





A Dataset of Vietnamese Junior High School Students' Reading Preferences and Habits

Quan-Hoang Vuong^{1,2,*}, Anh-Vinh Le³, Viet-Phuong La^{1,2}, Thu-Trang Vuong⁴,
Thu-Hang Do⁵, Ha-My Vuong⁶, Duc-Lan Do³, Phuong-Hanh Hoang³, Thi-Hanh Vu⁷,
Manh-Tung Ho^{1,2} and Manh-Toan Ho^{1,2,*}

¹ Center for Interdisciplinary Social Research, Phenikaa University, Ha Dong district, Hanoi 100803, Vietnam; phuong.laviet@phenikaa-uni.edu.vn (V.-P.L.); tung.homanh@phenikaa-uni.edu.vn (M.-T.H.)

² Faculty of Economics and Finance, Phenikaa University, Ha Dong district, Hanoi 100803, Vietnam

³ Vietnam National Institute of Educational Sciences, 101 Tran Hung Dao street, Hoan Kiem district, Hanoi 100000, Vietnam; leanhvinh@gmail.com (A.-V.L.); doduclan@gmail.com (D.-L.D.); hoangphuonghanh.hph@gmail.com (P.-H.H.)

⁴ Sciences Po Paris, Campus de Dijon, 21000 Dijon, France; thutrang.vuong@sciencespo.fr

⁵ Vuong & Associates, Dong Da district, Hanoi 100000, Vietnam; dohang.neu@gmail.com

⁶ Hanoi Amsterdam High School for the Gifted, Hoang Minh Giam street, Cau Giay district, Hanoi 100000, Vietnam; vuonghamy2003@gmail.com

⁷ School of Economics and International Business, Foreign Trade University, 91 Chua Lang street, Dong Da district, Hanoi 100000, Vietnam; hanhvt@ftu.edu.vn

* Correspondence: hoang.vuongquan@phenikaa-uni.edu.vn (Q.-H.V.); toan.homanh@phenikaa-uni.edu.vn (M.-T.H.)

Received: 20 March 2019; Accepted: 1 April 2019; Published: 2 April 2019



Abstract: Books are the invaluable, colossal storage of mankind's immense scholarship and are still commonly perceived as a more reliable source of knowledge even in this age of digitized information. Extensive reading is often promoted as being vital to cognitive development, especially for students in primary and secondary education. While it could now be considered common knowledge that reading is highly beneficial, reading habits vary among individuals in the same culture and receiving the same public education. This could be due to demographic variations and differences in socioeconomic status, or other factors such as family background and education. Despite the ample literature on reading habits, there still exists a lack of holistic approach with empirical results concerning the reciprocal interactions of reading and its relevant affecting factors. This data article presents a dataset of 1676 responses to the survey "Studying reading habits and preferences" of junior high school students in Vietnam. Result analysis facilitates evaluation of reading habits and their affecting factors, thus holding implications on education measures and policy. The dataset is available with the paper.

Dataset: The data is submitted as a supplementary file

Dataset License: CC-BY

Keywords: junior high school student; reading preferences; reading habit; Vietnam; STEM; K12 education; educational psychology; reading

1. Summary

The data article presents a dataset of 1676 responses from junior high school students from 8 schools in Northern Vietnam about their reading preferences and habits. The dataset was acquired from the first phase of the survey "Studying reading habits and preferences". The first phase ran from

December 2017 to January 2018. An expansion of the dataset will include data from 16 schools in total. The dataset contains information about junior high school students' reading behaviors such as preferred type of book, source of supply for books, actions after finishing a good book, and so on. Moreover, information about family background, academic performance, and future professional orientation is also acquired.

The dataset provides a thorough understanding of the reading habits and preferences of the current generation of Vietnamese junior high school students in particular. Based on the findings, we hope to generalize the implications on students in developing countries. Data analyses could enable understanding individual differences in perception of books and choices of book genres. The dataset would provide empirical evidence for the effect of reading on academic performance and future professional orientation. The dataset allows researchers as well as policy makers to evaluate some factors that influence reading preferences and wishes of students in order to make sound and scientific education policies.

In the next section, we will explain in detail all the coded variables of the dataset. Then, in the Methods section, the design of questionnaire, the survey process, the data analysis method, and examples of results are presented. Finally, the Conclusion will discuss the limitation and potential of the dataset.

2. Data Description

The data set contains 1676 responses acquired from the first phase of a survey about junior high school student's reading preferences and habits, parental influences, impacts of families' economic conditions, and the association between reading habits and academic performances and future professional decisions.

The questionnaire consists of multiple-choice questions and open-answer questions. For all multiple-choice questions, the surveyed is required to choose one single answer out of the options provided. Questions and answers were encoded into variables (discrete and continuous variables) and items of variables in the dataset. The questions were divided into two groups: (1) Questions on personal and family background, and (2) Questions on reading preferences and habits.

