

Review Essays / Essais critique

David Johnson. (2005). *Signposts of Success: Interpreting Ontario's Elementary School Test Scores*. Policy Study 40. Toronto: C. D. Howe Institute. 235 pages. ISBN:0-88806-649-X; ISSN: 0832-7912

W. Todd Rogers, professor, Faculty of Education, University of Alberta
Don A. Klinger, professor, Faculty of Education, Queen's University

Over the last four decades several scholars have studied factors that influence student and/or class/school academic performance (e.g., Raudenbush & Byrk, 2002; Willms, 1992). These efforts are part of the growing monitoring and accountability functions that mark education. Most of the provinces use some form of province-wide testing as part of their accountability systems. At the national level, the Council of Ministers of Education, Canada, administers national tests in the areas of reading, writing, mathematics, and science. Internationally, Canada and her provinces participate in the Programme for International Student Assessment (PISA) in these same subject areas.

Of concern is the use of the data to rank schools based on student performance. The best known rankings in Canada are those reported by the Fraser Institute (Cowley & Easton, 2003). The Fraser Institute rankings have been seriously questioned for their failure to adequately take into account school context. Although the Fraser Institute uses procedures to try to account for differences in socio-economic status to improve their rankings, the schools with students with high socio-economic status continue to outrank schools with students with lower socio-economic status.

In his book, *Signposts of Success: Interpreting Ontario's Elementary School Test Scores*, David Johnson (2005) employs regression procedures to separate out socio-economic factors from other factors that explain student performance and to control for socio-economic status when ranking schools. The first two chapters lay the foundation for the study Johnson completed. The first chapter addresses the need to ensure that comparisons among schools are valid and not open to misinterpretation

due to the presence of unaccounted for socio-economic differences among schools. The factors included in the study are then identified. Unfortunately, he does not acknowledge that the use of SES alone may in itself be a limitation as other relevant contextual variables may also be important. Chapter 2 contains (a) a brief history of “jurisdiction-wide” assessment programs in Canada and the United States followed by a brief, but informative description of the activities, costs, and the public face of and reactions to the province-wide assessments conducted in Ontario by the Education Quality and Accountability Office (EQAO), and (b) a brief review of previously completed studies of the variance accounted for among schools on the Ontario assessments and similar studies conducted in other jurisdictions.

The next four chapters contain the methodology and the results and a discussion of the results. Chapter 3 describes how the values of the socio-economic variables, identified in Chapter 1, were determined for each school in the samples using census data and the postal codes of the students who attended the school and correctly identifies the limitations of using the census-postal code approach. Chapter 4, the centre-piece of the book, contains a description of the regression analyses (presented in an Appendix to the chapter) and the application of these analyses to determine which of the factors described in Chapter 3 explains the variation in the proportion of students at Level 3 and Level 4 (*The Ontario Curriculum* sets Level 3 [student meets curriculum standards] as the provincial standard). In brief, Johnson suggests that approximately 25 per cent of the variation in the school proportions of students at Levels 3 and 4 is accounted for by the factors considered (pp. 73; 107). Unexpectedly and serving as a distraction, the influence of the same factors on the rate of exemptions is reported in Chapter 5. Chapter 6 (with its Technical Appendix) returns to the proportions of students who meet the curriculum standards and compares the Fraser Institute school rankings with school rankings based on Johnson’s method in which schools with a similar socio-economic standing are compared. Interestingly, Chapter 7 describes a study in which principals, teachers, and parents in 13 schools that outperformed other schools with a similar socio-economic status were interviewed to identify factors other than socio-economic status that led to increased performance. Lastly, Chapter

8 contains the conclusions reached on the basis of the findings presented in Chapters 4, 6, and 7, and a personal assessment of the effectiveness of the province-wide assessment in Ontario and recommendations for improving the process.

As indicated, the centre piece of *Signposts of Success: Interpreting Ontario's Elementary School Test Scores* is the fourth chapter. Two major analyses are conducted. The first involved an analysis of the trends in average achievement across time. Although Johnson devotes a number of pages in the main text and the appendix of this chapter to the analyses and interpretation of the trend analysis results, Johnson appears to dismiss them, pointing out that "They are not, however, the effects of central issue in our analysis" (p. 72). Johnson correctly suggests that interpreting the trend results might be problematic because different test forms, built using the same test specifications, may differ somewhat in difficulty. This points to the need to equate before the test results can be compared between years. Equating adjusts for differences in difficulty among forms that are constructed to be similar in difficulty and content (Kolen & Brennan, 2004). Since Johnson did not do this, the results of the trend analyses cannot be validly interpreted.

