



Collaborative learning in the Vietnam Escuela Nueva Model and students' learning behaviors: A mixed methods longitudinal study

Khoa Vu^{a,*}, Vu Dao^b, Joan DeJaeghere^b, Paul Glewwe^a

^a Department of Applied Economics, University of Minnesota, United States

^b Department of Organizational Leadership, Policy, and Development, United States

ARTICLE INFO

Keywords:

Vietnam
Basic education
Escuela Nueva
Learning behaviors
Peer learning

ABSTRACT

We study how Vietnam's Escuela Nueva, a pedagogical reform that promotes participatory and collaborative learning in primary schools, affects students' learning behaviors in short and long run. Using a propensity score matching approach, we find that the model increases the likelihood of giving and receiving feedback from peers as well as asking questions in class. The peer-learning effects appear to persist in the long run as students enter lower secondary school. Qualitative interviews with lower secondary teachers and school principals also provide evidence of these long-term effects. However, teachers also highlighted important concerns of this model.

1. Introduction

Escuela Nueva (EN) is a schooling model that embraces the popular learner-centered paradigm by promoting participative and collaborative learning; it is widely considered a promising approach to improve access and quality of education in rural areas. Escuela Nueva started in Colombia in the 1970s and has since been implemented in 14 countries (Le, 2018b). The model includes features across several domains, including peer interaction, teaching methods, and integration between schools and the community.¹ While much has been written about the impacts of this model on learning outcomes,² we know little about the channels through which the model shapes students' learning outcomes.

Vietnam offers a unique opportunity to better understand the mechanism of this learner-centered approach. In 2012, over a thousand primary schools across Vietnam adopted a version of this model known as the Vietnam Escuela Nueva (VNEN) program. This curriculum and pedagogy model was widely taught up until 2017, when it was replaced by a national competency-based curriculum that started in 2018. Previous evaluations of VNEN have found positive impacts on students' cognitive and non-cognitive outcomes (Parandekar et al., 2017; Dang et al., 2022).³ One possible explanation for these learning gains is that

the adapted EN curriculum and pedagogy promotes classroom participation and peer learning, which helps students learn more effectively. Peers play a major role in students' learning outcomes (Paloyo, 2020), often through group interaction and motivation (Eisenkopf, 2010; Feld and Zölitz, 2017). Thus, it is important to understand whether this pedagogy effectively promotes such collaborative and peer-learning behaviors.

We study how the VNEN's teaching practices affect students' participatory and peer-learning behaviors in primary schools. Since students who learned the VNEN curriculum in primary schools had to go back to the traditional curriculum in lower secondary schools, we also examine how these behavioral shifts persist in secondary schools. We draw data from two previous impact evaluation studies on the VNEN program (Parandekar et al., 2017; Dang et al., 2022). Using detailed teacher questionnaire data, we first document the extent to which different Escuela Nueva practices were implemented inside the Vietnamese classrooms in 2014–2015 and 2015–2016, including group seating arrangement, and teaching practices that encourage peer learning and classroom engagement.

We then use a propensity score matching approach to evaluate the impacts on students' learning behaviors in the short run between 2014

* Correspondence to: 231 Ruttan Hall, 1994 Buford Ave, St. Paul, MN 55108, United States.

E-mail address: vuxxx121@umn.edu (K. Vu).

¹ See Kline (2002) and Parandekar et al. (2017) for detailed discussions about the design of the Escuela Nueva model and how its features are related to the relevant educational literature.

² For an extensive review on the Escuela Nueva model, see Le (2018b). For the VNEN model, see Parandekar et al. (2017), Le (2018a, b) and Dang et al. (2022).

³ Dang et al. (2022) found that the program increased cognitive skills by 0.15–0.21 standard deviations and non-cognitive skills by 0.15–0.22 standard deviations. Parandekar et al. (2017) found that changes in math test score increased by 18 points and Vietnamese test score by 15.9 points when a school moved from not implementing the VNEN model to fully implementing the model.

and 2015 (Grade 4) and 2015–2016 (Grade 5) when students were learning VNEN in primary school.⁴ We also estimate the long-run impacts on the same outcomes on the same sample of students between 2018 and 2019 (in Grade 7) and 2020–2021 (Grade 9) when the primary students went to lower secondary school. We also analyze qualitative data from teachers and principals about their perceptions of students’ learning behaviors at the secondary level.

We contribute to an expanding literature on the Escuela Nueva model, its pedagogical practices and effects. Early studies use linear regressions and find mixed evidence for the effects of the original Escuela Nueva model on cognitive and non-cognitive skills in Colombia (Psacharopoulos et al., 1993; McEwan, 1998; Forero-Pineda et al., 2006). A more recent study by Hammmer (2017) applies a multilevel approach and finds positive impacts of the model on test scores. In Vietnam, the Escuela Nueva model is argued to deviate from its approach in Colombia as it is too dependent on textbook and teachers’ teaching, much like the traditional teacher-centered model before VNEN was adopted (Le, 2018a, b). Le (2018b) documented instances where teachers switched to teacher- and textbook-centered practices when students had a hard time following the VNEN model. Parandekar et al. (2017) and Dang et al. (2022) show that the model still has positive effects on cognitive and non-cognitive outcomes of students in the short run.

Our study also highlights the importance of learning behaviors within a broader literature on pedagogical reforms in developing countries,⁵ especially those that push for learner-centered approach like the Escuela Nueva model.⁶ Some of this literature, and particularly econometric studies, tend to focus more on class attendance, grade completion, cognitive test score, and non-cognitive skills as outcomes, and less on how students interact with and learn from peers or how they participate in lessons. An extensive body of literature also focuses on the role of peers and peer interaction in learning,⁷ but this literature has not been discussed in relation to the EN model. Our results suggest that peer interaction and learning can be an important channel through which educational reforms can improve learning, especially since these behaviors appear to persist over time. We also document potential challenges to collaborative and peer learning in education reforms.

The rest of the paper is organized as follows. Section 2 provides a brief overview of the implementation of the Vietnam Escuela Nueva and how schools implement the program. In Sections 3 and 4, we explain our quantitative and qualitative methods, data and analyses. In Section 5, we discuss our findings, and in Section 6, we draw our policy implications from those findings.

2. Overview of the VNEN Model

The Vietnam Escuela Nueva (VNEN) model was adopted in Vietnam in 2010 via a \$84.6 million project funded by the Global Partnership for Education and supported by the World Bank (Le, 2018a). The program was first piloted in 24 schools in 2010 before being widely adopted in the summer of 2012. About 1447 out of Vietnam’s 14,000 public primary schools chose to adopt the VNEN model during this period. VNEN schools were also offered funding to implement the program.

⁴ While the original data are also available in 2013–2014 and 2015–2016, the relevant learning behaviors were not surveyed in these years.

⁵ See Kremer et al. (2013), Glewwe and Muralidharan (2016), and Evans and Mendez Acosta (2021) for a summary of educational reforms and interventions in developing countries.

⁶ See, e.g., Schweisfurth (2011), (2013), Tabulawa (2013); Bremner et al. (2022), for extensive discussions on learner-centered paradigm and reforms in developing countries.

⁷ Carrell et al. (2013), Sojourner (2013), Booji et al. (2017) documented the effects of peers’ ability on learning outcomes. Also see Paloyo (2020) for an extensive review on the peer effects literature.

