Does a National System of "Small-Group" Learning Improve Outcomes?
An Analysis of the "New Horizons" Education Reform

Hagit Glickman
Nurit Lipshtat
Tal Raz
David Ratner

National Authority for Measurement and Evaluation in Education (RAMA)
Ministry of Education, Israel

Overview

In 2007, the Israeli Ministry of Education launched a national reform program in primary schools known as "New Horizons". The "New Horizons" reform, which involving a significant salary increase for teachers, is intended to raise student achievement levels, improve school climate, and provide equal opportunities for all students. The design of the program centers on the formation of "small-group" learning formats in which teachers work with groups of 5-8 students on a daily basis. Complementing the "small-group" format, schools are provided with additional resources for professional development that are intended to prepare teachers to work effectively within the small group format.

Purpose

To determine the efficacy of the reform, and to provide information that would inform future adaptations, the National Authority for Measurement and Evaluation in Education (RAMA), undertook a broad evaluation of the reform. The current study focuses on the first two years of implementing the reform in more than 700 schools, comprising about half of all public primary schools in Israel. The study was intended to address the following questions:

1. What effect did the reform have on students' achievements in Reading and Writing in mother tongue, Mathematics, Sciences and English as a foreign language?
2. What effect did the reform have on school climate?
3. What were teachers', principals', and inspectors' perceptions of the reform's effect on their professional status, self efficacy, school quality, professional capability, and work load?

In addressing these questions, the current paper aims to provide an interim report on the reform's effects, but also to offer insights into the way in which national education policy can serve as a viable tool for improving teaching and learning, school climate, and professional development.

Providing definite conclusions regarding the reform effectiveness is challenging as the "New Horizons" reform (similar to many other educational interventions) was not designed and implemented, for various reasons, as a controlled experiment. There was no "treatment group" vs. "control group" as the participation in the reform was in most cases a decision of the school administration and was not based on random sampling. Statistical methods were chosen in order to overcome this obstacle; however, an appropriate caution must be taken when drawing conclusions.

Data/ Method

Questionnaires for teachers, principals, inspectors and parents were developed and administered for the purpose of the study. The questionnaires dealt with respondent's attitudes toward the reform, with its implementation and with their future expectations. Data were collected from a representative sample of (around) 4,700 teachers, 900 principals, 100 inspectors, and 1,000 parents, from schools that joined the reform as well as from schools that didn't. In order to get a better understanding of the implementation processes, the questionnaires were supplemented by in-depth interviews with principals, teachers and students from 70 schools.

The analysis of the reform's effects was based on data from standardized national assessments (the "Meitzav") that are administered to 5th grade students in the subjects of Reading and Writing in mother tongue (Hebrew or Arabic), Mathematics, Science and English as a foreign language; and also on data from a national school-climate survey that is administered to 5th and 6th grade students. The data were collected from a representative sample of 740 primary schools, were about 45% of them have already undergone 1-2 years of reform and the rest 55% have not yet joined the project.

The data were analyzed using hierarchical linear modeling (HLM) procedures, in a quasi-experimental approach. These statistical procedures can partly compensate the fact that the reform was not implemented in a controlled experiment form. More specifically, they allow monitoring background variables, such as school's achievements in the past and social-economic status that may affect student achievements (i.e., confounding variables). The HLM procedure also allows analyzing data at different levels, i.e., student-level data and school-level data, with the appropriate reference to measurement errors and sampling errors (e.g., Raudenbush and Bryk, 2002).