2.1. Group (1) Questions

The first two questions in group (1) are 'Grade' and 'Sex'. In Vietnam, junior high school covers grade 6th to grade 9th. The question on 'Sex' consists of two items: 'male' and 'female'. The following questions asked about the number of children in the family ('NumberofChi'), and the subjects' birth order in the family ('RankinF').

Distribution of answers is presented in Table 1. The ratio of male and female students joining the survey is approximately equivalent. Grade 6 is the modal category, occupying up to 28% of all the student body; however, the differences in share of student population among grades are insignificant. The families mainly have 2 to 3 children (~82%), and most of the students surveyed are first or second child (~78%).

The questions on academic performance asked about the average score of the most recent 45-minute tests of Mathematics, Physics, (Chemistry), and Biology, and the average score of the midterm examinations of Mathematics and proposed natural science subjects. The responses were encoded into two continuous variables: 'APS45' and 'APSVNEN'. Statistics show that 'APS45' range from 2.3 to 10 and 'APSVNEN' from 3.5 to 9.8; note that exams in Vietnam are graded on a scale of 0 to 10 with 10 being the highest score. Figure 1 is a histogram chart of 'APS45' in which the horizontal axis presents levels of performance and the vertical axis is the number of students. Each column is the number of students at a specific score. The number of students with below average score ($= < 5$ points) is little; and their columns have small height (occupying ~8%). For most students, the average score of the most recent 45-minute tests ranges from 6 to 8 points (~53.8%).

Table 1. Distribution of students according to demographic and reading behavior.

Coded Name	Explanation	Items	Frequency	Proportion
Grade	Current grade where students are studying	Grade 6	467	27.86%
		Grade 7	443	26.43%
		Grade 8	410	24.46%
		Grade 9	356	21.24%
Sex	Biological sex of the students	Male	853	50.89%
		Female	823	49.11%
RankinF	Subjects' birth order in the family	First child	695	41.47%
		Second child	617	36.81%
		Third child	251	14.98%
		Forth child	79	4.71%
		Fifth child	26	1.55%
		Sixth child	5	0.30%
		Seventh child	3	0.18%
NumberofChi	Number of children in the family	1 child	78	4.65%
		2 children	744	44.39%
		3 children	629	37.53%
		4 children	159	9.49%
		5 children	45	2.68%
		6 children	15	0.89%
		7 children	5	0.30%
		8 children	1	0.06%
Source	Main source of supply for books	Personal expense/Parents	617	36.81%
		Borrowed from friends/library	1016	60.62%
		Gift/Reward	43	2.57%
PrioAct	Prioritized actions when coming across interesting content	Tell friends or family	796	47.49%
		Note down	349	20.82%
		Apply the content to daily life	111	6.62%
		Reflect and connect with personal knowledge	420	25.06%
AftAct	Actions after finished a good book	Find more books with the same topic	825	49.22%
		Find books on related topics	277	16.53%
		Find books on new topics	242	14.44%
		Read again	332	19.81%
EncourAct	Preferences on joining reading promotion activities	Yes	1493	89.10%
		No	183	10.90%
MostlikedAct	Favorite book-related activities and events	Book exhibition	637	39.69%
		Storytelling competition	427	26.60%
		Story-writing competition	187	11.65%
		Creating book illustrations	354	22.06%
Bookcase	Evaluating classrooms' public bookshelves	Diverse and interesting	637	38.01%
		Missing good titles	455	27.15%
		Lacking books	225	13.42%
		No bookshelf	359	21.42%

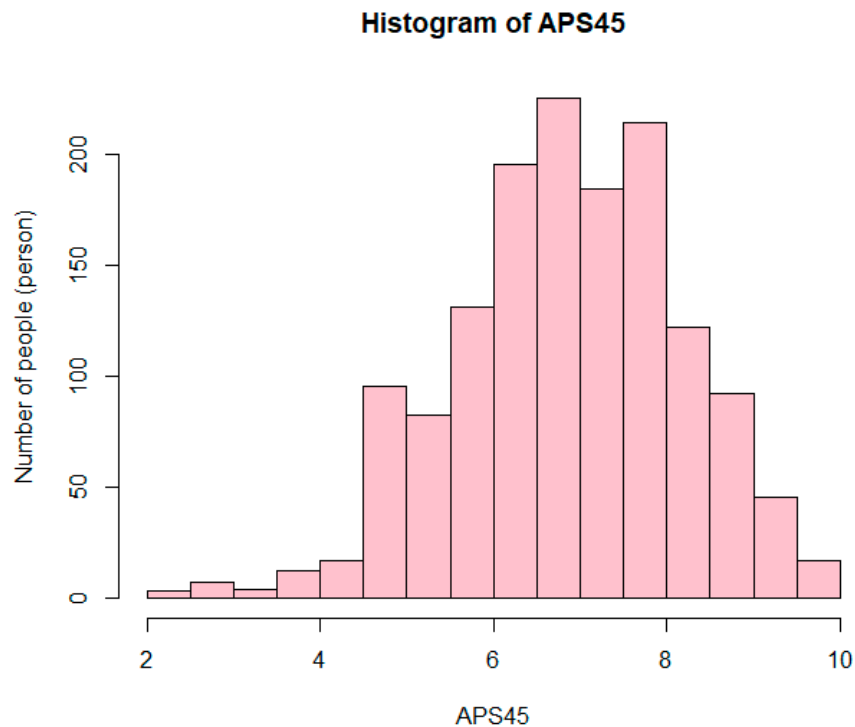


Figure 1. Average score of the most recent 45-minute tests of Mathematics, Physics, (Chemistry), and Biology.