The selection into the VNEN model is as follows. First, 63 provinces were classified as most disadvantaged (“Priority I”), somewhat disadvantaged (“Priority II”), and least disadvantaged (“Priority III”). Each province was allowed a number of VNEN schools based on their level of disadvantage, with most VNEN schools being in the Priority I provinces. Provinces then offered the VNEN program to disadvantaged schools based on (i) the proportion of ethnic minority students, (ii) being located in remote areas, and/or (iii) being located in low-income communes. Most schools accepted the offer since it came with funding (Dang et al., 2022).

In this section, we summarize the differences in designs between the VNEN model and the traditional school model. We then use teacher questionnaire data that the World Bank collected in the 2014–2015 and 2015–2016 school years to document how these features were implemented by teachers in both VNEN and non-VNEN schools. Details on the data can be found in Section 3.

2.1. Classroom organization

A VNEN classroom is organized differently than a traditional classroom as the new model prioritizes group learning and engaging students with real-life objects and their communities. In Table 1, we summarize these classroom organization features and the extent to which they were adopted by VNEN and non-VNEN schools in 2014 and 2015.

By design, Escuela Nueva classrooms need to have sufficient space for teachers and students to move around for group activities (Parandekar et al., 2017). The new model also frequently uses group seating, as opposed to lecture style in traditional classrooms, with all desks in row facing the front of the classroom. We observe that although VNEN and non-VNEN schools have similar classroom area (square meters) per student, the seating arrangements are indeed very different across the two models. Over 91% of VNEN schools report having group seating most of the time, while only over 21% of non-VNEN schools report having a similar arrangement. Surprisingly, the share of non-VNEN schools reporting having group seating most of the time increases from 21.3% to 26.6% although they are not required to do this. In

Table 1
Classroom organization by VNEN status and year.

	Non-VNEN	2014 VNEN	Diff.	Non-VNEN	2015 VNEN	Diff.
Classroom area and seating						
Class area per student (in log)	0.61	0.62	0.01	0.65	0.64	-0.01
Group seating (1/0)	0.21	0.91	0.70***	0.27	0.92	0.66***
Lecture seating (1/0)	0.56	0.07	-0.49***	0.54	0.08	-0.46***
VNEN artifacts (1/0)						
Community map	0.28	0.96	0.68***	0.31	0.95	0.63***
Learning corner	0.54	0.98	0.44***	0.51	0.96	0.45***
Classroom library	0.43	0.94	0.51***	0.47	0.94	0.47***
Student mailbox	0.31	0.96	0.66***	0.37	0.94	0.57***
Community corner	0.16	0.88	0.72***	0.22	0.87	0.65***
Organization chart	0.33	0.97	0.64***	0.38	0.95	0.57***
Learning guides (1/0)		0.43		0.10	0.94	0.84***

Note: Columns Non-VNEN and VNEN report means of the corresponding variables in each row for non-VNEN and VNEN schools. Columns Diff. show the difference in means between the two types. Statistical significance is calculated using standard errors clustered at the matched pair level (see text for more details).

contrast, lecture seating is still the most common arrangement in non-VNEN schools, with over 54% of them reporting having such arrangement most of the time, although this share also decreases slightly over time. Only over 7% VNEN schools report having lecture-style seating most of the time.

By design, a VNEN classroom also utilizes observable objects for teaching. These classrooms display various objects that help students relate their learning to real-life items as well as characteristics of their communities. These objects are often referred to as “artifacts”:

- *Community map*—showing each students’ home as well as local landmarks. The goal of this item is to integrate learning with the local community,
- *Classroom library*—including learning guides, reference material, storytelling books, and other reading materials,
- *Learning corner*—displaying students’ work,
- *Community corner*—displaying objects related to the life of the community,
- *Student mailbox*—where students can write notes to each other,
- *Student government chart*—displaying the structure of the student government and committees.

Besides utilizing the artifacts, the VNEN program also relies heavily on learning guides to substitute for textbooks. Learning guides have lessons and chapters similar to what is found in textbooks but also include a workbook with activities, exercises, and instructions on how to carry out these tasks.

These features are widely adopted by VNEN schools in practice. Over 80–90% of VNEN schools possess these items in their classroom. Surprisingly, a small share of non-VNEN schools also adopt these artifacts, even though they are not typical in a traditional school model. The adoption rate differs across artifacts, with 16–22% for community corners and 50.6–53.6% for learning corners.

2.2. Teaching Methods

The VNEN model also requires teachers to employ different teaching methods than the traditional model. We summarize the top 3 most used teaching practices as well as practices that teachers report using in more than half of their lessons in Table 2.

An important aspect of the VNEN pedagogy is its emphasis on relating lessons to real life and the use of games in teaching, allowing teachers to create the emotional spark for students and provide a well-rounded education for students (Parandekar et al., 2017). VNEN schools appear to adopt these strategies more than non-VNEN schools. Almost 93% of VNEN schools, as opposed to 70% of non-VNEN schools, report that they often encourage students to learn about themselves. Over 62% of VNEN schools report that they use game as part of their lesson as opposed to 42.4% of non-VNEN schools. Among VNEN schools, 41% report using role-play in Vietnamese lessons; this figure is only 38% for non-VNEN schools. Non-VNEN schools rely more often on lecturing, working with a few students on board, asking students to read out loud, and writing instruction on the board. Meanwhile, VNEN schools focus on encouraging students to learn about themselves and asking questions in class as well as finding new methods for group work and self-study.

Consistent with the frequent use of group seating, the VNEN pedagogy relies heavily on the use of group work. Students are organized into small groups and work together to complete assignments in class. Teachers work with individual groups or individual students most of the time, instead of lecturing to the whole class. This design can foster

Table 2
Teaching practices by VNEN status and year.

	2014		2015			
	Non-VNEN	VNEN	Diff.	Non-VNEN	VNEN	Diff.
Top 3 most used activities						
Group discussion	0.87	0.92	0.05***	0.88	0.95	0.07***
Visual materials	0.69	0.57	-0.12***	0.66	0.50	-0.16***
Assign in-class exercise	0.42	0.30	-0.12***	0.43	0.31	-0.12***
Encourage asking Qs	0.44	0.52	0.08***	0.56	0.63	0.07***
Self-study	0.20	0.33	0.13***	0.23	0.37	0.15***
Games, quizzes, and tests	0.14	0.17	0.03**	0.12	0.13	0.02*
Assign homework	0.09	0.04	-0.05***	0.08	0.05	-0.02***
Read from textbook	0.03	0.04	0.02**	0.03	0.04	0.01*
Write on boards	0.02	0.02	-0.00	0.03	0.02	-0.01***
Apply this practice more than half of the lessons (1/0)						
Lecturing	0.15	0.06	-0.09***	0.07	0.03	-0.04***
Work w/ 1–2 students in front of class	0.51	0.22	-0.29***	0.54	0.23	-0.31***
Ask students to read out loud	0.14	0.07	-0.07***	0.11	0.04	-0.07***
Write instruction on board	0.16	0.05	-0.11***	0.11	0.04	-0.08***
Encourage learning about themselves	0.70	0.93	0.23***	0.70	0.92	0.22***
Encourage asking Qs	0.81	0.91	0.10***	0.83	0.92	0.09***
Give feedback to class	0.84	0.80	-0.04*	0.86	0.85	-0.01
Individual feedback	0.79	0.79	-0.00	0.83	0.82	-0.01
Find new methods for group work	0.74	0.92	0.18***	0.72	0.92	0.20***
Find interesting way for self-study	0.66	0.82	0.16***	0.65	0.81	0.16***
Use games as part of lesson	0.42	0.62	0.20***	0.41	0.66	0.25***
Use role-play	0.38	0.41	0.03	0.36	0.42	0.06***

Note: Columns Non-VNEN and VNEN report means of the corresponding variables in each row for non-VNEN and VNEN schools. Columns Diff. show the difference in means between the two types. Statistical significance is calculated using standard errors clustered at the matched pair level (see text for more details). Under the Top 3 most used activities panel, each row shows the share of schools listing the corresponding activity as one of their top 3 most used activities.

participative and collaborative learning, enhancing “the probability for a student to be in his or her zone of proximal development,”⁸; and, thus, their ability to learn (Hogan and Tudge, 1999). The group work emphasis also encourages students to move around, connecting learning with movements.⁹ Over 92–94% VNEN schools report group discussion as one of their top three most used activities; 86–88% non-VNEN schools also report that group discussion is a top three most used teaching activities. These figures suggest that group discussion is the most popular

⁸ Zone of proximal development is the space between what learners can do without help and what they cannot do even with help from peers. It is the range that students can only learn with support from teachers and their peers (Hogan and Tudge, 1999).

