HOW DO SCHOOL PRINCIPALS UNDERSTAND AND USE THE STATISTICS IN REPORTS FROM NATIONAL LARGE-SCALE ASSESSMENTS

Iddo Gal[1] and Hani Shilton[2]
[1]University of Haifa, Israel
[2]National Authority for Measurement & Evaluation (RAMA), Ministry of Education, Israel
iddo@research.haifa.ac.il

This paper describes selected results from a project by the National Authority for Measurement & Evaluation in Education (RAMA) in Israel that examined how school principals understand the statistical information in national assessment reports they work with, and use the results for school improvement. The paper presents a new multidimensional conceptual model developed to guide the project and preliminary results based on interviews with school principals and supervisors. Based on the results, we reflect on the nature of the complex interpretation tasks faced by school principals who use reports on comparative large-scale assessments involving both achievement tests and school surveys, and on the link between principals' statistical literacy and evidence-based management practices as school leaders. The study examines implications for statistics education and professional development of school principals, and for the design of statistical reports.

INTRODUCTION

Most developed countries conduct national educational assessments or tests in multiple subject areas and grade levels. Results from such surveys or testing, which supposedly reflect achievement levels of students and schools and other aspects of students' attitudes and school climate, can inform critical decisions by school management regarding educational priorities and allocation of resources. Some prior research has examined how school teachers understand and use assessment to inform instruction at the individual or class level (e.g., Lyon, 2013). However, little is known about how well *school principals* understand the statistical information in such reports, what insights they gain from the analysis of such reports, and how such insights are used to inform managerial decisions, such as regarding the allocation of resources, assignments of teachers and other personnel, or other areas of school operation and management (Bair & Enomoto, 2013).

The lack of prior research on principals' understanding of results from national testing programs is surprising in light of the increasing interest around the world in evidence-based management, and the critical role that principals play as educational leaders in affecting a whole school and all its educational staff, students, and their community. There are mounting expectation in many countries that internal school policies will be based, to the extent possible, on solid and relevant data regarding students' achievement levels or educational performance, and regarding other school-related factors and correlates that may affect students' behavior and learning.

This paper describes a project by the National Authority for Measurement & Evaluation in Education (RAMA) in Israel that examined how school principals understand the statistical messages in national assessment reports they work with, and use the results for school improvement. The study was initiated as part of RAMA's emphasis on continuous improvement and its interest in seeking ways to improve statistical reports and its services to key information clients. The paper presents a multifaceted evaluation model developed to guide the project, and results based on interviews with school principals and supervisors.

BACKGROUND: NATIONAL EDUCATIONAL ASSESSMENT IN ISRAEL

Until recently, the Israeli school system has employed a national evaluation model called "Meitzav" (In Hebrew, an acronym of "Indicators of school effectiveness and growth"). In that system, students in grades 5 (in elementary schools) and grade 8 (in middle schools) were assessed in four subject areas: Home Language [Hebrew or Arabic], Science and Technology, Mathematics, or English. (Students in Grade 2 were assessed in Home Language only). In addition, both students and teachers fill extensive surveys regarding many aspects of the school climate and learning environment. Schools participate once every 2 years in an "external" assessment cycle whose results are reported to the public and to the Ministry of Education, as part of which achievement in two of the four subject areas listed above is tested. In parallel, in such a year the school implements

an internal assessment (i.e., not reported to the ministry), testing in the two other subject areas. In non-external years (i.e., every 2 years), a school runs a full "internal" assessment serving only the school and not reported outside, testing in all the four core subject areas.

As a result, a school principal would receive every two years an external detailed official report on the school performance in that assessment cycle. A typical report is a hefty information product; it holds 85-95 pages, including over 30 tables and 20 graphs and additional texts. Key elements of the report cover the following issues:

- *Attainment*. Statistical information about the performance of the students in the two subject areas included in the last assessment cycle (i.e., either Home language + Science and Technology; or English + Mathematics). The statistical information covers many issues and is conveyed through a combination of text, numerical displays (e.g., counts, percents, means & standard deviations on raw and standardized scales, etc) and various graphical displays, mainly bar charts. In many of these tables and graphs, comparisons are provided between the school statistics and those for other schools in Israel, using various types of classifications or tabulations (e.g., to schools speaking the same home language or to all schools in the nation, by SES level of the students, by the performance of the school on the same subject area in the last external assessment, using standardized scores, etc).
- *Climate*. Many tables and graphs show information about results of the surveys of school climate among pupils and teachers, on diverse topics such as attitudes towards the school and towards the learned subject areas, school safety, homework, and other topics.
- *Notes on interpretation and usage of the results*. Interspersed through the report is a lot of supporting text explaining the content of various tables and graphical displays, explanatory notes and reminders on how to read and make sense of key tables and statistics, caveats to take into account in order to avoid mis-interpretation of certain indices, and suggested general questions to keep in mind that aim to help the readers in linking the results to actual school processes such as in terms of resource allocation and teaching practices.