Closely aligned with the failure to equate tests prior to conducting a trend analysis is the questionable formation of the coded ("dummy") variables in some but not all of the trend analyses and all of the coded vectors in the regression analyses. To illustrate this concern, it is necessary to consider the basic regression equation used by Johnson:

$$C_{S,i,t} = \alpha + \delta_{i,t} D_{S,i,t} + \beta_{X_j} X_{S,j} + u_S + u_{S,i,t},$$

where $C_{S,i,t}$ is the proportion of students of students at Levels 3 or 4 at School S on assessment i (reading, writing, or mathematics in either Grade 3 or Grade 6) in academic year t (1998-99, 1999-2000, 2000-2001, or 2001-2002);

$D_{S,i,t}$ is a coded variable that has a value of 1 when the assessment at the school is in a specific assessment and zero if not;

$X_{s,j}$ is a measure of factor j , $j = 1, 2, \dots, k$, k the number of factors; and u_s and $u_{s,i,t}$ are, respectively, the variation among schools that is not explained by the factors considered and the coded variables.

In some of the trend analyses, some of the coded variables included comparisons between the performances on two different subject areas (e.g., 2000-01 with 1998-99). Although Johnson suggests these coded vectors have “a clear interpretation” (p. 71), they do not. First, these contrasts have different subject areas and, second, differences in the difficulty of the two tests will serve to confound differences in unpredictable ways.

The second set of analyses involved a series of regression analyses. First, it is not clear why step-wise regression analyses were not conducted. Generally the intent of a regression analyses in the present context is to find the set of predictor variables that maximally accounts for variation in the dependent variable. All factors were included in the analyses presented in the book. Second, there is insufficient justification to warrant reporting average results across subject area and school year. One might want to argue that if the intent was to reduce the number of results, then consideration might be given to combining across years within subject area. However, even this might be questionable if the trends across years noted by Johnson within subject area persist after equating. Further, the mean of four observations is subject to the presence of extreme results or differentially affected by differences in the variance of scores in different years or subject areas. The absence of separate results prevents the reader from assessing the consistency of the results across subject area and year and the potential adverse influence of extreme values. Given the lack of separate results, and the problems with the coded variables, it is not possible to interpret validly the results presented.

It is also not clear why Hierarchical Linear Model analyses, with students nested within schools (Raudenbush & Byrk, 2002), were not used. The concern in education is ultimately with the student.

Improvement of student performance will naturally lead to improvement in school performance.

Lastly, the interview study itself is suspect. First, given the usual problems with self-report interview data, this portion of the study would have been stronger had observational data been collected. Second, the interview results would have been much more interpretable if a set of low performing schools had been included. The results from both sets of schools would then illuminate the results of each other and point out differences, if any, between school and teacher practices amenable to change and the influence of policy initiatives. The absence of low performing schools prevents an unequivocal interpretation of the Chapter 8 results.

David Johnson states that a major foundation of the book was to provide comparisons among schools that were valid and not open to misinterpretation. A review of the regression approach he used to assess trends within and across subject areas is seriously incomplete and fails to meet his own standard. The use of a single level regression model and the questionable practices when conducting the regression analyses seriously undermines valid interpretation of the results presented by Johnson. Nor can it be forgotten that the rankings are based on results that only account for 25 per cent of the variability in the scores. Hence the quest to explain and understand the variation in student and school performance continues. Attention still needs to be given to the measurement of predictor variables at the student level that yields valid and interpretable information. Likewise, there will be advantages if multilevel models (e.g., HLM) are used. Such models are much more isomorphic with school structure and practice.

REFERENCES

- Cowley, P., & Easton, S. T. (2003). *Report card on Ontario's elementary schools, 2003 edition*. Vancouver, BC: Fraser Institute
- Kolen, M. J., & Brennan, R. L. (2004). *Test equating, scaling, and linking: Methods and practices*, (2nd edition). New York: Spring Business+Business Media.

Raudenbush, S. W., & Bryck, A. S. (2002). *Hierarchical linear models: Applications and data analysis methods*, (2nd edition). Thousand Oaks, CA: Sage Publications.

Willms, J. D. (1992). *Monitoring school performance: A guide for educators*. London, UK: Falmer.

Moira Monteith (Ed.). *ICT for Curriculum Enhancement* Bristol, England: Intellect Ltd., 2004. 160 pages. ISBN 1-84150-061-5 (paperback).

Bruce L. Mann, professor, Faculty of Education, Memorial University

This book follows from Ms. Monteith's earlier publication *ICT and Learning Enhancement*, which focused on ways in which Information Technology (IT) promotes learning by encouraging reflection among teachers in sharing good teaching practice in the UK. This book review will focus only on the latter work, namely *ICT for Curriculum Enhancement*.

Moira Monteith's introductory chapter in *ICT for Curriculum Enhancement*, "Re-modeling education" could easily be titled "how to improve one's teaching practice through self-assessment." Monteith also makes modeling with students in classrooms an issue, "...any learning model we employ also build in the valuable asset of teacher's pragmatism" (p. 18). Throughout the chapter, "encouragement" seems to be the key construct, never a bad thing in education. What I missed, however, was one or two shining examples, a really good school model, or even a promising theory. I was surprised with the omission of other models of school improvement in the United Kingdom, for example, Alma Harris in the Centre for Teacher and School Development at the School of Education, University of Nottingham (Harris, 2000). And UK expatriate Tony Bates (1999) provides a thorough "How-To" plan, covering the importance of leadership, vision, and strategic planning, using a task force for change, and highlights several approaches to introducing technology into teaching. In the United States for example, the Co-nect for K - 12 schools model of school reform (2004) has

customized on-line/on-site training and personal support, recommends project-based learning, peer and progress review programs, leadership processes for whole-school technology integration. And what of the previous notable work on school reform theory by well-respected change theorists like Michael Fullan or Michael Apple. In fact, the chapter skims lightly over many important concepts, including "learning" and "cognition," and uses uncommon terms (i.e., "computer cognition") without definition or discussion. I interpreted her meaning as "student motivation for technology," as in Papert's fascination with his motorcycle, or Turkle's child programmers.