⁹ E.g., Vazou et al. (2012).

pedagogical practice, although more VNEN schools tend to employ this as one of their main methods.

The design of VNEN also differs from the traditional model in terms of assessment approach. VNEN emphasizes formative instead of summative assessment, i.e., assessments are used to provide information on how students can improve their skills (Dang et al., 2022). This is done through quizzes, tests, and informal conversations between teachers and students and among teachers. Because of the focus on group work, students are encouraged to use group assessments (e.g., checking in pair, group leaders checking their members, or teacher checking the whole group).

Yet there is surprisingly little difference in terms of assessments in practice between the two types of school. Over 84–86% non-VNEN schools give feedback to class in more than half of the lessons, while for VNEN schools, this figure ranges between 80% to over 85%. Similarly, over 79–83% non-VNEN schools and almost 79–82.5% VNEN schools use individual feedback for more than half of the lessons.

In Table 3, we examine assessment types in the two models in detail. VNEN schools tend to use self-assessment, pair-checking, and group-checking more often, while non-VNEN schools use more traditional assessment methods such as the whole class checking the answer of one student or students writing answers on a slate to be checked by the teacher. Interestingly, both types of schools often assess non-cognitive skills, even though only the VNEN model has an emphasis on developing these skills.¹⁰

Two other important distinctions between the two school models are (1) the extensive use of student government within classroom to develop leadership and communication skills and (2) parental engagement. We lack data on how schools implement student government, so we do not discuss it here. In Table A1, we summarize the use of parental engagement for the VNEN and non-VNEN schools. Teachers in VNEN schools tend to engage more with students' parents, especially by communicating with them about students' behaviors and well-being as well as parents' concerns and wishes. VNEN schools also engage more with parents by inviting them to school and classroom activities.

2.3. Variation in implementation

Two surprising observations arise from examining the implementation of the Escuela Nueva in Vietnam. VNEN schools vary in the extent to which they adopt different components of the program. For example, most VNEN schools adopted the learning corner, but fewer schools adopted the community corner. Non-VNEN schools also appear to adopt some components of the VNEN model, although only the VNEN schools received the funding to implement these changes. Dang et al. (2022) document how the VNEN model was adopted in both VNEN and non-VNEN schools by constructing a composite index for VNEN implementation.

We also constructed this index and plot its distribution in Fig. 1.¹¹ On average, VNEN schools implement the program more than the non-VNEN schools, but many non-VNEN schools also implement some parts of the model. In 2015, we also note that the non-VNEN distribution spreads out to the right and the VNEN distribution also moves further to the right. In other words, both VNEN and non-VNEN schools implement even more of the VNEN model in 2015.

Why did non-VNEN schools implement some components of the VNEN model, especially when they did not have funding specifically for applying such a model? The reason, according to Parandekar et al. (2017), is that the Ministry of Education and Training explicitly allowed non-VNEN schools to adopt the model. The VNEN model was also understood by the schools to be the Ministry's pilot for a national educational reform that focuses on competency-based education, as opposed

to knowledge-based education (Duong et al., 2023). Anecdotal evidence suggests that non-VNEN schools tried to get ahead of the national reform by adopting some of the VNEN teaching practices.

3. Methods

We employed a mixed method longitudinal design to understand how the VNEN model shapes students' learning behaviors and whether these behavioral effects persist in the long run. The quantitative analysis uses a propensity score matching research design to quantify the treatment effect of the VNEN model on primary level students' behaviors in the short run using data from a World Bank impact evaluation study (Parandekar et al., 2017). We then examine the long-term effects on the same outcomes on a follow-up subsample as the students reached lower secondary level. We also conducted a qualitative study including interviews with school principals and interviews and video-recorded lessons with teachers from the lower secondary sample to understand how they perceived VNEN students' learning behaviors. This section briefly discusses the research design and methods we use for the analyses.¹²

It is important to note that we use the terms "VNEN students" to refer to students who studied in a primary school that elected to implement the VNEN model and "non-VNEN students" to refer to students who went to a non-VNEN primary school. For the lower secondary sample, these terms also refer to whether students were exposed to the VNEN model at the primary level.

3.1. Propensity score matching

Because schools voluntarily decide whether to adopt the VNEN model or not, simply comparing the average outcomes of VNEN and non-VNEN school would suffer from selection bias due to confounders that drive both adoption decision and students' learning behavior (Angrist and Pischke, 2009; Cunningham, 2021).

The World Bank conducted an initial impact evaluation on the impacts of the VNEN program on learning outcome using a propensity score matching research design (Parandekar et al., 2017). The authors first matched the VNEN schools with comparable non-VNEN schools based on several observable characteristics using a primary school census dataset,¹³ then collected data on the matched sample of schools to conduct the impact evaluation. These characteristics were selected, and the matching was done *prior* to the main data collection. This design has two advantages. First, it does not impose any functional form assumption, unlike standard OLS regression with additive controls. Second, the authors cannot cherry-pick control variables or specifications to achieve significant results.

Given these advantages, we use the same research design and data from Parandekar et al. (2017) to examine how the VNEN model impacts students' behaviors. We estimate the following OLS model on the matched sample:

$$Y_{i,s} = \alpha + \tau_{OLS} \cdot VNEN_s + \epsilon_{i,s}$$

where $Y_{i,s}$ denotes the behavioral outcome of student i in school s , $VNEN_s$ denotes the VNEN status of school s , and τ_{OLS} is an estimator for ATT of the VNEN model. Following Abadie and Spiess (2022), we cluster the standard error at the matched school pair.

The internal validity of this estimator depends on the selection process of the VNEN program. As noted previously, disadvantaged schools in each province were offered the program based on a number of characteristics, including the proportion of ethnic minority students and being located in remote and/or low-income areas. The propensity score

¹⁰ The only exception is with assessing leadership skills.

¹¹ The construction of these indices is discussed in Section 3.

¹² See Appendix A1 for a formal discussion of our research design and estimation approach.

¹³ The matching process is discussed in detail in Appendix A2.

Table 3
Assessment practices by VNEN status and year.

