%. To the extent that these data are representative of anemia in China's poor provinces, this means that more than twenty million children in China's poor, rural areas are severely undernourished. Many studies have shown that anemia adversely affects the cognitive development and academic performance of children at school (Lozoff, Jimenez, Hagen, Mollen, & Wolf, 2000; Luo, Shi, Zhang, Liu, *et al.*, 2012). Anemic children are also more likely to have lower incomes (and lower quality of life) when they grow up (Bobonis, Miguel, & Puri-Sharma, 2006; Halterman, Kaczorowski, Aligne, Auinger, & Szilagyi, 2001).

Despite this high prevalence and the proven consequences of childhood anemia, the problem is often misunderstood and/or neglected by educators and parents (Luo *et al.*, 2011). Luo *et al.* (2010) demonstrate that over 90% of rural primary school principals—almost certainly among the most educated persons in China's rural communities—either had no knowledge of anemia or did not think anemia was a problem among their students. Back in students' homes (in rural villages), parents have also been reported to lack knowledge about anemia (Shi *et al.*, 2012). Learning about anemia (or being aware of it), absent formal health education, or in-hospital blood testing, is difficult, given the lack of specific symptoms, the difficulty of observing the micronutrient content of individual diets, and the lagged relationship between dietary change and symptoms (Luo, Shi, Zhang, Zhang, *et al.*, 2012). It is therefore unsurprising that schools and households fail to take action to fight anemia.

Although there may be a role for the government to play in solving the problem, attempts to do so have run into implementation problems in the past. One example is to address anemia at school by improving school meals. In rural China, a large proportion of students eat at least some of their meals at school (Kim, Haines, Siega-Riz, & Popkin, 2003). Because school meals are mainly starch-based, with little meat, vegetables or fruit, iron deficiency anemia among students is common (FAO, 1999; Luo, Shi, Zhang, Liu, Rozelle, & Sharbono, 2009). In order to improve the nutritional content of student diets, in 2012 China's central government allocated 16 billion yuan (2.5 billion USD) to launch the Healthy Lunch Program, which is supposed to provide healthy school lunches to poor rural students (NBC, 2012). According to the plan, the government provides three RMB per student per day to fund a healthy lunch, which is supposed to contain sufficient meat, eggs, and milk to give each child a balanced meal. However, it has been reported that the meals provided by the schools are insufficient either in calories or in micronutrients (China Youth Daily, 2012). Our research team conducted a survey among 300 rural primary schools in three provinces in

* We would like to thank the Ford Foundation and the Stanford University Global Underdevelopment Action Fund for their generous support of this project. We would also like to extend a thank you to Paul Cavey for his financial assistance and other kind support. We are appreciative of the careful and diligent research assistance provided by Rong Li, whose efforts proved essential in the initial design and launch stages of this project. Finally, we are especially grateful to Madhouse, a mobile advertising firm headquartered in Shanghai, for their invaluable help in providing technical and logistical support for all of the text messages sent in this project. Final revision accepted: July 15, 2014.

northwest China and found that only 26% of the schools regularly included meat, eggs, or milk in school meals (Sylvia, Luo, Zhang, Shi, Medina, & Rozelle, 2013).

Other studies in developing countries suggest that nutrition problems may also be attacked at home (Zlotkin, Christofides, Hyder, Schauer, Tondeur, & Sharieff, 2004). Families are likely to be more motivated than schools to improve child nutrition. Hence, in theory, health education that targets caregivers has the potential to increase health/nutrition knowledge and to instill healthy behaviors that can help improve children's nutritional intake (Black *et al.*, 2004; Rivera *et al.*, 2001).

If health education programs targeting caregivers could be shown to be effective, there would be considerable scope for such programs to help address China's anemia problem. Around 85% of the school-age children in rural China live at home with parents or grandparents (Luo *et al.*, 2009). Moreover, although rural households in Western China are still poor, most households have access to sufficient financial resources (from government subsidy programs, rising wages from off-farm income, and rising agricultural prices) to provide their child with a balanced diet. The primary obstacle therefore seems to be ignorance: studies show that many rural caregivers are unaware that feeding their child a more nutritious diet would have a positive effect on future educational attainment and other outcomes (Chen & Li, 2000; Du, Mroz, Zhai, & Popkin, 2004; Popkin, Conde, Hou, & Monteiro, 2006).

Unfortunately, previous attempts to persuade parents in rural China to address anemia have been unsuccessful. Luo, Shi, Zhang, Zhang, *et al.* (2012) field-tested three different approaches to conducting health education campaigns in rural primary schools in China. In one experiment, the authors sent a letter home to parents explaining anemia, revealing each child's anemia status, and outlining anemia prevention and treatment strategies. In a second experiment, the research team invited parents to a face-to-face meeting at their child's school to deliver essentially the same information as was contained in the letter home, this time incorporating the use of multimedia materials such as colorful pamphlets and an educational video. In a third experiment, multiple face-to-face meetings were held at the school—one meeting per academic semester. Unfortunately, across all three experiments, there was no impact of the educational campaigns on student anemia status.

In light of the failure of more traditional educational/training campaigns, the use of mobile technology to educate households becomes an attractive option. Over the past decade, mobile technology has been promoted as a potentially powerful health promotion tool in developing countries (Cole-Lewis & Kershaw, 2010). Mobile technology is now more widely available than ever before; up to two-thirds of the world's population own mobile phones and use them in daily communication (Union IT, 2010). In China, it has also been shown that a large percentage of households (around 90%) have mobile phones, even in rural areas (China Internet Network Information Center, 2012). Compared with traditional ways of delivering information, text messaging—defined as sending an electronic message over cellular phone networks—has been promoted as being both more accessible and more efficient (Fjeldsoe, Marshall, & Miller, 2009). It is already being used to address a myriad of health issues in the developing world, ranging from malaria (Zurovac, Sudoi, Akhwale, Ndiritu, Harner, & Rowe, 2011), to HIV (Pop-Eleches, Thirumurthy, Habyarimana, Zivin, Goldstein, & de Walque, 2011), to tuberculosis (Denkinger, Grenier, Stratis, Akkihal, Pant-Pai, & Pai, 2013).

Despite the recent popularity of health interventions using text messaging, there is little rigorous evidence showing whether health information transmitted via text messages is enough to change health behaviors and outcomes (Cole-Lewis & Kershaw, 2010). Of the handful of studies that aim to provide evidence on the effectiveness of text messaging, many of them fail to include a control group (Fjeldsoe *et al.*, 2009). Even among the more rigorous evaluations, the intervention is often bundled with other inputs, such as web-based information, assistance hotlines, office visits, or free/subsidized medicine (Franklin, Waller, Pagliari, & Greene, 2006; Hurling *et al.*, 2007; Ostojic, Cvoriscec, Ostojic, Reznikoff, Stipic-Markovic, & Tudjman, 2005; Rodgers *et al.*, 2005). Only a few studies have attempted to isolate the impact of text messaging on health behavior or other outcomes (Free *et al.*, 2013). Almost all of the existing evaluations have been "reminder" studies that used text messages as reminders for disease control or health management. For example, text messages were sent as reminders to regulate weight (Haapala, Barengo, Biggs, Surakka, & Manninen, 2009), to take insulin to manage diabetes (Franklin *et al.*, 2006), to wear condoms (Delamere *et al.*, 2006) or to stick to an exercise plan (Prestwich, Perugini, & Hurling, 2009). To our knowledge, no rigorous evaluation has ever attempted to investigate the impact of a health education campaign that uses text messaging on the health and education of project participants (especially when the participants are known to have zero or minimal health/nutrition knowledge).

The overall goal of this study is to test whether a health education campaign that uses text messaging to transmit information is effective at raising parental awareness and understanding of child nutrition, promoting child health, and improving children's educational outcomes. To meet this overall goal, we have four specific objectives. First, we report on the ability of the text messages to reach the caregivers of children in poor rural areas. Second, we conduct econometric analyses to examine the impacts of two different text messaging strategies on a set of health and educational outcomes, including parental health/nutrition knowledge, student hemoglobin levels (an indicator of anemia status), overall child health, and children's in-school behavior and standardized math test scores. Third, in order to better understand *how* text messaging may impact health and educational outcomes, we conduct econometric analyses to examine whether text messaging has any impact on household purchases of food and/or nutritional supplements. Finally, we test the heterogeneous effects of text messaging by type of message recipient in order to answer the question of whether the impact of the text messaging program varies if the child's mother receives the messages versus others in the same household.

(a) *Text messaging in health education*

In theory, health campaigns that use text messaging have a number of advantages over traditional campaigns. Text messaging makes information delivery convenient and efficient (Fjeldsoe *et al.*, 2009). Messages can be accessed wherever the recipient is located and the messages can be read whenever it is personally convenient to the recipient (Lim, Hocking, Hellard, & Aitken, 2008). Messages can be easily repeated to reinforce the information (Fry & Neff, 2009). Repetition of the information can also act indirectly as a reminder for behavioral change (Malow, Kershaw, Sipsma, Rosenberg, & Dévieux, 2007).

Another advantage of health campaigns that use text messaging is that mobile technologies allow for more opportunities for engagement and interaction between senders and recipients (Suggs, 2006). Commercial campaigns using text messaging have been growing rapidly because they have been found to be effective at stimulating responses from the consumers (Rettie, Grandcolas, & Deakins, 2005).

In the education, cognitive science and behavioral economics literature, one of the most traditional ways of encouraging engagement with the material is to test learners on what they have learned. There is a scientific basis for this strategy: Quiz-based interactions require active processing of the information and, more specifically, require mental retrieval, a process that can lead to changes in mnemonic representation of the information (Carpenter & DeLosh, 2006; Roediger & Karpicke, 2006). It is widely thought that these changes lead to improved knowledge retention. Studies on classroom learning have shown that taking a quiz (even one that has no influence on a student's final grade) can have a greater positive effect on the retention of knowledge than spending an equivalent amount of time restudying the material (Butler & Roediger, 2007). In the case of incomplete mastery of knowledge, self-assessment based on quiz questions has been shown to trigger effort, and continuous engagement in learning (Kornell & Son, 2009).