Next, we investigated the future professional orientation of students through the question: “In the future, which job would you like to do the most?” (‘FutureJob’). A variety of professions was proposed by the students such as doctor, policeman, scientist, teacher, and so on. Depending on particular research purposes, the professions will be categorized into different groups.

We also examined information about students’ parents: age, academic level, profession of the father (‘AgeFat’, ‘EduFat’, ‘CareerFat’); and, similarly, age, academic level, profession of the mother (‘AgeMot’, ‘EduMot’, ‘CareerMot’). Academic level contains 4 items: Under high school (‘UnderHi’), high school (‘Hi’), Undergraduate (‘Uni’), Graduate School (‘PostGrad’). Age of the parents is coded as continuous variables. The highest age of a father is 70 years old, while for the mother it is 76 years old. ‘EcoStt’ shows the economic condition of the family as perceived by the students themselves, consisting of 3 categories: Poor (‘poor’), Average (‘med’), and Rich (‘rich’). For further detail, we investigated whether or not the students were aware of the average monthly income of their household (‘KnowledgeInc’); the answers consist of ‘yes’ or ‘no’. 185 out of 1676 students were able to estimate said average monthly income, in which case their estimation is recorded in the variable ‘EstIncome’ (unit: million VND).

2.2. Group (2) Questions

Group (2) questions aim to examine the reading preferences and habits of students. After the question: “Do you like reading?” (‘Readbook’) with two choices ‘yes’ or ‘no’, students are invited to express their reading preferences by choosing from a list of reading topics (‘Topic’) which includes: math - physics (‘math.phy’), literature (‘literality’), foreign languages (‘language’), natural sciences, chemistry, and biology (‘nat.chem.bio’), history and geography (‘his.geo’), information technology (‘tech’). There are 79 students who did not choose any of the provided topics, whose responses are encoded as ‘notans’. Figure 2 shows the distribution according to preferred reading topics. Each column represents a topic and the number of students interested in said topic. Accordingly, most of the students enjoy literature books (around 500 out of 1676 students). Moreover, each column was divided into two colors: blue shows the number of students who enjoy the activity of reading, while

the orange represents students who do not like to read. More than 90% of the students claim to enjoy reading, illustrated by domination of the blue area over orange. On the other hand, most of the 79 students who did not give an answer on their preferred reading topic did not like reading.

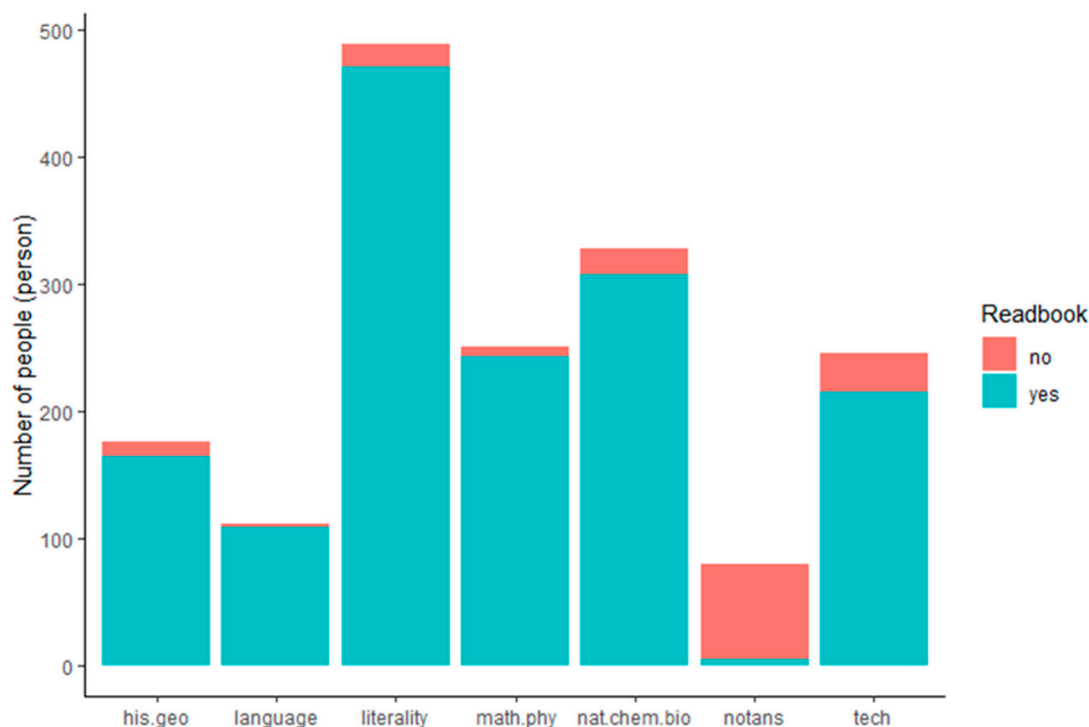


Figure 2. Distribution of students according to their preferred reading topics and their reading hobby.