	2014			2015		
	Non- VNEN	VNEN	Diff.	Non- VNEN	VNEN	Diff.
Math						
Self-assessment	0.71	0.81	0.10***	0.78	0.85	0.07***
Checking in pair	0.78	0.90	0.12***	0.82	0.95	0.13***
Group leader checks	0.66	0.90	0.24***	0.69	0.92	0.23***
Other group members check	0.47	0.58	0.11***	0.50	0.68	0.18***
The whole class checks	0.74	0.34	-0.40***	0.75	0.37	-0.38***
Teacher checks answers on slates	0.43	0.13	-0.30***	0.46	0.17	-0.29***
Teacher checks notebooks during class	0.88	0.88	-0.00	0.91	0.91	-0.01
Teacher checks notebooks after class	0.65	0.71	0.06**	0.71	0.72	0.01
Teacher gives in-class quizzes	0.39	0.35	-0.04	0.36	0.34	-0.02
Teacher gives feedback to group	0.65	0.72	0.07**	0.67	0.71	0.04
Vietnamese						
Self-assessment	0.70	0.80	0.10***	0.76	0.84	0.07***
Checking in pair	0.76	0.89	0.13***	0.80	0.94	0.14***
Group leader checks	0.64	0.89	0.25***	0.69	0.91	0.22***
Other group members check	0.46	0.57	0.11***	0.48	0.66	0.18***
The whole class checks	0.73	0.34	-0.40***	0.73	0.37	-0.36***
Teacher checks answers on slates	0.37	0.12	-0.25***	0.39	0.16	-0.23***
Teacher checks notebooks during class	0.87	0.87	-0.00	0.89	0.90	0.00
Teacher checks notebooks after class	0.66	0.71	0.05*	0.68	0.70	0.01
Teacher gives in-class quizzes	0.38	0.37	-0.02	0.35	0.34	-0.01
Teacher gives feedback to group	0.64	0.72	0.09***	0.65	0.70	0.05*
Non-cognitive skills						
Communication	0.94	0.94	0	0.94	0.95	0.01**
Creativity	0.78	0.79	0.01	0.72	0.77	0.05***
Leadership	0.59	0.74	0.15***	0.46	0.66	0.19***
Critical thinking	0.62	0.64	0.02	0.57	0.61	0.04**
Problem solving	0.87	0.89	0.02	0.90	0.92	0.03***
Cooperation	0.89	0.95	0.06***	0.93	0.97	0.04***

Note: Columns Non-VNEN and VNEN report means of the corresponding variables in each row for non-VNEN and VNEN schools. Columns Diff. show the difference in means between the two types. Statistical significance is calculated using standard errors clustered at the matched pair level (see text for more details).

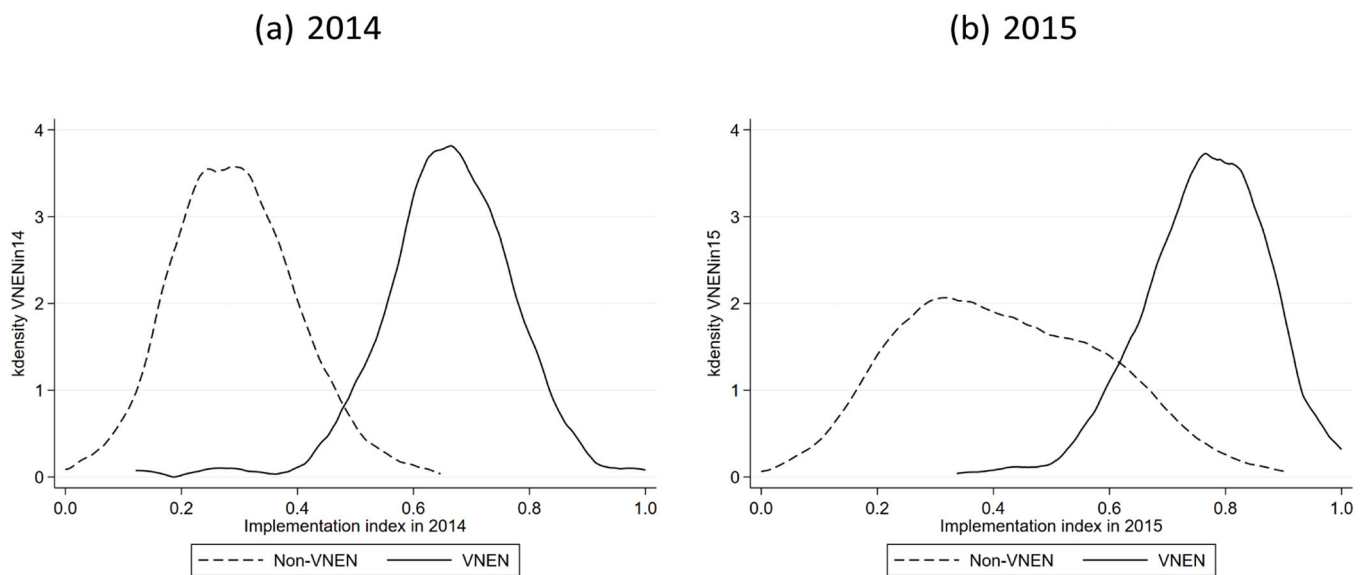


Fig. 1. : VNEN Implementation index. Note: Implementation index is constructed separately each year following Dang et al. (2022).

matching process thus accounts for these observable characteristics and other related factors at the school level.

3.2. Instrumental variable

As discussed in Section 2, some non-VNEN schools were implementing some components of the VNEN model. Therefore, the propensity score matching estimator may underestimate the true effect of the VNEN program because of this spillover. Dang et al. (2022) account for this problem by constructing a continuous index for VNEN

implementation across all VNEN and non-VNEN schools and estimating the effect of this index, denoted by $impl_{i,s}$, instead of the binary VNEN status:

$$Y_{i,s} = \alpha + \beta \cdot impl_{i,s} + \epsilon_s$$

Under the same assumptions about conditional independence and

common support above, along with an additional assumption about monotonicity,¹⁴ we can instrument the implementation index with the binary VNEN status. In other words, the first-stage equation is

$$impl_s = \eta + \theta.VNEN_s + u_i$$

where θ represents the effect of the VNEN model on the implementation index and β captures the average causal response of outcome Y to an increase in the implementation index.¹⁵ In other words, this IV model allows us to quantify the effect of going from not implementing the VNEN program at all to fully implementing it, accounting for the fact that non-VNEN schools also implemented some part of the program.

We briefly summarize the procedure to construct this index; for a more complete discussion of this method, see Dang et al. (2022). The VNEN implementation can be categorized into eight fundamental components: classroom space, group seating, classroom artifacts, learning guides, teaching methods, group work, assessment, student government, and parental engagement. Using the teacher data from the World Bank evaluation of the VNEN model, one can construct eight subindices for these fundamental components using principal component analysis (PCA). PCA can then be used again to combine these eight indices into one composite index for each year.

3.3. Qualitative interviews

Qualitative data were critical in this study to examine teachers' perspective of VNEN and their pedagogical practices that supported student learning. To understand teacher pedagogical practices, the research team¹⁶ conducted semi-structured interviews with the school principal and one math and one literature teacher at each school, as well as video-recorded three lesson periods.¹⁷ Interviews with principals helped understand if and how they supported changes to teachers' pedagogical practices, including those introduced by VNEN and then in the new competency-based curriculum. Interviews with teachers were conducted before we video-recorded their lessons to understand their perspective about pedagogical reforms, their general teaching style and pedagogical practices, and specifically what they intended to do in their lessons. We also conducted video-cued reflection interviews about their pedagogical practices. While we have conducted analyses of the video-recordings of classroom teaching and learning, for this paper, we focus primarily on interview data (DeJaeghere et al., 2023a, b). In addition to background information about whether teachers had VNEN training and the number of VNEN students in their classroom, we asked about how teachers and principals considered the benefit and/or challenges of the VNEN models. When teachers commented or compared VNEN and non-VNEN students' learning behaviors, we probed further to understand how they saw these differences. We also asked how they adapted their teaching styles when teaching VNEN or non-VNEN students and why they decided to teach certain ways.