Studies have further shown that attention and effort can be increased when there is an expected reward to effort (Sarter, Gehring, & Kozak, 2006). Rewarding the correct answers to quiz questions may lead to increased attention to the material being learned and to increased effort in learning that material (Fryer, 2011). The general increase in attention and learning may further serve to make the messages more effective reminders for action (Dale & Strauss, 2009).

Finally, marketing theory suggests that interactions between senders and recipients in text messaging may lead consumers (recipients) to place a higher perceived importance on the information that is delivered (Fjeldsoe *et al.*, 2009). Studies on mobile marketing find that interaction between consumers and the company using mobile technology appears to increase consumer trust in such a way that consumers are more likely to think that the messages are relevant or useful (Lee, 2005). Studies have demonstrated that recipients are more likely to scrutinize information when they are more involved (Bouhlel, Mzoughi, Hadiji, & Slimane, 2011). They also are more likely to place a higher value on the product or service that is being promoted by the messages (Liu & Shrum, 2002). A higher level of perceived relevance and usefulness is thought to be one of the strong drivers for creating favorable attitudes among consumers toward a promoted service (Bauer, Barnes, Reichardt, & Neumann, 2005). This positive attitude can result in higher adoption rates of the promoted services (Zhang & Mao, 2008).

Empirically, however, no study has rigorously evaluated the relative effectiveness of text messaging in delivering health information with or without the opportunity for quiz-based interaction. As has been discussed above, the existing literature provides little evidence on any form of health education using text messages alone (Cole-Lewis & Kershaw, 2010). As far as we know, no study has attempted to compare the effectiveness of different strategies for using text messaging as an education tool that may ultimately promote better health and education outcomes.

2. METHODS

(a) *Sampling*

Using strategies derived from the experimental economics literature (e.g., Duflo, Glennerster, & Kremer, 2007; Imbens & Wooldridge, 2009), we conducted a field experiment to assess the effectiveness of a text message-based health education campaign at improving health and educational outcomes in rural areas of northwest China.[1] A total of 900 fourth grade students in 51 primary schools in Ningxia Autonomous Region were involved in our study. On average, each student/household had two caregivers who owned mobile phones. We randomly sampled participants at the student level.

The randomization process for our study consisted of six steps. First, we obtained a list of all counties in Ningxia Autonomous Region. Ningxia Autonomous Region was selected as the study site because of previous work by the study team (Luo, Shi, Zhang, Zhang, *et al.*, 2012) showing high rates of anemia in this region. Ningxia is also one of the poorest provinces in China. The average per-capita income is RMB 3,180 (USD 522), 23% below the mean national income (National Bureau of Statistics of China, 2008). It is representative of around 1.7 million children living in rural areas in Northwest China.

Second, we randomly selected three counties from the list that met the official criteria for impoverished counties. Third, we obtained a list of all primary schools in the sample counties using official records. Fourth, we conducted a canvass survey to identify all schools that had six grades ("complete primary schools," or *wanxiao* in Chinese) and this list of schools became our formal sampling frame. (We focused on complete primary schools because China's government is currently consolidating rural schools into complete primary schools; as a result, these schools serve more than 90% of all rural school children.) Fifth, we randomly selected 51 primary schools from the sampling frame to serve as our sample schools. Finally, we randomly selected around half of the fourth grade students (900 students) in each sample school for inclusion in the study.[2, 3]

Once the sample selection was complete, we randomly assigned each student into one of three study groups: two treatment groups and a control group with no intervention (Figure 1). The randomization is conducted at the individual level without any clustering. In other words, within each class, some students were assigned to Treatment Group 1; some students were assigned to Treatment Group 2; and some students were assigned to the control group. Specifically, a total of 225 students were randomly assigned to Treatment Group 1, 225 students were assigned to Treatment Group 2 and 450 students were assigned to the control group.[4]

Although we managed to track each student whose contact information was valid and who was still attending school in the sample counties, we had an attrition rate of 12.8% between the baseline and endline surveys. This attrition was mainly due to illnesses and student transfers out of our sample areas (typically to an urban school where one or both of their parents were living and working as migrants). At the time of the endline survey, we were able to track and find 193 students in Treatment Group 1, 201 students in the Treatment Group 2, and 391 students in the control group (Figure 1). In other words, we were able to follow up with 785 out of the initial 900 students at the time of the endline survey.

In Table 1, we show that student attrition is independent of study arm assignment (Tables 1 and 2). The difference in the attrition rates across the three groups is not statistically significant (Table 2, row 11). Moreover, the attrited students had (statistically) similar characteristics across the three experimental groups (Table 1, columns 2 and 3).

By comparing a set of student characteristics among Treatment Group 1, Treatment Group 2, and the control

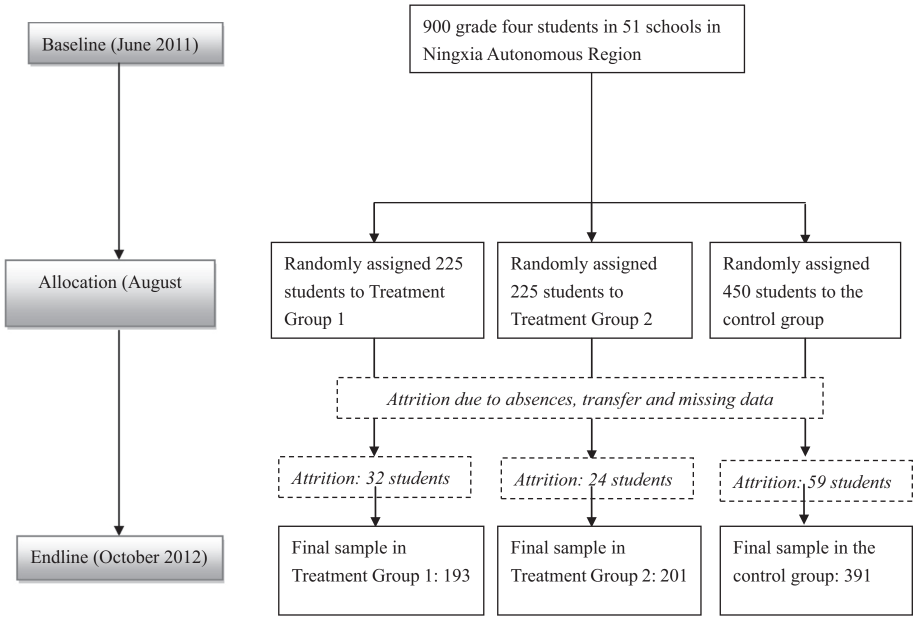

Figure 1. *Experiment profile. Note: Total sample size of students with both baseline and endline data was 193 + 201 + 391 = 785.*

Table 1. *Comparisons of the student characteristics between the attrited students and those remaining in the sample and the characteristics of attrited students between the two treatment groups and the control group*

		Sample: sample + attrition obs. Dependent variable: attrition (1 = attrited, 0 = remained) [1]	Sample: attrition obs. Dependent variable: treatment of weekly health messages (1 = yes, 0 = no) [2]	Sample: attrition obs. Dependent variable: treatment of weekly health messages and monthly quiz questions (1 = yes, 0 = no) [3]
[1]	Baseline hemoglobin level (g/L)	0.00	−0.00	0.00
		(0.00)	(0.01)	(0.01)
[2]	Baseline math test score (standard deviation)	−0.03**	−0.00	−0.03
		(0.01)	(0.08)	(0.07)
[3]	Baseline parental health/ nutrition knowledge (standard deviation)	−0.02	0.07	0.05
		(0.01)	(0.10)	(0.09)
[4]	Baseline student health level (standard deviation) [a]	−0.02	0.05	0.05
		(0.01)	(0.06)	(0.07)
[5]	Student gender (1 = male, 0 = female)	−0.01	−0.20	0.09
		(0.02)	(0.12)	(0.16)
[6]	Student age (years)	−0.06***	0.05	−0.01
		(0.02)	(0.05)	(0.06)
[7]	Father has at least junior high education (1 = yes, 0 = no)	−0.03	0.09	−0.17
		(0.02)	(0.16)	(0.15)
[8]	Mother has at least junior high education (1 = yes, 0 = no)	−0.04	−0.00	−0.01
		(0.03)	(0.19)	(0.29)
[9]	Household asset (1 = higher than the median, 0 = lower or equal to the median)	−0.01	−0.03	−0.00
		(0.03)	(0.19)	(0.12)
[10]	Baseline student self-efficiency scale (0–10 points)	0.03	−0.01	−0.05
		(0.02)	(0.11)	(0.14)
[11]	School dummy	Yes	Yes	Yes
[12]	Observations	900	115	115
[13]	*R*-squared	0.175	0.389	0.378

*, Significant at 10%, **, significant at 5%, ***, significant at 1%. Clustered standard errors at the school level in parentheses.
[a] We used the information on the incidence of health problems to generate a dummy variable for each symptom (equaling one if the child did not have the symptom and zero if the child had the symptom). We then took the mean over all symptoms and normalized the variable to create an indicator of general health (student health level), where high values indicate a healthier child.