In order to evaluate the reading habits of the students, we surveyed the students on their daily pastimes. The ‘Hobby’ variable represents activities that students like to do the most during their free time: reading (‘a’), watching TV or listening to music (‘b’), doing housework or helping at the farm (‘c’), observing nature (‘d’), hanging out with friends and family (‘e’), and other (‘f’). Notably, while Figure 2 shows that a large proportion of students like reading, Figure 3 reveals that only 331 out of 1676 students (20% of the sample) consider reading their favorite hobby. Item ‘b’ is the modus among items of the ‘Hobby’ variable, meaning that students prefer watching TV or listening to music (up to 47%) to reading. Similarly to Figure 2, columns in Figure 3 were also divided into two colors: The blue coded for male and the orange coded for female. It can be observed that more female students like reading, doing housework, or helping at the farm, while male students prefer watching TV, listening to music, and hanging out with friends.

Daily time spent reading science books (‘TimeSci’) and daily time spent reading literature and social sciences-related books (‘TimeSoc’) were divided into three items: ‘less30’ for reading time under 30 m, ‘b3060’ for reading time from 30 m to 60 m, and ‘g60’ for more than 60 m. All durations are personal estimations provided by the students themselves. 50% of the students spend more than 30 m/day on science books; while for literature and books related to social sciences, only 36% of the students sending more than 30 m/day reading them. Besides personal preferences and passions, the amount of time spent reading also depends on a person’s literacy [1].

Students were also asked to give a ‘yes’ or ‘no’ response for the question as to whether their parents buy them books or not (‘Buybook’); and whether their parents read them stories at home or not (‘Readstory’). Scientists have found that parents reading for children associates with the development of children’s language, reading, and writing ability [2–5]. This dataset shows that while 1447 students were given books by their parents, only 424 students have ever been read or told a story by their parents.

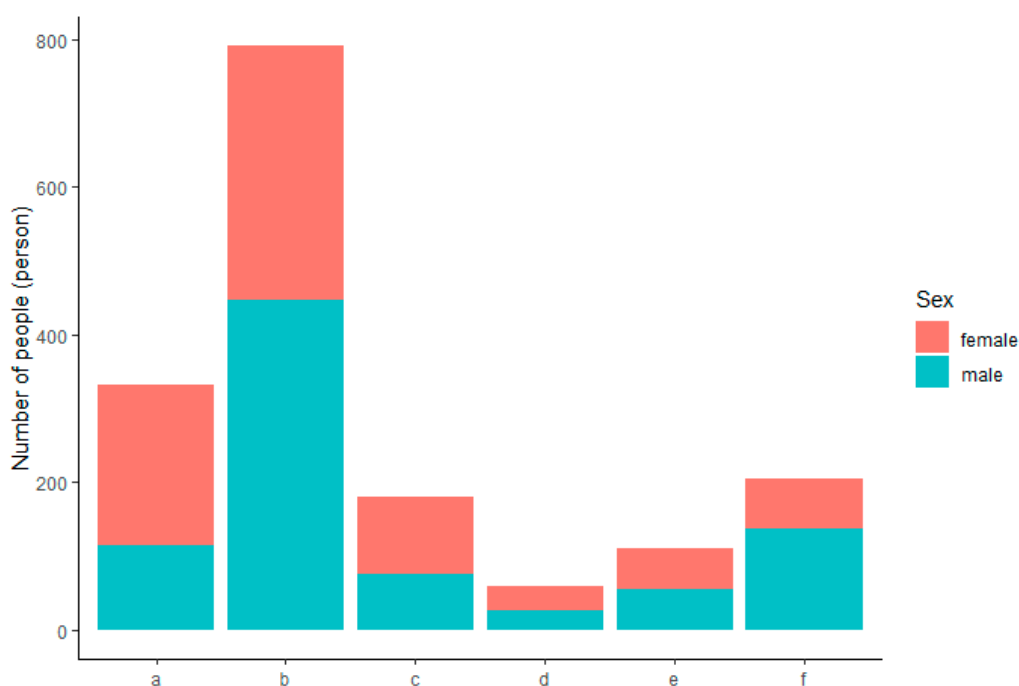


Figure 3. Distribution of students according to favorite hobbies.

We further examined the reading preferences of the students through the question: “Besides textbooks, if someone offers to gift you a book, what kind of book would you choose?”. The nominal variable is ‘Typebook’ and the choices were Novel (‘a’), Biography (‘b’), Popular Science (‘c’), Arts (‘d’), Vocational instruction (‘e’), and Other (‘f’). Up to 41% of the students chose Novel, while only 7% chose Biography. The rest of the proposed choices are almost equally shared at around 10% each. Regarding the reason for their book choice (‘Reason’), the following options are proposed: ‘a’ = personal preferences, ‘b’ = recommended by parents, ‘c’ = recommended by teachers/friends, ‘d’ = serendipity. Most of the students (up to 85%) choose books based on their personal preferences.