Interviews were transcribed in Vietnamese and translated into English so that non-Vietnamese researchers could discuss and supervise coding and analysis. After listening to and reading the interviews,

¹⁴ Specifically, we assume that $impl_s(T=1) - impl_s(T=0) \geq 0, \forall s$ (monotonicity). In other words, the VNEN status does not increase the implementation for some schools, while decreasing the implementation for some other schools at the same time.

¹⁵ It is important to note that this parameter only applies to the range of $impl_s$, for which the implementation index gets shifted by the binary VNEN status (Angrist and Pischke, 2009).

¹⁶ The data were gathered by a group of Vietnamese researchers. All of them were trained in qualitative methods and qualitative data collection by two of the authors who also supervised the data obtaining process. These two authors also gathered data in one province in the final years.

¹⁷ For the scope of this paper, we only analyzed principals' and teachers' interviews.

researchers discussed key themes. Based on these themes and interviews questions, the codebook was developed. Then Vietnamese researchers, including one of the authors, coded the Vietnamese version of the interviews in Transana, including various themes related to implementation of VNEN and its pedagogical practices. While the qualitative data were gathered at the same time as the quantitative, we proceeded with separate analyses of the qualitative and quantitative data. We then discussed the findings for this analysis, asking what we could learn from the findings of the different data sets. After reviewing the quantitative findings about learning behaviors, we reanalyzed the qualitative data from VNEN and non-VNEN teachers and classrooms. We particularly drew on themes that relate to teachers' and principals' perceptions about VNEN and students' learning behaviors in the VNEN program and in the secondary classrooms.

4. Data

Our quantitative analyses use two data sources: (1) VNEN evaluation data collected by the World Bank for 2013–2016 at the primary school level (Parandekar et al., 2017); and (2) VNEN follow-up data collected by the Research on Improving Systems of Education (RISE) Project in Vietnam for 2017–2020 (Dang et al., 2022). The qualitative analysis is conducted on a sub-sample of the quantitative data. In this section, we summarize the sampling and sample size of each dataset.

The primary school sample was collected by the World Bank is comprised of 651 primary schools (323 VNEN schools and 328 non-VNEN schools).¹⁸ The original study collected data from students, teachers, and parents for the 2013–2014, 2014–2015, and 2015–2016 school years. Our quantitative analysis uses student questionnaire data from 2014 to 2015 (12,050 observations) and 2015–2016 (11,584 observations) because the relevant variables were only available in these years.

The lower secondary school sample collected by RISE Vietnam is comprised of 99 lower secondary schools. The RISE Project randomly sampled 100 primary schools from the 651 primary schools in the World Bank sample: 49 VNEN and 51 non-VNEN schools. The Project then followed the students from these primary schools to their lower secondary schools;¹⁹ since two of the 100 primary schools were connected to the same lower secondary school, the final sample of the RISE data only has 99 lower secondary schools. The Project sampled 2959 students from these 99 lower secondary schools, half of which were from the World Bank Study. Due to the budget constraints, the Project randomly drew half of the original sample to administer a questionnaire. This led to a final sample of 1199 lower secondary students with detailed information about learning behaviors. Data were collected for 2017–2018, when these students were in grade 7, and in 2019–2020, when they were in grade 9.

The qualitative study was conducted with a subsample of 20 secondary schools from the 99 in the quantitative study. These schools were from 10 different provinces across the Northern, Central, and Southern regions of the country. We selected schools in which the majority of their students came from VNEN primary schools, although not all students from the observed classrooms studied VNEN model in their primary level. In addition, a few of the lower secondary schools in our sample applied the VNEN model, while other schools had stopped the model before the research was conducted or had never applied the VNEN model. We selected grade 7 classrooms that included students who were given the student assessment and questionnaires in 2017–2018. Teachers from these classrooms were interviewed and video-recorded. We

¹⁸ This is an already-matched sample.

¹⁹ During fieldwork, the Project was able to track over 80% of the student sample.

followed these students through grade 8 and 9, interviewing their math and language teachers. Over the three years, 91 teachers²⁰ and 24 principals²¹ participated in the study. This analysis compared teacher data from secondary classrooms with students from VNEN primary schools as well as some who were not from VNEN primary schools. See Appendix A3 for a more detailed discussion of the qualitative data collection process.

4.1. Peer-learning and classroom participation outcomes

The quantitative analyses focus on peer-learning and participatory behaviors as the main outcomes. Peer-learning behaviors include how often students receive feedback from classmates, provide feedback to classmates, and communicate ideas among their group. Participation includes how often students raise questions in class and prepare for next-day lessons. Although both the World Bank and RISE data share similar outcome variables, these categorical variables use different categories to document the frequencies of each activity. The World Bank questionnaires use a subjective scale, e.g., (1) Always, (2) Usually, (3) Sometimes, and (4) Seldom/Never. The RISE questionnaires ask survey respondents the number of times a specific activity takes place, e.g., 3–5 times/day. To make our estimates easier to interpret and comparable across the two surveys, we dichotomize these variables so that they are zero if the frequencies are below the median, and one if they are above the median. The treatment effect thus reflects the changes in likelihood of engaging in an activity more than the median. We summarize these peer-learning variables by VNEN status and year in Table A3.

5. Results

5.1. Quantitative analysis: Short and Long-term effects on learning behaviors

Recall that we use two different empirical strategies: propensity score matching (PSM) and instrumental variable (IV), which measure two different causal parameters; PSM estimates the effect of the binary VNEN status and the IV estimates the effect of the continuous implementation index. For each identification strategy, we use two specifications. In the main specification, we only include the treatment variable on the right-hand side. Under the assumption that VNEN schools and non-VNEN schools are comparable across observable and unobservable characteristics, this specification would give us an unbiased estimate of the average treatment effect. As a check for this assumption, we include basic demographic and economic variables as control variables in the regression. If the two groups are indeed comparable, including such control variables would not alter our results substantially. Our controls include student's ethnicity, whether father and mother can speak and write Vietnamese, and parents' monthly income. To provide the correct inference, we cluster the data at the matched pair of schools (Abadie and Spiess, 2022). For the IV estimates, we also report the OLS and Pflueger effective F-statistics for weak instrument which is robust to heteroskedasticity (Olea and Pflueger, 2013).

We present the PSM estimations for the effects on learning behaviors in Table 4. Given the binary outcomes, our PSM model is equivalent to the linear probability model. We find that the VNEN increases the likelihood of all peer-learning activities considered. Students are more likely to help their classmates with in-class problems and check their answers; they are also more likely to receive the same treatment.

from their classmates. The effect sizes are meaningful. In 2014, about 36.4% non-VNEN students help classmates with in-class problems, so an

effect of 8.1 percentage points means that the VNEN increases the likelihood of this peer-learning activity by almost 22%. The effects on peer-learning activities in 2015 are statistically significant and much larger. For example, VNEN raises the probability of discussing lessons with classmates by 11.9 percentage points. Given that only 12.6% of the control group discusses lessons with classmates (Table A1), this effect is equivalent to an increase of almost 95%. Students are also 13.9 percentage points more likely to give comments to classmates (equivalent to a 39% increase) and 11 percentage points more likely to receive comments from their classmates (equivalent to a 30% increase).