Table 2. *Comparison of student and family characteristics of Treatment Group 1, Treatment Group 2, and the control group*

		Text Message-only Group (Treatment Group 1) (193 obs.)		Quiz group (Treatment Group 2) (201 obs.)		Control group (391 obs.)		*p*-Value		
		Mean [1]	SD [2]	Mean [3]	SD [4]	Mean [5]	SD [6]	H_0: [1] = [3] [7]	H_0: [1] = [2] [8]	H_0: [2] = [3] [9]
[1]	Baseline hemoglobin level (g/L)	132.96	10.68	133.69	10.75	132.94	10.30	0.50	0.99	0.41
[2]	Baseline math test score (standard deviation)	0.09	0.98	0.02	1.02	−0.01	0.97	0.46	0.25	0.79
[3]	Baseline parental health/nutrition knowledge (standard deviation).	0.05	1.03	0.00	0.92	0.01	1.03	0.61	0.67	0.89
[4]	Baseline student health level (standard deviation)[a]	0.05	0.93	0.02	0.99	0.02	0.99	0.78	0.71	0.95
[5]	Student gender (1 = male, 0 = female)	1.46	0.50	1.48	0.50	1.50	0.50	0.67	0.33	0.63
[6]	Student age (years)	12.47	1.04	12.49	1.08	12.45	1.06	0.88	0.84	0.71
[7]	Father has at least junior high education (1 = yes, 0 = no)	0.83	0.38	0.83	0.38	0.84	0.37	0.96	0.76	0.80
[8]	Mother has at least junior high education (1 = yes, 0 = no)	0.47	0.50	0.47	0.50	0.46	0.50	0.98	0.71	0.69
[9]	Household asset (1 = higher than the median, 0 = lower or equal to the median)	0.56	0.50	0.62	0.49	0.58	0.49	0.25	0.59	0.43
[10]	Baseline student self-efficiency scale (0–10 points)	2.36	0.52	2.42	0.51	2.41	0.52	0.25	0.28	0.81
[11]	Attrition rate[b]	0.14	0.35	0.11	0.31	0.13	0.34	0.25	0.69	0.36

[a] We used the information on the incidence of health problems to generate a dummy variable for each symptom (equaling one if the child did not have the symptom and zero if the child had the symptom). We then took the mean over all symptoms and normalized the variable to create an indicator of general health (student health level), where high values indicate a healthier child.
[b] The statistics of attrition rate is calculated using the total sample (225 obs. in Text Message-only Group, 225 obs. in Quiz Group and 391 obs. in the control group). All the other statistics are calculated using the sample after attrition.

group, it can be seen that our random assignment was valid (Table 2). We found that none of the differences in student characteristics across the three experimental groups was statistically significant. In addition, the differences that do exist are small in magnitude.

(b) *Interventions*

The experiment includes two intervention arms and a control group with no intervention. At the time of the baseline survey, the mobile phone numbers of all the students' caregivers (typically the parents and/or the grandparents) were collected. An average of two phone numbers per student was collected, for a total of 1,744 phone numbers.

The caregivers of the students in Treatment Group 1 were assigned to receive weekly health messages about anemia and health (henceforth, the *Text Message-only Group*). The caregivers of the students in Treatment Group 2 were assigned to receive the same weekly health messages as in Treatment Group 1; in addition, they were also assigned to receive a monthly quiz question testing caregivers about the health/nutrition knowledge covered by the weekly health messages (henceforth, the *Quiz Group*). A correct response to the monthly quiz question was rewarded by a 10 RMB payment (1.61 USD), made directly to the respondent's mobile device. The phone numbers in the control group did not receive any mobile messages.

In partnership with a mobile marketing company based in Shanghai, the weekly health messages were sent to both the Text Message-only Group and the Quiz Group for a total of 51 weeks. Week 1 was dedicated to introducing the study. In total, there were 52 messages sent out, one per week for one year. The messages included information about: (a) what anemia is (and its prevalence in the sample counties); (b) the impact of untreated anemia on children's educational performance; (c) how to feed children a nutritious balanced diet that can keep them anemia-free, healthy, and performing well in school; and (d) other strategies that could be used by households to increase the level of nutrition (and health and education) of their children. For example, the messages advised caregivers that children require a balanced diet containing both iron-rich foods, such as lean meat, as well as foods rich in vitamin C, such as fruits and vegetables. The messages also stated that micronutrient supplementation, such as a daily multivitamin with iron or the use of iron-fortified flour or soy sauce during meal preparation, could be used to address anemia. The messages also stressed that when children become healthier, their school performance and test scores will improve. The messages were short and written in plain language that could be understood by individuals with relatively poor reading skills and nutritional knowledge. All messages were prefaced with the statement, "According to doctors at Xi'an Jiaotong University's School of Medicine..." All 51 messages were different, though they all focused on the themes described above.

In addition to the 51 weekly health messages, the caregivers of children in the Quiz Group also received a multiple-choice quiz question each month. If the caregiver answered the question correctly, within 24 h they were rewarded with a monetary prize transmitted directly to his or her mobile device. The quiz questions asked about the health messages that had been sent over the course of the previous four weeks. Recipients were invited to respond to the quiz questions and were informed that if they answered correctly, they would be rewarded with a 10 RMB ($1.61 USD) phone credit. The amount of the prize is large enough to attract attention. However, since it is only approximately 2.5% of the annual income of a local farmer, it is unlikely to generate an income effect. Each respondent who answered correctly was sent a congratulatory text message and received their monetary award. Each respondent who answered incorrectly was sent a text message with the correct answer and was encouraged to participate in future quizzes.

(c) *Data*

The research group conducted a total of two rounds of surveys: one at baseline and one at endline. Each survey was administered to our grade four sample students in the 51 schools (Figure 1). The baseline survey was conducted in June, 2011 (at the end of the spring semester), and the text messages were sent to the two intervention groups starting in August, 2011 and continuing through July, 2012. The endline survey was conducted in October, 2012 (at the beginning of the fall semester), 15 months after the baseline survey and two months after the conclusion of the text messaging interventions. All student and teacher surveys were administered by trained members of the research team, composed of around 50 graduate students and research staff from the Ningxia University, Chinese Academy of Sciences, and the Northwest University.

In each round of the survey, the enumeration team collected information from all students participating in the study, as well as from their caregivers and their homeroom teachers. The team visited each school and administered a three-block survey module to all participating students. Students also took a survey form home for their parents to complete and returned it to school the following day. The homeroom teachers were also asked to complete a short survey form evaluating the in-school behavior of their students.

The three-block student survey included a standardized math test, a student health module, and a module on student demographic characteristics. The math test included 30 questions selected from the Trends in International Mathematics and Science Study (TIMSS) test data bank. Students were required to finish the test in 30 min. Our enumeration team monitored the test, strictly enforced the time limit, and took steps to minimize cheating. We normalized the test scores to create a measure of educational performance (*math test score*).[5]

In the second block, the team collected data on student health. We measured each student's anemia status by testing his or her hemoglobin level and asking about his or her general health. Two trained nurses were assigned to each enumeration team to measure the hemoglobin concentrations using blood samples collected through fingerpricks. Hemoglobin levels were measured onsite using portable HemoCue Hb 201+ diagnostic machines. Information was also collected on the incidence of health problems for each student. Students were presented with a list of symptoms of common diseases (such as diarrhea and the flu), and asked whether they had experienced any of these symptoms during the one month prior to the survey.[6] We used the information on the incidence of health problems to generate a dummy variable for each symptom (equaling one if the child did not have the symptom and zero if the child had the symptom). We then took the mean over all symptoms and normalized the variable to create an indicator of general health (*student health level*), where high values indicate a healthier child.

In the third block of the survey, we collected information on student characteristics. With information from this part of the survey we were able to generate a number of control variables, including student's *gender*, *age,* and *self-efficacy*.[7]

Caregivers were also asked to complete a survey form. The caregiver form included around 15 questions testing caregiver's knowledge of anemia and health. The total score was normalized to represent *parental health/nutrition knowledge*. The questions in the caregiver form were also used to generate variables for parental education (*father has at least junior high education* and *mother has at least junior high education*) and a measure of each household's wealth as measured by the value of the primary household assets (*household asset*). In the caregiver form, we also collected information on household health/nutrition practices, such as whether the family regularly bought meat, nutritional supplements, or iron-fortified flour or soy sauce. We also asked the parents to specify the type of nutritional supplements that they gave their children, if any.

Homeroom teachers were asked to evaluate the behavior of each student; this information was used to create two indicators of student in-class behavior. Teachers were asked to record (from classroom records) each student's school *attendance* (i.e. whether the student had been absent from class due to illness in the one month prior to the survey). The teacher also created a subjective measure of each student's *concentration* (how often the student was distracted during class). The measure was created on a scale of one to four (equaling one if a student was often distracted during class and four if the student was never distracted during class). All teachers were blind to all study interventions. During the endline survey, we asked teachers whether they knew that there was an intervention being conducted. Among the teachers we interviewed, 95% of them indicated they had no idea that there was an intervention being implemented at all, and none of the teachers knew which students were in which experimental group.

(d) *Statistical methods*

Through our statistical analysis, we aim to examine the impact of our text messaging campaign on health and educational outcomes. Our hypothesis ex ante was that both of the text messaging interventions would have significant and positive impacts on both parental health/nutrition knowledge and student health and educational outcomes. To test this hypothesis, we estimated the treatment effect in both the parameters of the intention to treat (ITT) and the local average treatment effect (LATE). These estimation strategies were adopted to deal with partial compliance, since not every individual who was assigned to receive the text messages actually received the messages.[8] Partial compliance is commonly found in text messaging interventions (Fafchamps & Minten, 2012).

In order to accommodate partial compliance, we let Z_i^1 equal one if student i was assigned to the Text Message-only Group and T_i^1 equal one if any of the caregivers of student i actually received at least one of the weekly health messages. Likewise, we let Z_i^2 equal one if student i was assigned to the Quiz Group and we let T_i^2 equal one if any of the caregivers of student i actually received at least one of the weekly health messages *and* at least one of the monthly quiz questions. Partial compliance implies that it is possible for $T_i^1 = 0$ even though $Z_i^1 = 1$. The same is true for the Quiz Group.

We believe the measurement of T to be quite accurate. The Shanghai mobile marketing firm allowed us access to the sending logs that their system creates when any text messages are sent. This system automatically records the success or failure of each attempt to send a message.

The intention to treat model compares the outcomes between students assigned to the control group and students assigned to the Text Message-only Group or the outcomes between students assigned to the control group and those assigned to the Quiz Group. It is important to consider the ITT parameters because any policy or program that would use text messaging for health education is likely to have a certain degree of noncompliance (e.g., people change phone numbers, switch service providers, etc.).