With the above in mind, overall, the national assessment report that principals have to read, interpret, and use is a complex information product which has much potential to inform their decisions and managerial practices, but which presents various challenges due to its multiple elements. Despite the centrality of such reports, we were unable to identify in the literature solid prior efforts to study school principals' understanding of the statistical information in such reports and how they transform this information into insights and action.

RESEARCH MODEL

In light of the above, this study was designed as exploratory research, using interviews with school principals and some supervisors of principals. The choice of a qualitative methodology based on interviews seemed self-evident, in light of the practical problems associated with administering any form of written structured test to school principals, and the need to build rapport and collect information about many processes and considerations that are unique to each school, explained below, in order to improve the validity of the information collected.

To further inform the planning of this study, based on a literature review and our own experience, we designed a new conceptual model that assumes that a principal's ability to make good use of the results of a national assessment and its associated report depends on three separate but related components:

- *Component 1 - Statistical aspects*: Based on Gal (2002), Gal & Murray (2011), and other sources, we assume that the ability of a user of a statistical report to read and comprehend the statistical aspects of a report depend on the interaction between two groups of factors: *(1) task factors*, i.e., characteristics of the information, displays, and texts in the statistical product, and of the information space in which messages are found (e.g., explanatory notes), and *(2) person factors*, i.e., characteristics of the individuals engaged with messages or products, such as knowledge of statistics and mathematics, linguistic skills, prior experiences and habits of mind, and attitudes and beliefs.
- *Component 2 - Educational evaluation and testing*: Based on extensive literature on educational evaluation (e.g., Birenbaum, Kimron, Shilton & Shahaf-Barzilay, 2009; Zuiker

& Whitaker, 2013), and in light of the content of the assessment results presented to principals in Israel as described above, it is expected that full understanding of a report from a national assessment requires that principals are also familiar with principles of formative and summative assessment, and with norm-references and criterion-referenced testing, as well as other elements related to using evaluation data. This for example stems from the reality that testing results are reported to principals in terms of student scores on a regular 0-100 scale (a criterion-referenced system, similar to how test results are used by teachers in Israel), yet principals also have to work within a norm-referenced system and understand comparative information, e.g., about the school standing compared to other school in terms of percentiles, or standardized scores when achievement is compared across years.

- *Component 3 - evidence-based school improvement processes*: The interpretation of the national assessment report further requires understanding of how change processes can occur and can be initiated and led within schools, using educational evidence. This requires among other things the need to interpret qualitative and qualitative information, integrate information from multiple sources, evaluate possible explanations and raising and testing hypothesis about possible factors affecting student performance or teacher performance, deriving implications and translating them into school-wide decisions.

METHOD

Based on the above model, we designed a semi-structured interview and administered it to 18 principals from an intentional sample of a wide range of schools, including for example schools in the top and bottom 10% on national assessments, from Hebrew-speaking and Arab-speaking urban and rural centers, and principals with different level of prior experience with national assessment results. In addition, two supervisors who mentor and monitor groups of schools and principals were also interviewed. The interviews lasted about 1.5 hours. The protocol included questions about the school background and principal's own work, and how reports are used within the school. Each principal was also presented with several key tables and graphs from the school's own most recent report (i.e., actual data he or she should be familiar with) and asked to explain his or her interpretation of the display and how this was used or how it informed educational decisions.

Figure 1 shows an example for one of the tables that appears in the national assessment report and shown to principals (though using actual data for their school). The table is written in Hebrew, read from right to left. It shows performance data in the home language for a certain grade level, broken down to 4 subareas (from top to bottom: reading and writing accuracy, reading comprehension, written comprehension, vocabulary, and total school score). For each of the 4 subareas and the total, the report shows, on a 0-100 scale, the average score and the standard deviation (below, in parentheses) for the school (in this case for 51 students) and for all schools of same home language (leftmost column). Additional explanatory notes appear below the table.