In contrast, we find little effects on peer-learning activities in 2018 as students entered lower secondary schools. The effect sizes on all outcomes were small and statistically insignificant. In 2020, the effects are larger and statistically significant for discussing lessons with classmates and giving comments to classmates. It is somewhat surprising that we find null results in 2018 and significant results in 2020. One possible explanation is that students are keen to conform to teacher rules/norms in their earlier years of secondary school, but by their final years, they are willing to engage in behaviors that they previously developed through VNEN. Another potential explanation is that peer-learning activities become more important during the post-COVID period, as students just came back from their schools being closed for three months (early February to the end of April in 2020). This result, however, can also be a result of statistical noise. We use qualitative interviews with lower secondary teachers to verify the existence of such long-term effects (see more in Section 5.2).

Additionally, the results indicate that the VNEN program raises class participation, as VNEN students are more likely to raise questions in class by 3.9 percentage points (or 17.3%) in 2014 and by 3.9 percentage points (or 22%) in 2015. These participatory effects appear to fade out in lower secondary schools, as the estimates are no longer significant. The VNEN model does not affect students' preparation for lessons at home or their overall time studying at home.

The credibility of these results depends on whether the propensity score matching provides a comparable group of non-VNEN schools as the control group. If the adoption of the VNEN model is driven by socioeconomic factors that were not included in the propensity score estimation or unobserved factors such as teachers' ability or political economy among schools (e.g., political influence of principals), then our results might be biased. As a falsification check, we estimate the same regression on various socioeconomic factors (e.g., household wealth, parental education and monthly income) and detailed information about teachers and principals (e.g., teaching experience, teaching awards, wealth index, and self-reported political influence) that were collected in 2014 with the binary VNEN status as the only dependent variable, while clustering standard error at the matched school pair level. VNEN status should not have any effect on these variables if VNEN and non-VNEN schools are comparable.²² The results are reported in Table A4. As expected, the effects on most outcomes are statistically insignificant and close to zero. The only variable with a significant result is whether the teacher received a teaching award at the national or provincial level – this is expected given the false positive rate of 5%. The magnitude of the difference is also small: 6.1% of non-VNEN and 8.1% of VNEN teachers reported to have a teaching award.

We summarize the IV estimations in Table A5. Recall that the implementation index is normalized to range between zero and one, where zero means no implementation at all and one means full implementation of the model. The IV results are larger than the PSM results, suggesting that if a school goes from zero implementation to 100% implementation, the effects are larger than implied by the PSM results. For example, the PSM results indicate that adopting VNEN would increase the likelihood of helping classmates by 8%; the IV results indicate

²⁰ There were 41 math teachers and 44 literature/Vietnamese language teachers.

²¹ A few schools had a new principal during the study period.

²² This falsification test is in the same spirit as a balance test commonly used in the RCT literature (Bruhn and McKenzie, 2009).

Table 4
Effects of VNEN on peer-learning activities – PSM estimates.

	Primary		Lower secondary					
	2014		2015		2018		2020	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Peer-learning								
Help classmates with problems	0.080***	0.080***						
S.E.	(0.013)	(0.013)						
Check classmates' answers	0.086***	0.086***						
S.E.	(0.012)	(0.012)						
Classmates help with problems	0.082***	0.084***						
S.E.	(0.013)	(0.013)						
Classmates check my answers	0.101***	0.103***						
S.E.	(0.012)	(0.012)						
Discuss lesson with classmates			0.120***	0.117***	-0.011	-0.007	0.064*	0.071**
S.E.			(0.010)	(0.010)	(0.029)	(0.030)	(0.033)	(0.034)
Give comments to classmates			0.138***	0.135***	-0.013	0.005	0.067*	0.079**
S.E.			(0.012)	(0.012)	(0.040)	(0.040)	(0.035)	(0.036)
Classmates give comments			0.108***	0.108***	-0.015	-0.003	0.015	0.018
S.E.			(0.012)	(0.012)	(0.040)	(0.042)	(0.037)	(0.037)
Other learning practices								
Raise questions in class	0.039***	0.040***	0.038***	0.038***	0.035	0.037	0.001	0.008
S.E.	(0.013)	(0.013)	(0.011)	(0.011)	(0.038)	(0.039)	(0.036)	(0.036)
Prepare for next day lessons	0.011	0.013	0.004	0.005	-0.046	-0.015	-0.014	0.001
S.E.	(0.014)	(0.013)	(0.012)	(0.011)	(0.041)	(0.041)	(0.043)	(0.045)
Study at home	-0.009	-0.008	-0.017	-0.015	-0.034	-0.001	-0.025	0.029
S.E.	(0.010)	(0.010)	(0.010)	(0.010)	(0.050)	(0.050)	(0.054)	(0.049)
N	11,391	11,391	11,181	11,181	1199	1199	1086	1086
Additional Control	No	Yes	No	Yes	No	Yes	No	Yes

Note: Each cell in the table reports the result for the coefficient of a VNEN binary variable for the corresponding outcome variable in each row. The coefficient is estimated by a linear regression on a matched sample, and equivalent to the ATT estimated from a propensity score matching estimator. Two specifications are estimated. In columns (1), (3), (5), (7), outcome variables are regressed against the VNEN status variable. In columns (2), (4), (6), and (8) we control for the child's ethnicity, parents' ability to speak and write Vietnamese, and household wealth index.

that moving from not implementing to fully implementing the VNEN model would increase the likelihood of this activity by 21%. This is of course only true under the assumption that the instrument for the implementation index is valid and that schools would fully implement the VNEN model. The IV results echo a recent quantitative literature on the importance of implementation in practice in external validity and scaling of an educational intervention (Angrist and Meager, 2023; Angrist et al., 2023).

We also conduct heterogeneity analysis and report the results in Table A6. The short-term effects on peer-learning activities are stronger among boys, Kinh Majority students (relative to Ethnic Minorities students) and students from wealthier households. Students with parents who have completed secondary education or higher also experienced slightly higher impacts, but the difference is not statistically significant. In contrast, the long-term effects on peer-learning activities are mostly concentrated among students from wealthier households and with more educated parents.

5.2. Qualitative findings

5.2.1. Teachers' perception of student learning behaviors in secondary school

The qualitative data offer further evidence of teachers' assessment of students' communication, peer-learning, and engagement who were taught in VNEN classrooms, even when these pedagogical practices are not formally taught through VNEN curriculum and pedagogy at the secondary level. A majority of teachers we interviewed, regardless of whether they had VNEN training or experience teaching the model, agreed that the VNEN program equipped students with confidence in communicating and cooperating in groups, and a high level of engagement in their studies that is not typical in the more teacher-led classrooms. A teacher who had training on VNEN and whose school applied the model for only a short time shared that "I can see so many advantages of the VNEN model. First, it creates [a condition for the development of] students' activeness and creativity, meaning they are so

confident. That is a great point. Secondly, students' communication in that model is very good" (Math teacher, grade 8). "Active" and "confident in communication" were the two characteristics that most teachers and principals commented about students who had been taught in VNEN classrooms. They also acknowledged that these students felt freer to "express their opinions" as well as were able to "criticize each other" than students who had not learned in such classrooms. This way of learning engagement is not common in classrooms not implementing VNEN, where opportunities for quality peer feedback were limited.

Through discussions with teachers and principals, it can be inferred that they credited VNEN's focus on group work and self-construction of knowledge as supporting students to develop these characteristics. For example, a math teacher teaching grade 8 shared that "[VNEN] students must self-explore knowledge in a group and then present the content they like and understand". A grade 7 literature teacher agreed that this VNEN teaching style allowed students to "take initiative in preparing lesson and in group discussion", which helped students develop self-study skills. Sharing this point, a grade 9 literature teacher also stressed the importance of group work and ability to communicate well among students who were taught in VNEN classrooms. She said: "[these] students take initiative in working in groups, follow questions in the textbook, share opinions with each other, give feedback and comment more". While group work as a pedagogical practice was a major change in the 2006 curriculum reform in Vietnam, many classrooms continued to use teacher-centered lecturing and question/answer (Math teacher, grade 9). Group work in the VNEN program seems to be effective in developing students' cooperation and communication and other skills. The following quote is from a principal whose school had used the VNEN program but then discontinued it (for lack of resources, as it was no longer supported formally in secondary schools).