The ITT analysis was conducted in two steps. First, we used both unadjusted and adjusted ordinary least squares (OLS) regression analysis to estimate how health and educational outcomes were changed by the treatment assignment. Our unadjusted analysis regressed the endline outcome variables on the dummy variables of the treatment assignment (Z_i^1 and Z_i^2). We used an adjusted analysis as well to improve precision. In all regressions, we corrected for clustering of standard errors at the school level.

First, the unadjusted model is:

$$y_i = \beta_0 + \beta_1 Z_i^1 + \beta_2 Z_i^2 + \varepsilon_i \tag{1}$$

where y_i is the endline outcome variable for child i and ε_i is a random disturbance term.

To improve the efficiency of the estimation, we build on the unadjusted model in Eqn. (1) by including a set of control variables and school fixed effects:

$$y_{1is} = \beta_0 + \beta_1 Z_i^1 + \beta_2 Z_i^2 + \gamma y_{ois} + X_{ois}\theta + \varphi_s + \varepsilon_{is} \tag{2}$$

where all the variables and parameters are the same as those in Eqn. (1), except that we added the baseline outcome variable (y_{ois}), the school fixed effects term (φ_s), and a set of control variables (X_{0is}). The variables in X_{0is} include the time-invariant characteristics of students and parents (*gender, age, father has at least junior high education, mother has at least junior high education, household assets,* and *student self-efficacy*). When estimating the equation by controlling for the baseline outcome variable, such an estimation approach is a difference-in-difference estimator (Angrist & Pischke, 2008).

In order to learn about the impact of the actual treatment of text messaging on student and parent outcomes, we also estimate a model that produces an estimate of the Local Average Treatment Effect (or LATE). Since no one in the control group received any of the messages, our LATE estimate is also the Average Treatment Effect on the Treated estimate (ATT) (Duflo *et al.*, 2007). We estimated our LATE/ATT by using the random assignment (Z_i^1 and Z_i^2) as an instrument for the actual treatment (T_i^1 and T_i^2).

To produce an estimate for our LATE/ATT estimator, we can estimate a IV model:

$$y_{1is} = \beta_0 + \beta_1 T_i^1 + \beta_2 T_i^2 + \gamma y_{ois} + X_{ois}\theta + \varphi_s + \varepsilon_{is} \tag{3}$$

where ε_{is} is an error term possibly correlated with T_i^1 and T_i^2 but uncorrelated, by design, with the instruments Z_i^1 and Z_i^2. The first stage equation of the IV model is:

$$T_i = \pi_0 + \pi_1 Z_i + \omega_i \tag{4}$$

where the actual treatment T_i (T_i^1 and T_i^2) is predicted by the treatment assignment Z_i(Z_i^1 and Z_i^2).[9]

Spillovers across project arms are unlikely, especially since we took actions to minimize the possibility of spillovers or contamination (an important part of study design for a project like this one). On average, each project school served six surrounding villages (*project villages*), each of which had a population of around 300 individuals. Each project village had an average of only 1.5 students whose caregivers were assigned to receive the text messages and an average of only 1.5 students whose caregivers were assigned to be in the control group. In addition, we made sure that the villages in the sample were all rural, poor and, as such, were comprised of multiple hamlets. This reduces even further the probability that the families of any of the students would live near each other. According to our endline data, students reported that they were unaware of any friends whose caregivers received weekly health messages or monthly quiz questions. It is also unlikely that the messages were forwarded from the treatment groups to the control group. In the last part of the endline survey, we asked all caregivers in the control group whether they ever received any health text messages that had the same beginning words in our messages. The data suggest that no one in the control group recognized the messages that were sent to the treatment groups. Even if there were contamination in the control group (e.g., a few families in the control group heard about the messages and changed their health behaviors), our estimate of the treatment effect would then be the lower bound of the program impact.

3. RESULTS

(a) *Receipt of the text messages and quiz question response rate*

Our logs of message receipt show that the weekly health messages reached the majority of the recipients (Table 3). Ninety-one percent of the Text Message-only Group received at least one message, while nearly 60% (59%) received over 80% (41–51 messages) of the messages (row 1). These numbers were slightly lower in the Quiz Group, with 81% of the Quiz Group receiving at least one weekly health message and 48% receiving over 80% of the messages (41–51—row 5).

Participants in the Quiz Group received the majority of the quiz questions; each quiz was received by an average of 70% of the targeted recipients (Table 4, row 1, columns 1–11). Although the initial response rate was low, it grew steadily over time from 15% (Table 4, row 2, column 1) to 28% by the last week (column 11). In each round of quizzes and responses, 79–100% of the responses were correct (except for Quiz 6—row 3, columns 1–11).

Since it is possible that the demographic composition of the household might affect the outcome of the interventions, it is important to ensure that household composition was equivalent across treatment and control groups. In our entire sample, parents account for 60% of caregivers. This is true in the Test Message-only Group (59%), the Quiz Group (60%) and the control group (60%). The data also show that the proportion of grandparents as a share of total caregivers is statistically identical across all three experimental arms: 30% in the Text Message-only Group, 30% in the Quiz Group, and 29% in the control group. Lastly, we find that in around 10% of sample households, the caregiver is a non-parent, non-grandparent individual (e.g., an uncle or other relative), and this proportion was statistically equal for the two treatment groups and the control group.

When examining who actually received our text messages, the data show that the majority of the caregivers who received the text messages were parents (Table 3). Parents account for 71% of all text message recipients in the Text Message-only Group (rows 2–4, column 1) and 75% of all recipients in the Quiz Group (rows 6–8, column 1). The shares of the recipients who were parents are statistically identical. We suspect that this might be due to the fact that parents in rural families are more likely to own mobile phones than are other caregivers.

The families who actually received the messages share similar characteristics with the ones who did not receive the messages

Table 3. *Number and percentage of students' families that received the text messages about nutrition/anemia by the two treatment groups*

		Total sample	Number of messages received[a] (Total number of messages sent = 51)		
			0	Fewer than or equal to 80% (1–40)	More than 80% (41–51)
		[1]	[2]	[3]	[4]
Text Message-only Group (Treatment Group 1)	1. Total	193	19	61	113
		100%	10%	32%	59%
	2. Mother only	33/193	0	6	27
		17%	0%	3%	14%
	3. Father only	72/193	1	24	47
		37%	1%	12%	24%
	4. Both mother and father	33/193	1	8	24
		17%	1%	4%	12%
Quiz Group (Treatment Group 2)	5. Total	201	38	66	97
		100%	19%	33%	48%
	6. Mother only	34/201	5	15	14
		17%	2%	7%	7%
	7. Father only	85/201	11	25	49
		42%	5%	12%	24%
	8. Both mother and father	32/201	3	4	25
		16%	1%	2%	12%

[a] The number of messages received is defined as maximum non-repeated messages that any single caregiver received.

Table 4. *Number and percentage of students' families that received and answered the quiz questions*

		Total number of quiz sent = 11										
		Quiz 1 [1]	Quiz 2 [2]	Quiz 3 [3]	Quiz 4 [4]	Quiz 5 [5]	Quiz 6 [6]	Quiz 7 [7]	Quiz 8 [8]	Quiz 9 [9]	Quiz 10 [10]	Quiz 11 [11]
Total sample = 201	Received	143	140	141	138	147	136	139	140	138	138	136
	(% of total)	71%	70%	70%	69%	73%	68%	69%	70%	69%	69%	68%
	Answered	21	17	16	23	25	23	27	27	26	28	38
	(% of received)	15%	12%	11%	17%	17%	17%	19%	19%	19%	20%	28%
	Accuracy	18	15	16	20	24	11	22	26	22	22	38
	(% of answered)	86%	88%	100%	87%	96%	48%	81%	96%	85%	79%	100%

(Table 5). In both groups, none of the student or family characteristics are significantly different between those who actually received the messages and those who did not. This suggests that conditional on geography and other common characteristics clustered at the school level, successfully receiving (or not receiving) text messages is likely to be a random process.

(b) *Impact on parental health knowledge, student health, and education*

(i) *Parental health/nutrition knowledge*

After receiving the text messages, the results show that caregivers' health/nutrition knowledge increased in the Quiz Group relative to the control group (Table 6). When using the unadjusted model in Eqn. (1), the ITT estimate (using Eqn. (1)) suggests that the health/nutrition knowledge of caregivers in the Quiz Group improved by 0.21 standard deviations (significant at the 5% level, row 2, column 1). This translates to two additional correct answers (out of a total of 15) from caregivers in the Quiz Group compared to those in the control group. When control variables are added (using the adjusted model—using Eqn. (2)), the ITT estimate rises slightly (to an improvement of 0.23 standard deviations) and it becomes more significant (significant at the 1% level, row 2, column 2). The LATE/ATT estimate (using Eqns. (3) and (4)) shows that the treatment of weekly health messages combined with the monthly quiz questions improved caregivers' health/nutrition knowledge by 0.34 standard deviations (significant at the 1% level, row 2, column 3).