The analysis of students' achievements is comprised of three models, with each being an expansion of the previous one. The first model is an unconditional means model (one-way random effect ANOVA) that aims to estimate the within schools and between schools variation, and to provide a baseline against which one can compare more complex models. The second model is a conditional model which includes school level and student level background variables. In this model, a separate regression line is fit to the data for each school to describe the relationship between student's score and his/her social-economic status (SES). The intercept of the regression line represents the school mean score and is estimated after taking account of school's background characteristics. The slope of the regression line indicates the extent of inequality attributable to SES. Steeper lines indicate a greater impact of SES on student performance – that is, more inequality - while more gradual lines indicate a lower impact of SES – that is less inequality, as illustrated in the diagram below.
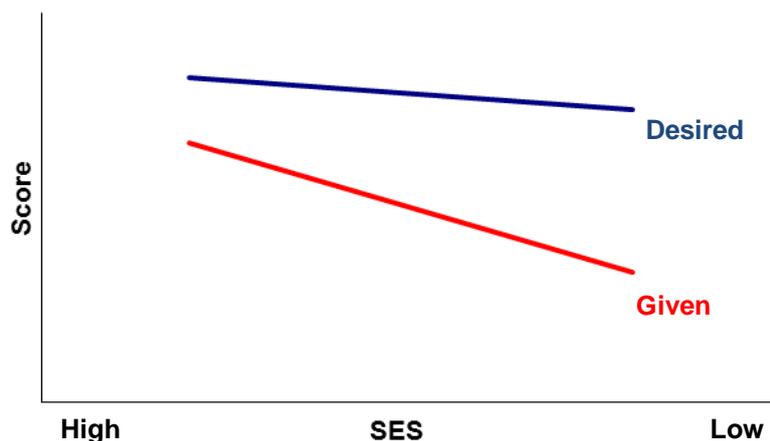
The third model examines the effect of the "New Horizons" reform on students' academic achievements. The average regression line for schools which joined the reform is compared to that of schools which did not, where higher intercept would indicate that the reform helps to raise achievements, and more moderate slope would indicate that the reform helps to reduce gaps between

students from different social-economic backgrounds. (For a mathematical description of the models see Appendix.)

The effect of the reform on school climate is also studied using hierarchical linear modeling. The analysis focuses on five indicators of school climate: students' behavior in class, involvement in violent incidents, personal relations between students and teachers, teachers' expectations for success, and general attitude towards school. The statistical models used to assess the reform's effect on school climate are similar to the ones used for academic performance; except for the use of individual SES. Here, the question of gaps between students from different social-economic backgrounds is irrelevant.

Results

Initial results indicate depict wide-ranging support for continued implementation of the reform: 99% of inspectors and principals, 88% of teachers, and 85% of parents report that they wish it to continue. In addition, 96% of inspectors, 93% of principals, and 71% of teachers (in participating schools) indicate that the reform contributed to an improvement in school's work. Teachers, principals, and inspectors perceive improvement in the status of the teacher, and teachers report a rise in professional efficacy, and ability to cope with student needs.

The "small-group" learning format is widely perceived as an outstanding feature that strengthens the relations between teachers and students, and allows the tailoring of instruction to student needs.

Based on principals' reports, 52% of the students in schools (on average) participate in the "small-group" learning formats. Teachers report that 41% of the additional hours are devoted to students with learning difficulties, 32% to students with average achievements, and about 10% to outstanding students. The rest 17% of the hours are devoted to individuals with various characteristics such as social difficulties and behavioral problems.

Teachers also report that 56% of the hours are allocated to reinforcement of core subjects (especially Reading and Writing in mother tongue and Mathematics), 28% are dedicated to expanding and enriching knowledge of specific subjects; and 16% are devoted to individual personal dialogue with students.

The effect of the reform on student achievement is less encouraging so far. Despite some pockets of improvement, no real difference in student achievement on state assessments was found between schools that joined the reform and schools that didn't. The most notable exception is in Mathematics for 5th grade, where among Hebrew-speaking secular schools, reform schools scored significantly higher than non-reform schools. That is, after controlling schools past achievements and social-economic status, a difference of about 10 points on the Meitzav perennial scale (0.10 SD) was found in favor of the reform schools (see Table 1). In the area of reducing the gaps between students from different social-economic backgrounds, no significant difference was found between the two groups. In addition, no significant differences in students' performance were found among Hebrew speaking religious schools and Arabic speaking schools

Another notable effect was found in the area of school climate, where among Hebrew speaking secular schools, reforms schools have shown better results on the five indicators than non-reforms schools. More specifically, after controlling past levels of school climate and social-economic status, significant differences were found in the following indicators in favor of the reform schools: 4.2 points (0.12 SD) in students' behavior in class, 0.9 points (0.05 SD) in students involvement in violence events, 3.4 points (0.10 SD) in personal relations between students and teachers, 1.6 points (0.06 SD) in teachers' expectations for success, and 2.0 points (0.06 SD) in students general attitude toward school (see Table 2). No significant differences were found among Hebrew speaking religious schools and Arabic speaking schools.

Discussion

Our research indicates several signs that the "New Horizons" reform has made a positive contribution to the schools. The "small-group" learning format is perceived as a helpful tool for the advancement of students' academic achievements and social-emotional state. The teachers report an improvement in their sense of capability as well as in their satisfaction from their salary and professional status. The vast majority of those participating in the reform - inspectors, principals, teachers and parents - would like to see the program continue. Moreover, students in schools that joined the reform report better learning atmosphere in class, less violence and closer student-teacher relations.