Table 1 provides more information cultivated from the datasets. We asked students about their source of book supply (‘Source’). Most students (approximately 61%) access books by borrowing from friends or libraries (‘borrow’). Other common sources include buying books using their own or their parents’ money (‘buy’) (about 37%), and receiving books as gifts or rewards (‘gift’) (about 25%).

Simply reading a book is one thing, but thoroughly understanding the message and implications is a different task, especially when it concerns younger audiences. Many studies have provided significant insights on critical reading and on how to read properly [6,7]. Each individual has a unique way to read critically, and the differences reflect the ways in which the brain processes and stores information [8]. We aim to explore how students understand a book and react to its content by questioning the first action undertaken by the student when they come across a piece of content that piques their interest (‘PrioAct’). The options proposed are: telling friends or family (‘a’), noting (‘b’), applying the content to daily life (‘c’), reflecting and relating to personal knowledge (‘d’). In addition, variable ‘AftAct’ shows their first action after finishing a good book: ‘a’, find more books on the same topics; ‘b’, find books about related topics; ‘c’, find books on new topics; ‘d’, read the book again. Most of the students would tell friends or parents when they come across something interesting, and find more books on the same topics after finishing a good book.

In order to understand students’ wishes regarding books, we asked the them to answer the following questions: (1) ‘EncourAct’ – “Do you like participating in reading promotion activities?” with the answers are ‘yes’ or ‘no’; (2) ‘MostlikedAct’ – If yes, then which of the following activities would you rather take part in: Book exhibition (‘a’), Storytelling competition (‘b’), Story-writing competition (‘c’), Drawing book illustrations (‘d’); (3) ‘Bookcase’ – “How would you describe the common bookshelf

in your classroom?" Diverse and interesting ('a'), Missing good titles ('b'), Lacking books ('c'), No bookshelf ('d'). The statistics show that around 90% of the students enjoy participating in reading promotion activities. However, 71 students did not answer question (2); of the total 1605 responses, book exhibition occupied around 40%.

Finally, there are two open-answer questions: "Name two books you like the most" ('Read.like'); and "Name three books you really wanted to read but have never had the chance to because they are not available." (Notread.like').

2.3. Potential Research Questions

Reading is a healthy hobby that contributes positively to vocabulary learning aptitude, reading and writing skills, and cognitive development [9–11]. Specifically, a study examining 288 third grade students reported that despite similarities in their self-evaluation of reading ability, young female students value reading more than their male counterparts [12]. Elementary students show more interest in reading than secondary or high school students [13]. Moreover, researchers have highlighted the lack of children's access to sciences and philosophy books despite their role on children's development of scientific perception [14].

Based on the dataset, we present potential research questions in following list.

Possible research questions arising from the data set:

- What socio-economic and socio-cultural factors that could affect reading habits of students?
- How habits, preferences and wishes of students influence the efficacy of reading promotion campaigns and events?
- How does the birth order associate with reading habits?
- How does reading habits associate with academic performance and professional orientation?
- Is there any difference in reading habits conditioned on biological sex? For example, do female students behave differently when coming across interesting content?
- How does students' perception of classrooms' bookshelves associate with reading habits?

3. Methods

To the best of our knowledge, there is presently no thorough dataset containing adequate information on the reading behavior of Vietnamese students. Therefore, we conducted a survey named "Studying the reading preferences and habits of students" at 16 junior high schools in North Vietnam. The logic for designing the study is similar to what is described in [15–18]. The survey was divided into two phases, and this dataset was acquired in the first phase. The first phase started in December 2017 and finished with 8 schools on January 2018. The list of high school students is provided in Table A1. In the first phase, we collected 1676 responses for this dataset.

The subjects were selected randomly and participated by giving their responses to the questionnaires. Examples of the questionnaires were provided in Supplementary Files S1–S4. Surveyed students were required to write down their school name before answering the questions. The structure of the questionnaires consists of 5 parts with 4 to 7 questions in each part. Part 1: Personal information; Part 2: Family information; Part 3: Habit/Preference; Part 4: Book choosing habits; Part 5: Questions on classroom's bookshelf. Questions can be either multiple-choice with only one single answer allowed, or open-answer.

The participants are junior high school students, so we constructed short and easy to understand questions, with the request for students to answer as truthfully and accurately as possible. After collecting the questionnaires, two members of the research team input the data into MS Excel for validation. Hence, the dataset is perfectly matched with the answers in the questionnaire.

The project consists of 5 periods: (1) Questionnaire design; (2) Surveying the students; (3) Validating and checking the quality of the answers; (4) Data design and input; (5) Mining the data and analyzing the results.

3.1. Materials and Methods

Raw data from the questionnaire were input as an MS Excel file (see Supplementary File S5). The data will be saved as CSV (see Supplementary File S6) format for processing in R. The analysis employed BCL model [19].