In contrast, there was no significant impact of the Text Message-only Group on caregiver knowledge (Table 6). In each of the models, the change in caregiver health/nutrition knowledge between the baseline and endline was statistically identical in the Text Message-only and control groups. In other words, in the absence of monthly quiz questions, the weekly health messages alone did not affect caregiver health/nutrition knowledge (row 1, columns 1–3). The estimated ITT and LATE/ATT parameters are all close to zero and none of the coefficients are significant (row 1, columns 1–3).[10]

(ii) *Student nutrition and health*

In contrast to the impact on parental health/nutrition knowledge, the impact of the interventions on student Hb levels and health is less straightforward (Table 6). The coefficients on the treatment variables for both the Text Message-only Group and the Quiz Group are positive for both the ITT and LATE/ATT models (rows 1–2, columns 4–6). Using the adjusted models in Eqns. (2) and (3), the point estimates suggest that Hb levels rose between 0.40 g/L (ITT) and 0.44 g/L (LATE/ATT) among students in the Text Message-only Group. The point estimates also suggest that Hb levels rose between 0.70 g/L (ITT) and 0.94 g/L (LATE/ATT) among students in the Quiz Group. However, none of these coefficients are statistically significant.[11]

Table 5. *Ordinary least squares estimators of the characteristics of the students/households who received the text messages in the Text Message-only Group and both text messages and quiz questions in the Quiz Group in the 51 primary schools in Ningxia*

		Dependent variable: Received any of the weekly health messages (1 = yes, 0 = no)	
		Sample: Text Message-only Group [1]	Sample: Quiz Group [2]
[1]	Baseline parental health/nutrition knowledge (0–10 points)	0.01	0.02
		(0.03)	(0.04)
[2]	Baseline hemoglobin level (g/L)	0.00	−0.00
		(0.00)	(0.00)
[3]	Baseline student health level (0–1 scale) [a]	0.03	0.01
		(0.04)	(0.05)
[4]	Baseline math test score (standard deviation)	0.02	0.05
		(0.03)	(0.04)
[5]	Student gender (1 = male, 0 = female)	0.00	0.04
		(0.08)	(0.08)
[6]	Student age (years)	−0.03	−0.01
		(0.03)	(0.04)
[7]	Father has at least junior high education (1 = yes, 0 = no)	−0.04	0.10
		(0.06)	(0.08)
[8]	Mother has at least junior high education (1 = yes, 0 = no)	−0.09	0.09
		(0.09)	(0.12)
[9]	Household asset (1 = higher than the median, 0 = lower or equal to the median)	−0.07	−0.10
		(0.06)	(0.09)
[10]	Baseline student self-efficiency scale (0–10 points)	0.09	−0.04
		(0.06)	(0.06)
[11]	School dummy	Yes	Yes
[12]	Observations	193	201

*, Significant at 10%, **, significant at 5%, ***, significant at 1%. Clustered standard errors at the school level in parentheses.
[a] We used the information on the incidence of health problems to generate a dummy variable for each symptom (equaling one if the child did not have the symptom and zero if the child had the symptom). We then took the mean over all symptoms and normalized the variable to create an indicator of general health (student health level), where high values indicate a healthier child.

Table 6. *Ordinary least squares and IV estimators of the impacts of treatment arms on parental knowledge about health/nutrition, student hemoglobin level, and student health level in the 51 primary schools in Ningxia*

		Dependent variable								
		Endline parental knowledge about health/nutrition (standard deviation)			Endline hemoglobin level (g/L)			Endline student health level (standard deviation)		
		(1) OLS (unadjusted)	(2) OLS (adjusted)	(3) IV	(4) OLS (unadjusted)	(5) OLS (adjusted)	(6) IV	(7) OLS (unadjusted)	(8) OLS (adjusted)	(9) IV
[1]	Treatment of weekly health messages (1 = yes, 0 = no)	0.01	−0.02	−0.03	0.39	0.40	0.44	−0.08	−0.07	−0.08
		(0.09)	(0.08)	(0.09)	(1.10)	(1.13)	(1.26)	(0.08)	(0.09)	(0.10)
[2]	Treatment of weekly health messages and monthly quiz questions (1 = yes, 0 = no)	0.21**	0.23***	0.34***	0.48	0.70	0.94	0.16*	0.15*	0.24**
		(0.09)	(0.07)	(0.10)	(1.24)	(0.91)	(1.29)	(0.08)	(0.08)	(0.11)
[3]	Control variables [a]	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes
[4]	Observations	785	785	785	785	785	785	785	785	785
[5]	*R*-squared	0.008	0.364	0.363	0	0.392	0.392	0.007	0.119	0.111

*, Significant at 10%, **, significant at 5%, ***, significant at 1%. Clustered standard errors at the school level in parentheses.
[a] The control variables include all the independent variables that are listed in Table 5 (including the school dummies).

Looking at other dimensions of student health, however, our analysis shows that the Quiz treatment did lead to improvements in the general health level of the students (Table 6). The estimated ITT parameter suggests that being in the Quiz Group increased student health levels by 0.16 standard deviations when using the unadjusted model and 0.15 standard deviations when using the adjusted model (both coefficients are significant at the 10% level, row 2, columns 7 and 8). Intuitively, this means that students in the Quiz Group were 5% less likely to have had any medical symptoms of common diseases. The LATE/ATT estimate suggests that receiving the weekly messages in combination with the monthly quiz questions makes students healthier by 0.24 standard deviations (significant at the 5% level, row 2, column 9). No significant treatment effect is found on the Text Message-only Group regardless of which model is used (row 1, columns 7–9).[12]

(iii) *Effects in the classroom*

The effects of the Quiz treatment were not limited to student health; they also appeared in the classroom. We tested the treatment effect of text messaging on both school attendance

Table 7. *Ordinary least squares and IV estimators of the impacts of treatment arms on student school attendance, student concentration, and standardized math test score in the 51 primary schools in Ningxia*

		Dependent variable					
		Student school attendance/ absence due to illness during one month before survey (1 = was never absent, 0 = has been absent)		Student concentration (1–4 points, 1 = was often distracted in class and 4 = was never distracted in class)		Endline standardized math test score (standard deviation)	
		(1) OLS (adjusted)	(2) IV	(3) OLS (adjusted)	(4) IV	(5) OLS (adjusted)	(6) IV
[1]	Treatment of weekly health messages (1 = yes, 0 = no)	0.02 (0.05)	0.02 (0.06)	0.13 (0.09)	0.14 (0.09)	0.03 (0.08)	0.04 (0.09)
[2]	Treatment of weekly health messages and monthly quiz questions (1 = yes, 0 = no)	0.00 (0.05)	0.00 (0.07)	0.15** (0.07)	0.19* (0.10)	0.14** (0.07)	0.19** (0.10)
[3]	Control variables [a]	Yes	Yes	Yes	Yes	Yes	Yes
[4]	Observations	785	785	785	785	785	785
[5]	R-squared	0.139	0.139	0.241	0.24	0.492	0.492

*, Significant at 10%, **, significant at 5%, ***, significant at 1%. Clustered standard errors at the school level in parentheses.
[a] The control variables include all the independent variables that are listed in Table 5 (including the school dummies).

and student concentration. We found no significant impact of either intervention on school attendance, regardless of estimation method (Table 7, rows 1 and 2, columns 1 and 2). By contrast, students in the Quiz Group were rated by their teachers as better able to concentrate during class (Table 7). The ITT estimate suggests that student concentration improved by 0.15 points (on a scale of one to four points) for students in the Quiz Group (significant at the 5% level, row 2, column 3). The LATE/ATT estimate suggests that receiving the weekly messages plus the monthly quiz questions improved student concentration by 0.19 points (significant at the 10% level, row 2, column 4).[13] Among students in the Text Message-only Group, the estimates of the treatment effect are not statistically significant, although the coefficients are positive (row 1, columns 3 and 4).

While we cannot identify the precise mechanism (possibly better concentration in class or better health), the standardized math test scores of students in the Quiz Group improved by 0.14 standard deviations (according to the adjusted and unadjusted models) after the treatment (Table 7, row 2, column 5). The coefficient is significant at the 5% level (row 2, column 5). The LATE/ATT estimate suggests that among the Quiz Group students whose parents actually received the weekly messages plus the monthly quiz questions, the improvement is larger, at 0.19 standard deviations (significant at the 5% level, row 2, column 6). Consistent with previous results, there is no evidence to suggest that weekly health messages alone (that is, the Text Message-only treatment) affect students' math test scores (row 1, columns 5–6).[14, 15]

(c) *Impact on health practices*

In order to gain a better understanding of the possible mechanisms behind the improvements in health and educational outcomes, we tested the treatment effect of text messaging on a set of family health practices. Following the list of health practices that were suggested by the weekly health messages, we asked caregivers in the endline survey whether the family bought (a) meat; (b) nutritional supplements; (c) vitamins; or (d) iron-fortified flour or soy sauce. Using these as dependent variables, we estimated the impact of the treatments using both the ITT and the LATE/ATT parameters using Eqns. (2) and (3).

According to our analyses, the caregivers of children in the Quiz Group were more likely to buy nutritional supplements relative to the control group (Table 8). The ITT estimate suggests that caregivers in the Quiz Group were 8% more likely to have purchased nutritional supplements (significant at the 10% level, row 2, column 3). The LATE/ATT estimate suggests that when both the weekly health messages and the monthly quiz questions were received, the impact on the probability of purchasing nutritional supplements is even larger, at 12% (significant at the 5% level, row 2, column 4).[16, 17]

In contrast, no significant treatment effect is found on other health practices (Table 8). We do not find evidence of any treatment effect on the purchase of meat, vitamins, or iron-fortified flour or soy sauce among caregivers in the Quiz Group (row 2, columns 1, 2, and 5–8). Consistent with our other findings, caregivers in the Text Message-only Group did not show any change in health practices (row 1, columns 1–8). All of the coefficients are close to zero and none of them is statistically significant.

Based on the results above, a possible mechanism for the improvement in the health and educational outcomes might be an increase in feeding children a variety of nutritional supplements. From the list of nutritional supplements that the parents reported, it appears to be possible for rural caregivers to obtain access to a variety of nutritional supplements in local shops or pharmacies. According to surveys of local pharmacies and village clinicians, these supplements did not always include micronutrients. Instead, they contained a variety of ingredients, including glucose, protein, and calcium.