We had a classroom that was used only for this VNEN class. This classroom had a projector, shelves, cabinets, chairs, desks which can fully serve the students. In the classroom, we decorated exactly the same as what we were trained from MOET and as we observed from classes in other schools. I think teachers understood the VNEN model well. There

were some matters that they focused greatly on when they taught. For example, before we had applied this model, students' communication ability was very bad. They were natural and disordered. Since this class, they could be self-managed and were able to manage their study. Of course, when they could self-manage, the class would be more disciplined. Especially, their ability in communication improved greatly, so did their ability in cooperation. (Principal, school in the Central region).

5.2.2. Teachers' concerns of learning behaviors in a test-oriented environment

While this principal saw students' ability as disciplined, some teachers were concerned that students coming from VNEN classrooms were undisciplined. They felt they had bad writing skills and performed less well on tests. In addition, these students were comfortable working in groups, sharing ideas and interacting during class time at the primary level, and several teachers found that these students were not disciplined. Talking is not a norm that many teachers promoted in classrooms because they needed students to stay quiet when they lectured, which we found in our video-recorded lessons to take a majority of the class time (particularly in classes where group activities were not present). Hence, when comparing a class that had VNEN students and a class that had non-VNEN students, a Grade 7 literature teacher, trained in the VNEN program, considered the latter "more well-behaved" (ngoa hơn) and they did "less private talk" (ít nói chuyện riêng hơn). This teacher was teaching both a textbook-based class with students who did not go to a VNEN school and another class that applied VNEN pedagogies and textbooks. She further commented about group work activities in the VNEN classroom: "Actually, in VNEN model like in [name of the class], it creates a condition for students to do more [private] talk, because they are used to discussion and talking. [Students] take advantage of group discussion to do private talks. Only [name of the non-VNEN class] has a sense of learning". (Literature teacher, grade 7).

Concerns were also raised about the focus on group work and communication in the VNEN program not preparing students for the kinds of learning they needed in secondary classrooms, and for test taking. A few teachers mentioned that the focus on group work at the primary level neglected the importance of teaching students to write and take notes. As a result, teachers felt that these students struggled with notetaking and could not properly record knowledge in their notebooks during lessons so that they could review later. Teachers had to spend additional time on teaching students these skills in the secondary grades. Some principals were concerned with teachers spending their time in these ways: "In the VNEN program, they only sit in group, talk, and do group work; a group leader will present results of group work. That's why they don't have writing and note-taking skills. Consequently, secondary teachers lost much of their time in class." (Principal, school in the north). This principal felt that classroom time should have been used to review lessons and key knowledge. Instead, teachers in his school, especially grade 6 teachers, had to use their class time to re-teach students how to write.

Another area of concern in using VNEN pedagogy is that it created learning gaps between students. Some principals and teachers pointed out that group work is not an effective way for students with lower learning abilities to improve. For instance:

In group activities, teachers have to group good, average and weak students together so that they can support each other. However, the weak students usually let the good students do everything [...]. In VNEN, they [students] were only given a problem to solve without any sample solutions. Students have to find out the solution based on teachers' direction. Hence, only students with solid knowledge can solve the problem. Weak students will lag behind. (Principal, school in Central region)

The idea that students self-construct knowledge in groups and teachers serve as guides and facilitators in that process became a concern to teachers as they felt that students could not acquire the

knowledge without teachers helping them to do so. Furthermore, some students had more prior knowledge while others had less, and teachers believed that these students could not contribute much to groups and had to depend on their peers to finish group tasks. A math teacher in grade 8 shared the following:

I think one limitation of VNEN is that students have heterogenous levels of academic achievement. Weak students could be weaker because they cannot follow [their group members]. They can't communicate with their friends. Additionally, in the actual VNEN model, lessons should be in modules, not the same lesson for the whole class, meaning groups that finish first will continue to the next activity. Teaching like that is difficult for weak students unless teachers are very actively flexible. If we conclude knowledge [too soon], weak students cannot follow because good students already finish the 5th task, for example, but weak students are still at the first or second task. It's very difficult.

In sum, the qualitative findings show that secondary teachers regard students from VNEN classrooms as more communicative with their peers and more likely to participate in class, which is consistent with the long-term quantitative results. However, these learning behaviors may not persist as strongly, as also shown in the quantitative findings, if teachers do not value and actively teach with these practices. Principals and teachers in secondary schools felt that a focus on group work and collaborative learning undermines some aspects of learning, such as writing and taking notes. They also felt that a focus on group work and related competencies of communication and cooperation do not support students with lower learning outcomes because they need to learn content to achieve on tests.

6. Conclusion

Escuela Nueva is a long-standing pedagogical model that is widely adopted across different countries. Although the implementations may vary in different contexts, the model aims to foster non-cognitive skills alongside cognitive learning outcomes for students. Yet, whether and how the model effectively alters students' learning behaviors remains unclear.

We assess evidence for short and long-term effects on students' learning behaviors by studying the VNEN program, an adaptation of the Escuela Nueva model in Vietnam. We found substantial positive impacts on peer-learning activities and students' asking questions in class in the short run. In the long run, some of these learning behaviors persisted in 2020, 8 years after the students first studied under the VNEN model. Qualitative data from lower secondary school teachers and principals also show that they see these learning behaviors among students who have learned in VNEN classrooms, though these behaviors are not equally valued at the secondary level.

The usual caveats of observational studies using propensity score matching apply to our study. We cannot account for any selection on unobservable characteristics; if any unobservable characteristics of the school, e.g., political ties, affects both peer-learning and selection into the VNEN, then our results are no longer unbiased. Another caveat is that the follow-up sample is relatively small, making the long-term estimates noisier than the short-term results. Nonetheless, the standard errors are small, suggesting that we are observing true effects. The long-term effects are also supported by qualitative data from lower secondary school teachers' and principals' perspectives about students from VNEN classrooms.

These results underscore the importance to look beyond the usual cognitive and non-cognitive outcomes when evaluating a pedagogical reform. Since these reforms seek to change teachers' pedagogical practices, it is crucial to consider how such reforms affect students' learning behaviors, especially in the long run. In Vietnam, the VNEN program focuses strongly on participatory and collaborative learning that fosters communication, cooperation, and self-managed learning. These

outcomes are equally important as test scores and non-cognitive skills, and they are now included in the new competency-based curriculum as valued outcomes to be taught, learned and assessed. However, using peer learning to achieve these outcomes may come with important downsides that require further attention from researchers. Peer learning can be regarded by teachers as disruptive to learning content for tests. In addition, teachers felt that poorer performing students do not benefit particularly from group work as they may not be learning content sufficiently or not able to communicate effectively with their peers. The learning outcomes of students and how they are affected by these different pedagogies (e.g., group work, peer feedback or teacher feedback) is a question for further empirical research.

Author agreement statement

We the undersigned declare that this manuscript is original, has not been published before and is not currently being considered for publication elsewhere.

We confirm that the manuscript has been read and approved by all named authors and that there are no other persons who satisfied the criteria for authorship but are not listed.

We further confirm that the order of authors listed in the manuscript has been approved by all of us.