בתי"ס דוברי עברית*	בית ספרך מספר הנבחנים: 51	הציון המדווח	
70 (26)	70 (23)	ממוצע (סטיית תקן)	דיוק בכתיבה ובקריאה של מילים
76 (18)	80 (16)	ממוצע (סטיית תקן)	הבנת הנקרא**
70 (20)	72 (18)	ממוצע (סטיית תקן)	הבעה בכתב
80 (22)	86 (15)	ממוצע (סטיית תקן)	ידע לשוני
74 (17)	77 (14)	ממוצע (סטיית תקן)	הציון הכולל במבחן

לוח 3 : ההישגים במבחן בעברית לכיתות ב' תשע"ב - במבחן השלם ולפי נושאים

* הנתון המתייחס לסטיית התקן של קבוצות התקן של ההשוואה הוא ממוצע סטיות התקן של בתי"ס הנכללים בקבוצה זו.
** הציון בהבנת הנקרא מורכב מפריטים העוסקים בהבנת הנקרא בטקסט קצר ו...בהבנת הנקרא בטקסט ארוך.

Figure 1. Sample table from an actual assessment report

RESULTS

The analysis of the data is still underway, hence selected details will be reported at the ICOTS9 conference. Examples for key findings:

- The principals interviewed demonstrated a very wide range of interpretations associated with the assessment reports in general and with specific statistical elements in particular.
- Some of the principals interviewed appear to lack sufficient understanding of basic statistics or reach fragmented conclusions based on tables and graphs, although they should be familiar with these in light of their prior experience with the report.
- National assessment reports place many demands for *reading data in context* and for *integration of information* from multiple parts of the report (in text, numbers/statistics, and changing graphical and tabular displays). The need for integration presents numerous challenges to some principals and to their staff. Such integration is essential for generating a comprehensive picture and for examining possible explanations for the multi-faceted results of the assessment, pertaining both to test results and school climate data.
- We found *heterogeneous teamwork practices* in schools (e.g., collaborative interpretation of reports by the senior members of the school staff), and these affect the quality of the interpretation of the reports and their translation into decision making in schools.

DISCUSSION

While the study is based on a small, non-representative sample, its findings shed new light on how school principals understand the statistical messages in national assessment reports they work with, and point to a link between school principals' statistical literacy and interpretive practices and their evidence-based management practices. The pattern of results seems detailed and consistent with what is known about statistical literacy in the general adult population (e.g., Gal, 2002). Further, the results lend support to the three-component conceptual model developed to guide the study; they suggest that there are unique statistical literacy task demands related to working with reports of comparative assessment results, yet further that statistical aspects are just one of several knowledge bases that should co-exist and interact in the work and managerial practices of school principals and their teaching teams (Bair & Enomoto, 2013).

This study points to several implications for statistics education in general and for professional development of school principals in particular. The study helps to raise both new questions as well as suggestive directions regarding the design of statistical aspects of large-scale assessment reports that are expected to inform decision-making in schools, and regarding future research in this regard. The study also highlights the unique role that *supervisors* of principals could play in improving the interpretation of the information in the reports and translating them into action. Finally, the study helps to identify training needs of school principals as well as a need for supporting technologies that can together help school principals to improve interpretation and integration of information from complex statistical reports that are an inherent part of their job.

REFERENCES

Bair, M. A., & Enomoto, E. K. (2013). Demystifying research: What's necessary and why administrators need to understand it. *NASSP Bulletin*, *97*(2), 124-138.

Birenbaum, M., Kimron, H., Shilton, H., & Shahaf-Barzilay, R. (2009). Cycles of inquiry: Formative assessment in service of learning in classrooms and in school-based professional communities. *Studies in Educational Evaluation*, *35*(4), 130-149.

Gal, I. (2002). Adults' statistical literacy: Meanings, components, responsibilities. *International Statistical Review*, *70*(1), 1-25.

Gal, I., & Murray, S. (2011). Users' statistical literacy and information needs: Institutional and educational implications. *Statistical Journal of the IAOS, 27*(3-4),185-195.

Lyon, E. G. (2013). Conceptualizing and exemplifying science teachers' assessment expertise. *International Journal of Science Education*, *35*(7), 1208-1229.

Zuiker, S., & Whitaker, J. R. (2014). Refining inquiry with multi-form assessment: Formative and summative assessment functions for flexible inquiry. *International Journal of Science Education*, *36*(6), 1-23.