Our endline survey and post-study interview also allowed us to identify why the families in our intervention groups did not purchase the other items (i.e., meat, vitamins, or iron-fortified food products) that were suggested in the weekly text messages. One possible reason might be limited market access—meat is sometimes logistically difficult to purchase and store, and it is almost impossible to find vitamins or iron-fortified flour in rural areas. In fact, our data support such conclusions. Nearly three-quarters (74%) of the families who did not buy meat reported that this was because there were no shops close by or because they did not have refrigeration facilities at home in which to store meat. Most of the families who purchased vitamins (only 10% of families reported doing so) reported that they had to go to the county or prefectural seat to make

Table 8. *Ordinary Least Squares and IV estimators of the impacts of treatment arms on household purchase of meat and nutritional supplement in the 51 primary schools in Ningxia*

		Dependent variable							
		Student's family bought meat during one month before the survey (1 = yes, 0 = no)		Student's family bought nutritional supplement during the treatment period (1 = yes, 0 = no)		Student's family bought vitamin during the treatment period (1 = yes, 0 = no)		Student's family bought iron-fortified flour or soy sauce during the treatment period (1 = yes, 0 = no)	
		(1) OLS (adjusted)	(2) IV	(3) OLS (adjusted)	(4) IV	(5) OLS (adjusted)	(6) IV	(7) OLS (adjusted)	(8) IV
[1]	Treatment of weekly health messages (1 = yes, 0 = no)	−0.00	−0.00	−0.02	−0.02	−0.02	−0.02	−0.02	−0.03
		(0.00)	(0.00)	(0.05)	(0.05)	(0.03)	(0.03)	(0.02)	(0.03)
[2]	Treatment of weekly health messages and monthly quiz questions (1 = yes, 0 = no)	−0.01	−0.01	0.08*	0.12**	0.01	0.02	0.02	0.03
		(0.01)	(0.01)	(0.04)	(0.05)	(0.03)	(0.04)	(0.03)	(0.04)
[3]	Control variables[a]	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
[4]	Observations	785	785	785	785	785	785	785	785
[5]	*R*-squared	0.052	0.041	0.112	0.104	0.074	0.072	0.109	0.110

*, Significant at 10%, **, significant at 5%, ***, significant at 1%. Clustered standard errors at the school level in parentheses.
[a] The control variables include all the control variables and school dummies that are included in Table 5 (including the school dummies).

the purchase. On average, the trip was more than 100 km (one way). Not one family said they were able to purchase either iron-fortified flour or soy sauce; no one knew where to buy these products.

(d) *Heterogeneous effects by message recipients*

In order to explore whether the nature of the impacts varies according to which caregiver receives the messages, we conducted analyses to test for heterogeneous effects by message recipient. We estimated the LATE/ATT parameters by including a vector of variables indicating which household member actually received the messages.

The results in Table 9 suggest that the treatment is more likely to have a positive impact on the health and educational outcomes of children whose mother was the only message recipient in the family to receive both weekly health messages and monthly quiz questions. In the Quiz Group, when only the mother received the messages, caregiver health/nutrition knowledge increased by 0.35 standard deviations (significant at the 10% level, row 2, column 1), student health levels increased by 0.52 standard deviations (significant at the 1% level, row 2, column 2) and student math test scores increased by 0.31 standard deviations (significant at the 10% level, row 2, column 3). In the Quiz Group, when only the father received the messages, caregiver health/nutrition knowledge was improved by 0.23 standard deviations (significant at the 5% level, row 4, column 1) student health levels were not significantly improved (the estimate is 0.16 standard deviations but insignificant at the 10% level, row 4, column 2) and student math test scores were improved by 0.30 standard deviations (significant at the 1% level, row 4, column 3). The Wald test suggests that the difference in the impact on student health levels is significant at the 10% level between the group whose

Table 9. *IV estimators of the heterogeneous effects of the treatment arms by the person who received the messages (mother only, father only or both mother and father) in the 51 primary schools in Ningxia*

		Dependent variable		
		Endline parental knowledge about health/nutrition (standard deviation) (1)	Endline student health level (standard deviation) (2)	Endline standardized math test score (standard deviation) (3)
[1]	**Only mother** received any of the weekly health messages (1 = yes; 0 = no)	0.16	−0.05	0.11
		(0.14)	(0.18)	(0.15)
[2]	**Only mother** received any of the weekly health messages and any of the monthly quiz questions (1 = yes, 0 = no)	0.35*	0.52***	0.31*
		(0.21)	(0.18)	(0.18)
[3]	**Only father** received any of the weekly health messages (1 = yes; 0 = no)	−0.16	0.11	0.02
		(0.12)	(0.12)	(0.11)
[4]	**Only father** received any of the weekly health messages and any of the monthly quiz questions (1 = yes, 0 = no)	0.23**	0.16	0.30***
		(0.10)	(0.11)	(0.08)
[5]	Control variables[a]	Yes	Yes	Yes
[6]	Observations	785	785	785
[7]	*R*-squared	0.365	0.114	0.511

*, Significant at 10%, **, significant at 5%, ***, significant at 1%. Clustered standard errors at the school level in parentheses.
[a] The control variables include all the independent variables that are listed in Table 5 (including the school dummies).

mother was the only quiz recipient and the group whose father was the only quiz recipient. The differences in impact on parental health/nutrition knowledge and math test scores between these two groups are not significant.

There are several reasons why the Quiz Group is more likely to have an impact in improving child health when mothers receive them. One of the reasons could be that mothers are more likely to be the ones who cook for the child and purchase food for the family. In fact, our data support this conjecture. Our data shows that 80% of mothers cook at home; while only 1% of fathers do so. Our data also shows that 80% of mothers are in charge of purchasing food for the family; in contrast, only 19% of fathers are in charge of purchasing food.

4. CONCLUSION

In search of an effective way to educate the rural poor in China about health and nutrition and in an attempt to explore how text messaging can be designed to deliver health information to remote and disadvantaged communities, we conducted a randomized field experiment in one of the poorest areas in China, rural Ningxia Autonomous Region. In this study, we tested whether health education using text messaging was effective at raising parental knowledge of health/nutrition, promoting student health, and improving student academic performance. According to our analyses, the majority of the messages successfully reached the students' caregivers. In addition, when caregivers received both weekly health messages and monthly quiz questions, caregiver health/nutrition knowledge improved, and students experienced gains in both physical health and academic performance. However, no significant impact on caregiver knowledge or student outcomes was found among families who only received weekly health messages without the monthly quiz questions. We also found that among households in the Quiz Group, increased purchases of nutritional supplements were observed. Finally, we found that Quiz Group students whose mother was the only one in the family who received the messages had relatively larger gains in student health level.

Such stark differences in the health and educational impacts of the two interventions, the Text Message-only treatment and the Quiz treatment, point to the importance of interactive, rather than passive, text messaging programs. Future research is needed to disentangle the mechanisms for the impact of the quiz and the monetary prize. Behavioral economics and education studies have suggested several possible mechanisms by which the importance of the quiz messages might be explained, such as the retrieval effect of quiz questions, attention, and perception of importance/relevance of the information that are induced by the interaction or the prize (Fjeldsoe *et al.*, 2009; Roediger & Karpicke, 2006; Sarter *et al.*, 2006). Alternatively, it may be the case that while the quiz messages and prizes provided the initial incentive for parents in the Quiz Group to pay closer attention to the content of the messages (and to their child's health), ultimately it was visible improvements in their child's health caused by this initial incentivized push that fueled long-run adherence to a more nutritional diet for their child. Unfortunately, one limitation of this study is that there is no way to test which of these mechanisms is behind the success of the quiz messages.

Another important limitation of this study, and of any intervention that attempts to use text messaging to reach vulnerable populations, is that even with the high rates of cell phone ownership that we see in the developing world today, there is still a small fraction of the population without access to cell phone technology, and this fraction may in fact be the population most in need of health intervention. For example, families living in the most remote areas might not have the regular access to electrical or network infrastructure that is needed to support cell phone technology. These same families are likely to have the longest travel times to the nearest clinic or health facility, and the lowest access to other income-generating technologies. Overlooking these types of households is a major shortcoming of any intervention that aims to improve the lives of rural families through the use of mobile technology.

Our study also raises important questions about the potential role of text messaging as an education tool within existing health systems. For example, could text messages be used to educate the general public about not just specific health issues, but about logistical issues such as national health insurance options, reimbursement amounts and procedures, or local service availabilities? Research shows that misinformation about coverage options can be a major barrier to care in developing countries (see for example: Jehu-Appiah, Aryeetey, Agyepong, Spaan, & Baltussen, 2012).

Similarly, while we know from previous studies that health education pamphlets on their own are of limited effectiveness in terms of improving health outcomes (Luo, Shi, Zhang, Liu, *et al.*, 2012; Luo, Shi, Zhang, Zhang, *et al.*, 2012), we do not yet have evidence on how such pamphlets might interact with a text messaging campaign. Further research is needed to identify the most cost-effective way of reaching individuals who would most benefit from health information.

Overall, our study has important implications for both the policy environment within China, and for the global development community outside of China. Inside of China, our study underlines the vast potential of text messaging as a tool to reach and engage China's rural poor. The Chinese government is currently in the midst of unveiling a number of new and innovative health policy interventions targeting the poor, ranging from free maternal and child health care services, to expanded rural health insurance coverage, to free nutritional supplements for infants and toddlers. One low-cost way to potentially improve uptake of these new initiatives might be to implement a text message campaign that educates rural families about the details, benefits, and eligibility requirements of the new policies.

Outside of China, the results of our study can be used to inform NGOs and policymakers looking to design their own health education campaigns. We have shown that a successful health education campaign is not only sustained, but also interactive: a weekly text message on its own is not enough to change household behavior. The type of phone-based interactive health education strategy that we used in our study could easily be applied to settings where an in-person home visit might be met with embarrassment or distrust, such as efforts to increase rates of condom usage in African communities at-risk for HIV.

It is our hope that future research will build on the results from this study to identify more precisely what types of interactive text messaging campaigns are most effective for increasing adoption of desired health behaviors, and that the global community can continue to discover more and better applications for such technology.

NOTES

1. This project was funded by generous grants from the Ford Foundation and the Stanford University Global Underdevelopment Action Fund, as well as by a financial gift from Paul Cavey. None of the funders played any role in the study design; in the collection, analysis or interpretation of data; in the writing of the report; or in the decision to submit the article for publication.

2. Our power calculations suggested that we needed a total of 900 students to detect an effect size of 0.2 standard deviations per intervention with 80% power at the 1% significance level (even accounting for tests of multiple outcomes). We assumed a pre- and post-intervention correlation of 0.5 and 15% loss of effective sample due to attrition and noncompliance. The canvass survey showed that there were around 1,800 students in the 51 schools. Therefore, we decided to randomly select half of the fourth grade students to reach a sample size of 900. In total, we surveyed 937 students, 900 of whom were included in the experiment as either treatment or control students.