Despite the above encouraging findings, we found no real improvement in students' achievements (except for mathematics). This result is not so surprising, as the study was carried out only two years after the reform was first implemented. Educational studies which follow educational reforms do not usually find significant changes in students' achievements after such a short amount of time. Nevertheless, it seems that the "New Horizons" reform has the potential to advance the Israeli education system towards better performances.

The reform includes some important actions to improve teaching (and hence improve learning). Besides providing opportunities for more adaptive instruction (the "small-group" learning format), it also aims to enhance the quality of teachers. In particular, the reform involves a significant salary increase for young teachers, more effective processes for selecting the right candidates to become teachers, new tools for teachers' evaluation, and more rigorous professional development program. However, such changes do not occur overnight. Changing the quality of teachers is a long and complex process and it may take several years until it will yield the desired results.

References

Boissiere, M. (2004) *Determinants of Primary Education Outcomes in Developing Countries*, **Background Paper for the Evaluation of the World Bank's Support to Primary Education**.

Cotton, K. (1989) *"Educational Time Factors",* **School Improvement Research Series**, **close – up # 8** (http://www.nwrel.org/archive/sirs/4/cu8.html)

Barber, M. and Mourshed, M. (2007) *"how the world best performing school systems come out on top"* **McKinsey&Company**

Raudenbush, S.W. and Bryk, A.S. (1995). "*The Estimation of School Effects*", **Journal of Educational and Behavioral Statistics (20): 307-335.**

Raudenbush, S.W. and Bryk, A.S. (2002). **Hierarchical Linear Models: Applications and Data Analysis Methods**, Sage Publications.

Rosenbaum, P. and Rubin, D.B. (1983). "*The Central Role of the Propensity Score in Observational Studies for Causal Effects*", **Biometrica (70): 41-55.**

Singer D.S. (1998). *"Using SAS PROC MIXED to Fit Multilevel Models, Hierarchical Models' and Individual Growth Models",* **Journal of Educational and Behavioral Statistics (42): 323-355.**

Wills J.D. (2006). "*Learning Divides: Ten Policy Questions About the Performance and Equity of Schools and Schooling Systems*", **UNESCO Institute of Statistic Working Paper No. 5.**

Woessman, L. (2007) *Fundamental Determinants of School Efficiency and Equity: German States as a Microcosm for OECD Countries*, **CESIFO Working Paper no. 1981**.

Table 1: Hierarchal Linear Models for 5th Grade Mathematics Achievements - Hebrew Speaking Secular Schools

	Model 0		Model 1		Model 2	
	Estimate	(SE)	Estimate	(SE)	Estimate	(SE)
Effects						
Intercept	-0.033	(0.030)	-0.014	(0.025)	-0.059	(0.033)
School Past Results			**0.169**	(0.026)	**0.163**	(0.026)
School mean SES			**-0.076**	(0.029)	**-0.094**	(0.030)
School SD of SES			**-0.103**	(0.028)	**-0.102**	(0.028)
SES (student level)			**-0.138**	(0.006)	**-0.138**	(0.007)
"New Horizons"					**0.111**	(0.053)
"New Horizons" × SES					0.001	(0.011)
Variance components						
Schools	**0.141**	(0.018)	**0.091**	(0.012)	**0.089**	(0.012)
Students	**0.884**	(0.013)	**0.828**	(0.012)	**0.828**	(0.012)
Proportion of variance between schools	13.8%		9.9%		9.7%	
Proportion of between schools variance explained			35.4%		36.9%	
Proportion of within schools variance explained			6.3%		6.3%	

Note: The statistical significance of the effects marked in bold is in the order of $p<0.05$.