The majority of the dataset is discrete data, which received their value from categories that were built according to the design of the survey. Thus, the most suitable method of analysis for this kind of data is categorical regression, in which log-linear and logistic models can be equivalent [20]. However, logistic regression model is more efficient in explaining the relationship, either independence or associations, among variables. Moreover, logistic regression analysis also provides coefficients to estimate the probabilistic trends for each value of the dependent variables according to the conditions of the independent variables. Logistic regression analysis is more flexible for analyzing mixed sets of nominal/ordinal and interval variables [21]. For more detailed discussion on the comparative advantages of the two methods of analysis, please see [19–21].

The general equation for the logistic model is:

$$\ln \frac{\pi_j(\mathbf{x})}{\pi_J(\mathbf{x})} = \alpha_j + \beta_j^T \mathbf{x}, j = 1, \dots, J - 1$$

In which, \mathbf{x} is independent variables; and $\pi_j(\mathbf{x}) = P(Y = j|\mathbf{x})$ is the corresponding probability. Therefore, $\pi_j = P(Y_{ij} = 1)$ with Y is the dependent variable.

In the logit model, the probability of each item of dependent variables was estimated as follows:

$$\pi_j(\mathbf{x}) = \frac{\exp(\alpha_j + \beta_j^T \mathbf{x})}{1 + \sum_{h=1}^{J-1} \exp(\alpha_h + \beta_h^T \mathbf{x})}$$

Regression on the statistical software R was run from the distribution table of frequency of factors under the format CSV, similar to Supplementary Files S7 and S8 in the dataset. The usage of this file was meant to adjust the based items for each variable. The model can be run on the original CSV file, however, based item is at default and cannot be changed.

3.2. Examples of Data Analysis

Logistic regression method is employed with the dependent variable being reading habits ('Readbook'), and biological sex ('Sex') and school grade ('Grade') as independent variables. The regression coefficients are reported in Table 2.

Table 2. Estimating of independent variables 'sex' and 'grade' against dependent variable 'Readbook'.

	Intercept	"Sex"		"Grade"	
	β_0	"Male" β_1	"Grade7" β_2	"Grade8" β_3	"Grade9" β_4
Logit (yes no)	3.888 *** [14.147]	-1.124 *** [-6.011]	-1.125 *** [-3.986]	-0.786 ** [-2.648]	-1.650 *** [-6.003]

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1, z-value in square brackets; baseline category for: "Sex": "Female"; and, "Grade": "Grade6". Residual deviance: 4.931 on 3 degrees of freedom. Log-likelihood: -20.309 on 1.

Examples of code on R that were used to come up with the results in Table 2 is as follows:

```
> tab=read.csv("D:/VA/Data in Brief/tab.csv",header=T)
> attach(tab)
> fit.md=vglm(cbind(yes,no)~Sex+Grade,data=tab,family=multinomial)
> summary(fit.md)
```

More example of R code is provided in Supplementary File S9.

Regression model:

$$\ln \frac{\pi_{yes}}{\pi_{no}} = 3.888 - 1.124 \times Male - 1.125 \times Grade7 - 0.786 \times Grade8 - 1.650 \times Grade9$$

Based on the model, estimation of conditional probability produced Figure 4 (Table of detail probability is in Table A2).

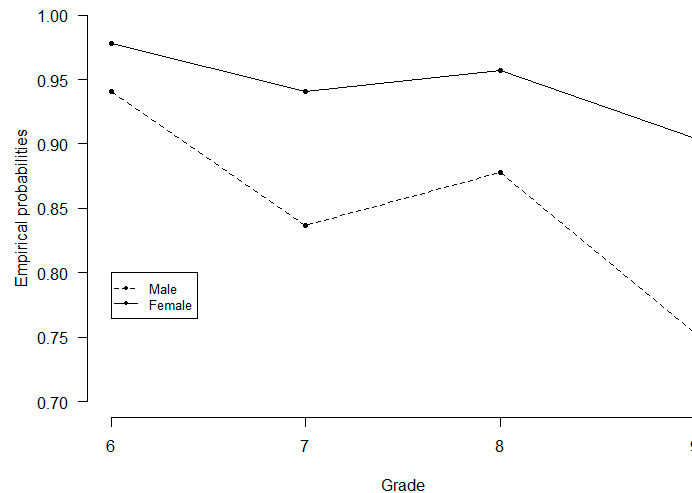


Figure 4. Probability to like reading according to biological sex and grade.

Table 3 presents another example of estimating impacts of socio-economic conditions and biological sex on students’ reading habits. Dependent variable is students’ reading habits (‘Readbook’); and independent variables are economic conditions (‘EconStt’) and biological sex (‘Sex’).

Table 3. Estimating the impact of biological sex (‘Sex’) and economic condition (‘EcoStt’) on reading habits (‘Readbook’).