3. We chose to work with a fourth grade population for two main reasons. First, our study team has conducted a number of nutrition-related studies targeting primary school students in rural China, and has found not only that anemia is a major problem among this population, but also that it affects their school performance (Luo *et al.*, 2011). Second, the internationally scaled standardized math test that we use as an outcome of educational performance in this study (described in more detail below) is specifically targeted toward fourth grade students.

4. Since we are interested in comparing each of the treatment groups with the control group separately and we put equal weight on each pair of the comparison, we need twice as many observations in the control group as in each of the treatment groups (Duflo *et al.*, 2007). Therefore, we assigned half of the sample to the two treatment groups and half of them to the control group.

5. Scores are normalized relative to the distribution of the baseline test scores of the control group. Specifically, we subtracted the mean of the control group and divided by the standard deviation. By normalizing the raw scores, student performance is measured by units of standard deviation from the mean. We use the distribution of the control group to rescale the raw scores because these students are not affected by the treatment and their scores represent how students would perform if there were no intervention. This method of normalization has been widely adopted in the field of education and economics (e.g., Duflo, Glennerster & Kremer, 2008).

6. The same list of symptoms was sent to the students' caregivers. The symptoms reported by the family highly resemble the ones reported by the students. Therefore, we believe the variable *student health level* is a good indicator of general student health.

7. Self-efficacy is a variable generated using the General Self-Efficacy Scale (GSE) by Jerusalem and Schwarzer (1979). The indicator was created to assess a general sense of perceived self-efficacy with the aim to predict their ability of coping with daily hassles as well as planning and taking actions to reach a particular goal.

8. There are many reasons why an individual might not have received the text messages sent to him/her. For example, the individual's phone may have been turned off at the time the message was sent, the individual may have changed his/her phone number, or the individual may have simply been in an area with poor reception.

9. We have also conducted robustness checks by running regressions using different endogenous variables, such as the number of weekly health messages and the number of monthly quiz questions that were received by the recipients. The results of using different endogenous variables have shown that the impact of weekly health messages and monthly quiz questions are consistent and robust across different models. The results are available upon request.

10. The F-test rejects the hypothesis that the coefficient on the treatment of weekly health messages is equal to the coefficient on the treatment of weekly health messages and monthly quiz questions (with a p-value of 0.06 when using the unadjusted OLS model in Eqn. (1); a p-value of 0.06 when using the adjusted OLS model in Eqn. (2); and a p-value of 0.05 when using the IV model in Eqn. (3)).

11. The F-test does not reject the hypothesis that the coefficient on the treatment of weekly health messages is equal to the coefficient on the treatment of weekly health messages and monthly quiz questions (with a p-value of 0.95 when using the unadjusted OLS model in Eqn. (1); a p-value of 0.77 when using the adjusted OLS model in Eqn. (2); and a p-value of 0.74 when using the IV model in Eqn. (3)).

12. The F-test rejects the hypothesis that the coefficient on the treatment of weekly health messages is equal to the coefficient on the treatment of weekly health messages and monthly quiz questions (with a p-value of 0.01 when using the unadjusted OLS model in Eqn. (1); a p-value of 0.02 when using the adjusted OLS model in Eqn. (2); and a p-value of 0.02 when using the IV model in Eqn. (3)).

13. In many respects, such effects are quite remarkable, given the way they were measured. Teachers were blind to the fact of the experiment and to experimental group assignment, yet they were able to independently identify students in the Quiz Group as those whose levels of concentration improved most, by a statistically significant margin.

14. We have also explored whether responding to more or fewer of the quiz questions makes a difference in the treatment effect among the Quiz Group. The results show that responding to more of the quiz questions is associated with larger positive impacts on parental health/nutrition knowledge, student health levels, and math test scores. On the one hand, it may suggest that more interaction between the recipients and senders made the treatment more effective in improving knowledge and changing behaviors. On the other hand, it may be the result of the fact that the recipients who responded more to the quiz questions are the ones who were more likely to think the messages were useful or important. They were thus more likely to take better care of the health and education of their kids. Despite the interesting results, our instrument, i.e. the random assignment of receiving the monthly quiz questions, cannot account for the selection of how many quiz questions to which the recipients responded. Therefore, we have not included the results in the paper (the results are available upon request).

15. The F-test rejects the hypothesis that the coefficient on the treatment of weekly health messages is equal to the coefficient on the treatment of weekly health messages and monthly quiz questions in the IV regression using Eqn. (3) (the p-value is 0.09). The hypothesis is not rejected in the OLS models (the p-value in the unadjusted OLS model without control variables is 0.41; and the p-value in the adjusted OLS model is 0.12).

16. The F-test rejects the hypothesis that the coefficient on the treatment of weekly health messages is equal to the coefficient on the treatment of weekly health messages and monthly quiz questions (with a p-value of 0.06 when using the unadjusted OLS model in Eqn. (1); a p-value of 0.06 when using the adjusted OLS model in Eqn. (2); and a p-value of 0.05 when using the IV model in Eqn. (3)).

17. We also investigated whether the recipients who received more prizes demonstrated different purchasing behavior compared to the individuals that received fewer prizes. In doing so, we conducted a test by regressing the purchase of nutritional supplements on the number of correct

responses (which had a one-to-one correspondence with prize pay-outs) of the recipients among the Quiz Group. We also included an additional dummy of whether the individuals ever responded to any of the quiz questions. The result shows that getting a higher rate of prize pay-outs is not associated with the likelihood of the family to purchase nutritional supplements. The coefficient on the variable of number of correct responses the recipient sent to answer the quiz questions is 0.01 and insignificant. In other words, we cannot reject the hypothesis that there is no difference between getting more prizes and getting fewer prizes in purchasing the nutritional supplement. The test suggests that the quiz questions (and the prizes) may not have worked through relaxing liquidity constraint of the families to purchase nutritional supplements. The resulting table is available upon request.

REFERENCES

Angrist, J. D., & Pischke, J. (2008). *Mostly harmless econometrics: An empiricist's companion*. Springer.

Bauer, H. H., Barnes, S. J., Reichardt, T., & Neumann, M. M. (2005). Driving consumer acceptance of mobile marketing: A theoretical framework and empirical study. *Journal of Electronic Commerce Research, 6*(3), 181–192.

Black, M. M., Baqui, A. H., Zaman, K., Persson, L. A., El Arifeen, S., Le, K., *et al.* (2004). Iron and zinc supplementation promote motor development and exploratory behavior among Bangladeshi infants. *The American Journal of Clinical Nutrition, 80*(4), 903–910.

Bobonis, G. J., Miguel, E., & Puri-Sharma, C. (2006). Anemia and school participation. *Journal of Human Resources, 41*(4), 692–721.

Bouhlel, O., Mzoughi, N., Hadiji, D., & Slimane, I. B. (2011). Brand personality's influence on the purchase intention: A mobile marketing case. *International Journal of Business and Management, 6*(9), 210–227.

Butler, A. C., & Roediger, H. L. III, (2007). Testing improves long-term retention in a simulated classroom setting. *European Journal of Cognitive Psychology, 19*(4–5), 514–527.

Carpenter, S. K., & Delosh, E. L. (2006). Impoverished cue support enhances subsequent retention: Support for the elaborative retrieval explanation of the testing effect. *Memory & Cognition, 34*(2), 268–276.

Chen, X., & Li, B. S. (2000). Depressed mood in Chinese children: Development significance for social and school adjustment. *International Journal of Behavioral Development, 24*(4), 472–479.

Chen, J., Zhao, X., Zhang, X., Yin, S., Piao, J., Huo, J., *et al.* (2005). Studies on the effectiveness of NaFeEDTA-fortified soy sauce in controlling iron deficiency: A population-based intervention trial. *Food & Nutrition Bulletin, 26*(2), 177–186.

China Internet Network Information Center (2012). *Survey report on the development of China's mobile network* (in Chinese). Available online: http://www.cnnic.net.cn/research/bgxz/ydhlwbg/201203/P020120329416039769203.pdf.

China Youth Daily (2012). *What's behind pupils' bread and milk lunch*? Available online: http://english.people.com.cn/90882/8038518.html.

Cole-Lewis, H., & Kershaw, T. (2010). Text messaging as a tool for behavior change in disease prevention and management. *Epidemiologic Reviews, 32*(1), 56–69.

Dale, A., & Strauss, A. (2009). Don't forget to vote: Text message reminders as a mobilization tool. *American Journal of Political Science, 53*(4), 787–804.

Delamere, S., Dooley, S., Harrington, L., King, A., & Mulcahy, F. (2006). Safer sex text messages: Evaluating a health education intervention in an adolescent population. *Sexually Transmitted Infections, 82*, A27.

Denkinger, C. M., Grenier, J., Stratis, A. K., Akkihal, A., Pant-Pai, N., & Pai, M. (2013). Mobile health to improve tuberculosis care and control: A call worth making. *International Journal of Tuberculosis and Lung Disease, 17*(6), 719–727.

Du, S., Mroz, T. A., Zhai, F., & Popkin, B. M. (2004). Rapid income growth adversely affects diet quality in China—Particularly for the poor! *Social Science & Medicine, 59*(7), 1505–1515.

Duflo, E., Glennerster, R., & Kremer, M. (2007). Using randomization in development economics research: A toolkit. *Handbook of Development Economics, 4*, 3895–3962.

Duflo, E., Glennerster, R., & Kremer, M. (2008). Using randomization in development economics research: A toolkit. In P. Schultz, & J. Strauss (Eds.), *Handbook of Development Economics* (4). Amsterdam: North-Holland.

Fafchamps, M., & Minten, B. (2012). Impact of sms-based agricultural information on Indian farmers. *The World Bank Economic Review, 26*(3), 383–414.