We understand that the Corresponding Author is the sole contact for the Editorial process. He/she is responsible for communicating with the other authors about progress, submissions of revisions and final approval of proofs.

CRedit authorship contribution statement

Dao Vu Phuc Phuong: Conceptualization, Formal analysis, Methodology, Writing – original draft, Writing – review & editing. **DeJaeghere Joan:** Conceptualization, Data curation, Formal analysis, Methodology, Writing – original draft, Writing – review & editing. **Vu Khoa:** Conceptualization, Formal analysis, Investigation, Methodology, Visualization, Writing – original draft, Writing – review & editing. **Glewwe Paul:** Conceptualization, Data curation, Formal analysis, Methodology, Writing – original draft, Writing – review & editing.

Acknowledgements

We thank the Research on Improving Systems of Education (RISE) for funding the data collection of this study. We are also grateful for the valuable comments from the discussant and audience at the Asian Development Bank (ADB) conference. Khoa Vu is thankful for the research grant and feedback from the ADB for earlier versions of this draft. He is also grateful for the thoughtful comments from Dr. Hiep Pham, attendants of the 2022 ADB-ISEAS Conference, CGDev seminar, and the APEC UMN seminar.

Appendix A. Supporting information

Supplementary data associated with this article can be found in the online version at [doi:10.1016/j.ijedudev.2024.103017](https://doi.org/10.1016/j.ijedudev.2024.103017).

References

Abadie, A., Spiess, J., 2022. Robust post-matching inference. *J. Am. Stat. Assoc.* 117 (538), 983–995.

- Angrist, J.D. and Pischke, J.-S. (2009). *Mostly harmless econometrics: An empiricist's companion*. Princeton university press.
- Angrist, N., & Meager, R. (2023). *Implementation Matters: Generalizing Treatment Effects in Education*. Available at SSRN 4487496.
- Angrist, N., Aurino, E., Patrinos, H.A., Psacharopoulos, G., Vegas, E., Nordjo, R., Wong, B., 2023. Improving learning in low-and lower-middle-income countries. *J. Benefit-Cost. Anal.* 1–26.
- Booij, Adam S., Leuven, Edwin, Oosterbeek, Hessel, 2017. Ability peer effects in university: Evidence from a randomized experiment. *Rev. Econ. Stud.* 84. 2 547–578.
- Bremner, N., Sakata, N., Cameron, L., 2022. The outcomes of learner-centred pedagogy: a systematic review. *Int. J. Educ. Dev.* 94, 102649.
- Bruhn, M., McKenzie, D., 2009. In pursuit of balance: randomization in practice in development field experiments. *Am. Econ. J.: Appl. Econ.* 1 (4), 200–232.
- Carrell, Scott E., Bruce, I.Sacerdote, James, E., 2013. West. From natural variation to optimal policy? The importance of endogenous peer group formation. *Econometrica* 81 (3), 855–882.
- Cunningham, S. (2021). *Causal inference*. Yale University Press.
- Dang, H.-A., Glewwe, P., Lee, J., and Vu, K. (2022). The impact evaluation of Vietnam's Escuela Nueva (new school) program on students' cognitive and non-cognitive skills. *DeJaeghere, Joan, Duong, Bich-Hang, Dao, Vu, 2023a. Quality of teaching and learning: the role of metacognitive teaching strategies in higher-performing classrooms in Vietnam. Educ. Res. Policy Pract.* 1–20.
- DeJaeghere, Joan, Dao, Vu, Duong, Bich-Hang, Luong, Phuong, 2023b. Learning inequities in Vietnam: Teachers' beliefs about and classroom practices for ethnic minorities. *Comp.: A J. Comp. Int. Educ.* 53 (3), 399–416.
- Duong, Bich-Hang, Dao, Vu, DeJaeghere, Joan, 2023. Complexities in teaching competencies: a longitudinal analysis of Vietnamese teachers' sensemaking and practices. *Pedagog., Cult. Soc.* 1–23.
- Eisenkopf, Gerald, 2010. Peer effects, motivation, and learning. *Econ. Educ. Rev.* 29 (3), 364–374.
- Evans, David K., Mendez Acosta, Amina, 2021. Education in Africa: what are we learning? *J. Afr. Econ.* 30 (1), 13–54.
- Feld, Jan, Zölitz, Ulf, 2017. Understanding peer effects: on the nature, estimation, and channels of peer effects. *J. Labor Econ.* 35 (2), 387–428.
- Forero-Pineda, C., Escobar-Rodriguez, D., and Molina, D. (2006). Escuela Nueva's impact on the peaceful social interaction of children in Colombia. In *Education for All and Multigrade Teaching*, pages 265–300. Springer.
- Glewwe, Paul, and Karthik Muralidharan. "Improving education outcomes in developing countries: Evidence, knowledge gaps, and policy implications." *Handbook of the Economics of Education*. Vol. 5. Elsevier, 2016. 653-743.
- Hammiller, K. (2017). *The Colombian Escuela Nueva school model: Linking program implementation and learning outcomes*. PhD thesis, Tulane University, Payson Center for International Development.
- Hogan, D.M., & Tudge, J.R. (1999). Implications of Vygotsky's theory for peer learning.
- Kline, R., 2002. A model for improving rural schools: Escuela Nueva in Colombia and Guatemala. *Current Issues in. Comp. Educ.* 2 (2), 170–181.
- Kremer, Michael, Brannen, Conner, Glennerster, Rachel, 2013. The challenge of education and learning in the developing world. *Science* 340 (6130), 297–300.
- Le, H.M., 2018b. The reproduction of 'best practice': Following Escuela Nueva to the Philippines and Vietnam. *Int. J. Educ. Dev.* 62, 9–16.
- Le, H.M., 2018a. Another textbook project? The implementation of Escuela Nueva in Vietnam. *Educ. Res. Policy Pract.* 17 (3), 223–239.
- McEwan, P.J., 1998. The effectiveness of multigrade schools in Colombia. *Int. J. Educ. Dev.* 18 (6), 435–452.
- Olea, J.L.M., Pflueger, C., 2013. A robust test for weak instruments. *J. Bus. Econ. Stat.* 31 (3), 358–369.
- Paloyo, Alfredo R. "Peer effects in education: recent empirical evidence." *The economics of education*. Academic Press, 2020. 291-305.
- Parandekar, S.D., Yamauchi, F., Ragatz, A.B., Sedmik, E.K., and Sawamoto, A. (2017). *Enhancing school quality in Vietnam through participative and collaborative learning*.
- Psacharopoulos, G., Rojas, C., Velez, E., 1993. Achievement evaluation of Colombia's "Escuela Nueva": Is multigrade the answer? *Comp. Educ. Rev.* 37 (3), 263–276.
- Schweisfurth, M. (2013). *Learner-centered education in international perspective: Whose pedagogy for whose development?* Routledge.
- Schweisfurth, M., 2011. Learner-centered education in developing country contexts: From solution to problem? *International journal of educational development* 31 (5), 425–432.
- Sojourner, Aaron, 2013. Identification of peer effects with missing peer data: Evidence from Project STAR. *Econ. J.* 123 (569), 574–605.
- Tabulawa, Richard. *Teaching and learning in context: Why pedagogical reforms fail in Sub-Saharan Africa*. African Books Collective, 2013.
- Vazou, S., Gavrilou, P., Mamalaki, E., Papanastasiou, A., Sioumala, N., 2012. Does integrating physical activity in the elementary school classroom influence academic motivation? *Int. J. Sport Exerc. Psychol.* 10 (4), 251–263.