	Intercept	“Sex”	“EcoStt”	
		“Male”	“poor”	“rich”
	β_0	β_1	β_2	β_3
Logit (yes no)	2.860 *** [17.936]	-1.107 *** [-5.980]	-0.014 [-0.045]	0.464 [1.529]

Signif. codes: 0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1, z-value in square brackets; baseline category for: “Sex”: “Female”; and, “EcoStt”: “med”. Residual deviance: 3.191 on 2 degrees of freedom.

4. Conclusions

The dataset provides a comprehensive view on the reading practices, habits, and preferences of students in Vietnam, which is a subject that to this day lacks systematic research despite being acknowledged as a crucial aspect of education. The inclusion of both self-reported personal preferences and objective data regarding school performance under the form of notes allows for multifaceted analyses. In addition to data regarding the activity of book reading, we have also collected demographic data and recorded the socio-economic background of the respondents, including estimated household income and parents’ profession. We believe this dataset has one of the most complete sets of variables out of all datasets in the same field in Vietnam, which would be especially beneficial in future models that require extensive control variables.

In terms of content, our dataset promises fertile grounds for future research on reading practices in adolescents both as a response variable—to demographic factors, family background, or personal preferences, for example—and a predictor variable—to, among others, academic performance or

professional orientation. The insights gained into the reading habits and preferences in Vietnamese pupils would serve as grounded suggestions for future policies and informed decisions [22], namely education, but also social policies geared towards families in rural areas and low-income households. Policy implications should concern not also Vietnam but perhaps also other emerging countries similar in economic development, or other East Asian/Southeast Asian countries with similar cultural attributes and family structures.

We acknowledge, however, that the dataset faces some limitations. First, the geographical area where the survey was conducted is limited to one province in North Vietnam. In the future, we hope to expand the dataset by broadening the surveyed geographical area to allow for more comparative perspectives between different regions; for example, urban versus rural. Second, the school grade range in the dataset could also be expanded to include primary school for a more panoramic view of reading habits among the youth, and potential investigations into differing habits between generations.

Moreover, there are values of the categorical variables that can be divided into two values. For instance, the categorical variable ‘Source’ now contains three values: borrowing from friends or libraries (‘borrow’), using their own or their parents’ money (‘buy’), and receiving books as gifts or rewards (‘gift’). The ‘borrow’ value can be divided into borrowing from friends and borrowing for libraries. Such an amendment in the categorization of data will provide a better insight into the usage of libraries. However, we can only address this in a future survey.

Supplementary Materials: The following are available online at <http://www.mdpi.com/2306-5729/4/2/49/s1>, Supplementary Files S1–S4: Survey questionnaire 1/2/3/4.pdf; Supplementary File S5: LVP_survey_180629.xlsx; Supplementary File S6: Data Survey_Readinghabit.csv; Supplementary File S7: tab.csv; Supplementary File S8: tab1.csv; Supplementary File S9: 20190221_CodeR.doc.

Author Contributions: Conceptualization, Q.H.V. and A.V.L.; methodology, Q.H.V., V.P.L.; software, V.P.L.; validation, Q.H.V., A.V.L., H.M.V., and V.P.L.; formal analysis, T.H.D.; data curation, M.-T.H. (Manh-Toan Ho), H.M.V., and V.P.L.; writing—original draft preparation, T.H.D., M.T.H. (Manh-Toan Ho), and P.H.H.; writing—review and editing, M.-T.H. (Manh-Toan Ho), M.-T.H. (Manh-Tung Ho), T.T.V., D.L.D., T.H.V., and H.M.V.; visualization, T.H.D.; supervision, Q.H.V. and A.V.L.; project administration, Q.H.V., A.V.L., and M.-T.H. (Manh-Toan Ho).

Funding: This research received no external funding.

Acknowledgments: We want to send our gratitude to research staff of Vuong & Associates (Hanoi, Vietnam) for assisting in collecting data, especially Do Minh Trang and Dam Thu Ha. Our appreciations also go on to personnel of junior high schools and provincial departments that provided supports during the survey.

Conflicts of Interest: The authors declare no conflict of interest.

Appendix A :

Table A1. List of surveyed junior high schools.

No.	School	Number of Surveyed Students	Female	Male
1	Yên Hòa junior high school	348	51.72%	48.28%
2	Khánh Dương junior high school	133	45.11%	54.89%
3	Khánh Thượng junior high school	281	54.09%	45.91%
4	Yên Mỹ junior high school	187	44.39%	55.61%
5	Yên Phong junior high school	221	48.87%	51.13%
6	Yên Thái junior high School	246	47.56%	52.44%
7	Khánh Thịnh junior high School	91	48.35%	51.65%
8	Mai Sơn junior high School	170	47.06%	52.94%

Expanded list of schools from the second phase: (9) Yên Thắng junior high School. (10) Yên Nhân junior high School. (11) Yên Đồng junior high School. (12) Yên Tử junior high School. (13) Yên Thành junior high School. (14) Yên Hưng junior high School. (15) Vũ Phạm Khải junior high School. (16) Yên Lâm junior high School

Appendix B :

Table A2. Probability of like/dislike reading according to biological sex and grade.