Fjeldsoe, B. S., Marshall, A. L., & Miller, Y. D. (2009). Behavior change interventions delivered by mobile telephone short-message service. *American Journal of Preventive Medicine, 36*(2), 165–173.

Food and Agriculture Organization of the United Nations (1999). *FAO – Nutrition country profiles (China)*. Rome: FAO, Available online at: http://www.fao.org/ag/agn/nutrition/chn_en.stm.

Franklin, V. L., Waller, A., Pagliari, C., & Greene, S. A. (2006). A randomized controlled trial of Sweet Talk, a text-messaging system to support young people with diabetes. *Diabetic Medicine, 23*(12), 1332–1338.

Free, C., Phillips, G., Galli, L., Watson, L., Felix, L., Edwards, P., *et al.* (2013). The effectiveness of mobile-health technology-based health behaviour change or disease management interventions for health care consumers: A systematic review. *PLoS Medicine, 10*(1), e1001362.

Fry, J. P., & Neff, R. A. (2009). Periodic prompts and reminders in health promotion and health behavior interventions: Systematic review. *Journal of Medical Internet Research, 11*(2).

Fryer, R. (2011). Financial incentives and student achievement: Evidence from randomized trials. *The Quarterly Journal of Economics, 126*, 1755–1798.

Haapala, I., Barengo, N. C., Biggs, S., Surakka, L., & Manninen, P. (2009). Weight loss by mobile phone: A 1-year effectiveness study. *Public Health Nutrition, 12*(12), 2382–2391.

Halterman, J. S., Kaczorowski, J. M., Aligne, C. A., Auinger, P., & Szilagyi, P. G. (2001). Iron deficiency and cognitive achievement among school-aged children and adolescents in the United States. *Pediatrics, 107*(6), 1381–1386.

Huang, J. K., & Rozelle, S. (2008). Agricultural development, nutrition and the policies behind China's success. Paper produced for the World Food Program.

Hurling, R., Catt, M., De Boni, M., Fairley, B. W., Hurst, T., Murray, P., *et al.* (2007). Using internet and mobile phone technology to deliver an automated physical activity program: Randomized controlled trial. *Journal of Medical Internet Research, 9*(2), e7.

Imbens, G. W., & Wooldridge, J. M. (2009). Recent developments in the econometrics of program evaluation. *Journal of Economic Literature, 47*(1), 5–86.

Jehu-Appiah, C., Aryeetey, G., Agyepong, I., Spaan, E., & Baltussen, R. (2012). Household perceptions and their implications for enrolment in the National health Insurance Scheme in Ghana. *Health Policy and Planning, 27*(3), 222–233.

Jerusalem, M., & Schwarzer, R. (1979). Self-efficacy as a resource factor in stress appraisal. In R. Schwarzer (Ed.), *Self-efficacy: Thought control of action* (pp. 195–216). Washington, DC: Hemisphere.

Kim, S., Haines, P. S., Siega-Riz, A. M., & Popkin, B. M. (2003). The Diet Quality Index-International (DQI-I) provides an effective tool for cross-national comparison of diet quality as illustrated by China and the United States. *The Journal of Nutrition, 133*(11), 3476–3484.

Kornell, N., & Son, L. K. (2009). Learners' choices and beliefs about self-testing. *Memory, 17*(5), 493–501.

Lee, T. (2005). The impact of perceptions of interactivity on customer trust and transaction intentions in mobile commerce. *Journal of Electronic Commerce Research, 6*(3), 165–180.

Lim, M. S., Hocking, J. S., Hellard, M. E., & Aitken, C. K. (2008). SMS STI: A review of the uses of mobile phone text messaging in sexual health. *International Journal of STD & AIDS, 19*(5), 287–290.

Liu, Y., & Shrum, L. J. (2002). What is interactivity and is it always such a good thing? Implications of definition, person, and situation for the influence of interactivity on advertising effectiveness. *Journal of Advertising*, 53–64.

Lozoff, B., Jimenez, E., Hagen, J., Mollen, E., & Wolf, A. W. (2000). Poorer behavioral and developmental outcome more than 10 years after treatment for iron deficiency in infancy. *Pediatrics, 105*(4), e51.

Luo, R., Kleiman-Weiner, M., Rozelle, S., Zhang, L., Liu, C., Sharbono, B., *et al.* (2010). Anemia in rural China's elementary schools:

Prevalence and correlates in Shaanxi Province's poor counties. *Ecology of Food and Nutrition, 49*(5), 357–372.

Luo, R., Shi, Y., Zhang, L., Liu, C., Rozelle, S., & Sharbono, B. (2009). Malnutrition in China's rural boarding schools: The case of primary schools in Shaanxi Province. *Asia Pacific Journal of Education, 29*(4), 481–501.

Luo, R., Shi, Y., Zhang, L., Liu, C., Rozelle, S., Sharbono, B., *et al.* (2012). Nutrition and educational performance in rural China's elementary schools: Results of a randomized control trial in Shaanxi province. *Economic Development and Cultural Change, 60*(4), 735–772.

Luo, R., Shi, Y., Zhang, L., Zhang, H., Miller, G., Medina, A., *et al.* (2012). The limits of health and nutrition education: Evidence from three randomized-controlled trials in rural China. *CESifo Economic Studies, 58*(2), 385–404.

Luo, R., Zhang, L., Liu, C., Zhao, Q., Shi, Y., Miller, G., *et al.* (2011). Anaemia among students of rural China's elementary schools: Prevalence and correlates in Ningxia and Qinghai's poor counties. *Journal of Health, Population, and Nutrition, 29*(5), 471.

Malow, R. M., Kershaw, T., Sipsma, H., Rosenberg, R., & Dévieux, J. G. (2007). HIV preventive interventions for adolescents: A look back and ahead. *Current HIV/AIDS Reports, 4*(4), 173–180.

National Bureau of Statistics of China (2008). *China statistical yearbook*. China Statistics Press.

NBC (2012). *Expired milk and a piece of bread: School lunch scandal sparks outrage in China*. Available online: http://behindthewall.nbcnews.com/_news/2012/12/01/15591510-expired-milk-and-a-piece-of-bread-school-lunch-scandal-sparks-outrage-in-china?lite.

Ostojic, V., Cvoriscec, B., Ostojic, S. B., Reznikoff, D., Stipic-Markovic, A., & Tudjman, Z. (2005). Improving asthma control through telemedicine: A study of short-message service. *Telemedicine Journal & e-Health, 11*(1), 28–35.

Pop-Eleches, C., Thirumurthy, H., Habyarimana, J. P., Zivin, J. G., Goldstein, M. P., & de Walque, D. (2011). Mobile phone technologies improve adherence to antiretroviral treatment in a resource-limited setting: A randomized controlled trial of text message reminders. *AIDS, 25*(6), 825–834.

Popkin, B. M., Conde, W., Hou, N., & Monteiro, C. (2006). Is there a lag globally in overweight trends for children compared with adults? *Obesity, 14*(10), 1846–1853.

Prestwich, A., Perugini, M., & Hurling, R. (2009). Can the effects of implementation intentions on exercise be enhanced using text messages? *Psychology and Health, 24*(6), 677–687.

Rettie, R., Grandcolas, U., & Deakins, B. (2005). Text message advertising: Response rates and branding effects. *Journal of Targeting, Measurement and Analysis for Marketing, 13*(4), 304–312.

Rivera, J. A., González-Cossío, T., Flores, M., Romero, M., Rivera, M., Téllez-Rojo, M. M., *et al.* (2001). Multiple micronutrient supplementation increases the growth of Mexican infants. *The American Journal of Clinical Nutrition, 74*(5), 657–663.

Rodgers, A., Corbett, T., Bramley, D., Riddell, T., Wills, M., Lin, R. B., *et al.* (2005). Do u smoke after txt? Results of a randomised trial of smoking cessation using mobile phone text messaging. *Tobacco Control, 14*(4), 255–261.

Roediger, H. L., & Karpicke, J. D. (2006). Test-enhanced learning taking memory tests improves long-term retention. *Psychological Science, 17*(3), 249–255.

Sarter, M., Gehring, W. J., & Kozak, R. (2006). More attention must be paid: The neurobiology of attentional effort. *Brain Research Reviews, 51*(2), 145–160.

Shi, Y., Chang, F., Su, X., Luo, R., Zhang, L., & Rozelle, S. (2012). Parental training, anemia and the impact on the nutrition of female students in China's poor rural elementary schools. *China Agricultural Economic Review, 4*(2), 151–167.

Suggs, L. S. (2006). A 10-year retrospective of research in new technologies for health communication. *Journal of Health Communication, 11*(1), 61–74.

Sylvia, S., Luo, R., Zhang, L., Shi, Y., Medina, A., & Rozelle, S. (2013). Do you get what you pay for with school-based health programs? Evidence from a child nutrition experiment in rural China. *Economics of Education Review, 37*, 1–12.

Union IT (2010). *The world in 2010: ICT facts and figures*. International Telecommunication Union.

Zhang, L., Kleiman-Weiner, M., Luo, R., Shi, Y., Martorell, R., Medina, A., *et al.* (2013). Multiple micronutrient supplementation reduces anemia and anxiety in rural China's elementary school children. *The Journal of Nutrition, 143*(5), 640–647.

Zhang, J., & Mao, E. (2008). Understanding the acceptance of mobile SMS advertising among young Chinese consumers. *Psychology & Marketing, 25*(8), 787–805.

Zlotkin, S. H., Christofides, A. L., Hyder, S. Z., Schauer, C. S., Tondeur, M. C., & Sharieff, W. (2004). Controlling iron deficiency anemia through the use of home-fortified complementary foods. *The Indian Journal of Pediatrics, 71*(11), 1015–1019.

Zurovac, D., Sudoi, R. K., Akhwale, W. S., Ndiritu, M., Harner, D. H., & Rowe, A. K. (2011). The effect of mobile phone text-message reminders on Kenyan health workers' adherence to malaria treatment guidelines: A cluster randomised trial. *The Lancet, 378*(9793), 795–803.