Table 2: Hierarchal Linear Models for Climate Indices, 5th – 6th Grade Students - Hebrew Speaking Secular Schools

Model 2	Students' Behavior in Class	Involvement in Violent Incidents	Teacher Student Relationships	Teachers' Expectations for Success	General Attitude Towards School
	Estimate	Estimate	Estimate	Estimate	Estimate
	(SE)	(SE)	(SE)	(SE)	(SE)
Effects					
Intercept	**31.206**	**15.643**	**56.596**	**85.027**	**71.785**
	(0.775)	(0.233)	(0.564)	(0.292)	(0.554)
School's past Index	**4.912**	**1.097**	**4.410**	**2.373**	**3.844**
	(0.580)	(0.191)	(0.441)	(0.251)	(0.423)
School mean SES	**1.588**	0.129	**1.506**	**1.210**	0.925
	(0.690)	(0.219)	(0.514)	(0.281)	(0.499)
School SD of SES	**-1.788**	**0.611**	-0.764	-0.259	**-0.940**
	(0.680)	(0.206)	(0.497)	(0.259)	(0.487)
"New Horizons"	**4.166**	**-0.875**	**3.362**	**1.574**	**2.048**
	(1.208)	(0.364)	(0.880)	(0.457)	(0.865)
Variance components					
Schools	**115.800**	**7.998**	**55.275**	**12.042**	**51.908**
	(9.369)	(0.870)	(5.013)	(1.322)	(4.865)
Students	**1014.140**	**357.060**	**1166.930**	**617.170**	**1272.420**
	(6.769)	(2.387)	(7.790)	(4.120)	(8.495)
Proportion of variance between schools	10.2%	2.2%	4.5%	1.9%	3.9%
Proportion of between schools variance explained	21.4%	20.9%	34.1%	45.7%	24.0%
Proportion of within schools variance explained	2.6%	1.1%	2.9%	2.5%	1.1%
Proportion of variance between schools - Model 0	12.7%	2.8%	6.7%	3.5%	5.1%

Note: The statistical significance of the effects marked in bold is in the order of $p<0.05$.

Appendix: Mathematical description of the models

The analysis of students' achievements is comprised of three models, as follows:

Model 0 (unconditional):

$$Y_{ij} = \beta_{0j} + \varepsilon_{ij}$$

$$\beta_{0j} = \gamma_{00} + u_{0j}$$

At level-1, the student level outcome (Y_{ij}) is expressed as the sum of an intercept for the student's school (β_{0j}) and a random error (ε_{ij}) associated with student i at school j. At level-2, the school level intercept is expressed as the sum of an overall mean (γ_{00}) and a series of random deviations from that mean (u_{0j}). Model 0 is simply a one-way analysis of variance model.

Model 1:

$$Y_{ij} = \beta_{0j} + \beta_{1j}X_{ij} + \varepsilon_{ij}$$

$$\beta_{0j} = \gamma_{00} + \sum_{k \geq 1} \gamma_{0k}W_{jk} + u_{0j}$$

$$\beta_{1j} = \gamma_{10}$$

This model expands model 0 in two ways. First, it includes a single level-1 predictor, SES (X_{ij}), stipulating that a student's achievement score is related to his or her SES. Second, at level-2, the model includes several school level predictors (W_{jk}), such as school's past results and the average SES of students within the school.

Model 2:

$$Y_{ij} = \beta_{0j} + \beta_{1j}X_{ij} + \varepsilon_{ij}$$

$$\beta_{0j} = \gamma_{00} + \sum_{k \geq 1} \gamma_{0k}W_{jk} + \gamma_{0k+1}R_j + u_{0j}$$

$$\beta_{1j} = \gamma_{10} + \gamma_{11}R_j$$

This model examines whether a difference exists between reform schools and non-reform schools. Variable R_j indicates whether a school joined the reform or not. Two more fixed effects are included: at level-2 (the school level) γ_{0k+1} is the effect of having joined the reform on school level intercept, and at level-1 (the student level) γ_{11} is the effect of having joined the reform on the slope of the linear relationship between student's achievement and SES.

The effect of the reform on school climate is also analyzed using hierarchical linear models, as described below. These models are similar to the ones used to analyzed achievement data except for the fact that they do not include SES as a student level predictor.

Model 0: an unconditional model as defined for achievements data.

Model 1 (school climate):

$$Y_{ij} = \beta_{0j} + \varepsilon_{ij}$$

$$\beta_{0j} = \gamma_{00} + \sum_{k \geq 1} \gamma_{0k} W_{jk} + u_{0j}$$

Model 2 (school climate):

$$Y_{ij} = \beta_{0j} + \varepsilon_{ij}$$

$$\beta_{0j} = \gamma_{00} + \sum_{k \geq 1} \gamma_{0k} W_{jk} + \gamma_{0k+1} R_j + u_{0j}$$

Data were analyzed using SAS PROC MIXED. The significance tests are Type III, that is, they test the significance of a specific effect given that all the other effects are incorporated in the model.