'Sex' 'Grade'	'Readbook'							
	'Yes'				'No'			
	'Grade6'	'Grade7'	'Grade8'	'Grade9'	'Grade6'	'Grade7'	'Grade8'	'Grade9'
'Male'	0.941	0.837	0.878	0.753	0.059	0.163	0.122	0.247
'Female'	0.978	0.941	0.957	0.904	0.022	0.059	0.043	0.096

References

1. Perfetti, C.A. *Reading Ability*; Oxford University Press: Oxford, UK, 1985. [CrossRef]
2. Cheng, K.-H.; Tsai, C.-C. Children and parents' reading of an augmented reality picture book: Analyses of behavioral patterns and cognitive attainment. *Comput. Educ.* **2014**, *72*, 302–312. [CrossRef]
3. Farrant, B.M.; Zubrick, S.R. Early vocabulary development: The importance of joint attention and parent-child book reading. *First Lang.* **2011**, *32*, 343–364. [CrossRef]
4. Manolitsis, G.; Georgiou, G.K.; Tziraki, N. Examining the effects of home literacy and numeracy environment on early reading and math acquisition. *Early Child. Res. Q.* **2013**, *28*, 692–703. [CrossRef]
5. Whitehurst, G.J.; Falco, F.L.; Lonigan, C.J.; Fischel, J.E.; DeBaryshe, B.D.; Valdez-Menchaca, M.C.; Caulfield, M. Accelerating language development through picture book reading. *Dev. Psychol.* **1988**, *24*, 552–559. [CrossRef]
6. Fountas, I.C.; Pinnell, G.S.; Le Verrier, R. *Guided Reading*; Heinemann: Portsmouth, UK, 2001.
7. Smith, F. *Understanding Reading: A Psycholinguistic Analysis of Reading and Learning to Read*; Routledge: London, UK, 2004. [CrossRef]
8. Daneman, M.; Carpenter, P.A. Individual differences in working memory and reading. *J. Verbal Learn. Verbal Behav.* **1980**, *19*, 450–466. [CrossRef]
9. Krashen, S.D. *The Power of Reading: Insights from the Research*; ABC-CLIO: Stan Barbara, CA, USA, 2004.
10. Hargrave, A.C.; Sénéchal, M. A book reading intervention with preschool children who have limited vocabularies: The benefits of regular reading and dialogic reading. *Early Child. Res. Q.* **2000**, *15*, 75–90. [CrossRef]
11. Sikora, J.; Evans, M.; Kelley, J. Scholarly culture: How books in adolescence enhance adult literacy, numeracy and technology skills in 31 societies. *Soc. Sci. Res.* **2019**, *77*, 1–15. [CrossRef] [PubMed]
12. Marinak, B.A.; Gambrell, L.B. Reading motivation: Exploring the elementary gender gap. *Lit. Res. Instr.* **2010**, *49*, 129–141. [CrossRef]
13. Anderson, M.A.; Tollefson, N.A.; Gilbert, E.C. Giftedness and reading: A cross-sectional view of differences in reading attitudes and behaviors. *Gift. Child Q.* **1985**, *29*, 186–189. [CrossRef]
14. Pigliucci, M. The Case for More Science and Philosophy Books for Children. Available online: <http://nautil.us/blog/-the-case-for-more-science-and-philosophy-books-for-children> (accessed on 21 March 2019).
15. Vuong, Q.-H. Survey data on Vietnamese propensity to attend periodic general health examinations. *Sci. Data* **2017**, *4*, 170142. [CrossRef]
16. Vuong, Q.H. Survey data on entrepreneurs' subjective plan and perceptions of the likelihood of success. *Data Brief* **2016**, *6*, 858–864. [CrossRef]
17. Vuong, Q.-H.; Nguyen, T.-K. Data on Vietnamese patients' financial burdens and risk of destitution. *Data Brief* **2016**, *9*, 543–548. [CrossRef] [PubMed]
18. Vuong, Q.H. Data on Vietnamese patients' behavior in using information sources, perceived data sufficiency and (non)optimal choice of health care provider. *Data Brief* **2016**, *7*, 1687–1695. [CrossRef] [PubMed]
19. Agresti, A. *Categorical Data Analysis*, 3rd ed.; Wiley: Hoboken, NJ, USA, 2013.
20. Fienberg, S.E.; Junker, B.W. Categorical Data Analysis. In *International Encyclopedia of Education*, 3rd ed.; Peterson, P., Baker, E., McGaw, B., Eds.; Elsevier: Oxford, UK, 2010; pp. 58–65. [CrossRef]
21. Tansey, R.; White, M.; Long, R.G.; Smith, M. A comparison of loglinear modeling and logistic regression in management research. *J. Manag.* **1996**, *22*, 339–358. [CrossRef]

22. Vuong, Q.H. The (ir)rational consideration of the cost of science in transition economies. *Nat. Hum. Behav.* 2018, 2, 5. [[CrossRef](#)] [[PubMed](#)]



© 2019 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).