The strengths of Martine Owen's chapter "Just a tool: The computer as the curriculum" reside in statements about ongoing changes to the learning environments due in part to technological progress -- from solitary activity to social constructions of knowledge. In some ways, Owen's chapter is reminiscent of Marshall McLuhan's ideas (Sanderson & Macdonald, 1989). Owen writes, "Changing the technology causes us to think differently about what we can know from books. What we can teach from paper is mediated by the technologies and systems we have devised for the construction and storage of printed media" (p. 28). Owen's message is that there is an inevitable impact of technological progress such as writing and later print technology on the oral tradition of education. Aside from its likeness to early literary opinion, this chapter reads like a set of anthropological notes: on Bateson's learning types: Learning I, Learning II, Learning III, on Robins and Webster's Mode 1 and Mode 2 knowledge, on Engestrom's Activity Theory, or even Skinner's views on operant conditioning.

In contrast to the general overviews offered throughout this book, Peter Twining's chapter "Computer practice framework: A tool to enhance curriculum development in ICT" is a systemic treatment of numerous factors. The "computer practice framework" is his three dimensional approach to generating student practice with a computer. His three dimensions are "quantity," "focus," and "mode." Under "quantity" Twining mentions available learning time or proportion of the school day, though I would've liked to see the *how* the time would be used as well -- also known as "quality of time on task." A missing dimension, perhaps, is the number of computers (networked lab vs. one

per classroom vs. three or four computers), and how they have been configured (classroom vs. lab). This issue was introduced in the 1990's and brought out fully in technology integration textbooks. Twining's modal approach to curriculum is reminiscent of writers in the late 1980's who preferred to think of knowledge structures, and students mentally acquiring and accommodating those structures through practice. Practice has always been closely linked to feedback, which is missing in this chapter. This chapter could have been much improved with a more comprehensive review of the literature, especially in technology integration in classrooms.

Libby Jared's "Curriculum enrichment: Using online resources balancing creativity with the readily available" is a broad overview of what a teacher can do with a connection to the Internet. Jared could have much improved this chapter with a more comprehensive review of the literature, especially methods of classifying online learning resources. For example, Mann (2005, 2000, 1999a, 1999b,) classified fifteen types of web-based resource-based teaching frequently used by instructors, and categorized by purpose, adapted from a taxonomy of conventional learning resources developed in the UK in 1994 by Parsons and Gibbs.

I enjoyed reading about the instructional and assessment strategies discussed in Pilkington and Kuminek's chapter "Using role-play activity with synchronous CMC to encourage critical reflection on peer debate." First I liked the "issues for discussion". Second, I liked the "roles given to students" and the way it was reported in this chapter, I also liked the method used of student tallying as a quick means of self-assessment. But the authors ignored several issues. They make no distinction in mixing synchronous and asynchronous data, which of course is different. Pie charts assume there is a finite number, a ceiling on numbers of postings from each student share, which is not the usual approach to analyzing discourse. Three of the seven students that were absent in session B chat subsequently submitted their views into the asynchronous discussion board, and their data included (p. 76). I was surprised that the instructors of a 12-week Psychology and Cognitive Science course in Language, Cognition, and Collaborative Learning would choose to assign their students two media comparison studies (i.e., by Sullivan and Pratt, and by Warschauer), given the long-known evidence of

confounding in comparing one medium such as classroom instruction with another such as learning from a computer program (Bernard et al., 2004; Clark, 2001; Mayer 2001). I can only assume it was to arouse discussion about the problems in media comparison research. Another surprise was to see that the authors had published the full names of the participants in their research. It is well established in educational research that the identity and privacy of research participants should be protected. "The principle means of ensuring anonymity is not using the names of the participants" (Cohen, Manion, & Morrison, 2000, p. 62). To my understanding, this appears to have compromised ethical standards.

Jocelyn Wishart in her chapter "Computers in schools and colleges: The user in control," discussed four projects with different populations of learners. She concluded, "...the key factor for the successful application of ICT in the curriculum is the enabling of user control over the computer and consequent learning through it" (p. 85). Setting aside for a moment the confounding errors in these "research projects" (i.e., hypotheses, randomization, control group, etc) and considering them as either separate illuminative case studies (Stake, 1994) or multiple case studies (Yin, 2003), I still don't see how the cases support the conclusion: User control over what – access? Program control? Pacing? Also, Wishart should have provided a rationale for reporting these studies in this way, given the long-running concern over the effects of media on learning, and motivation to learn (Clark, 2001; Mayer 2001).

The chapter "ICT and curriculum provision in early years" by Deidre Cook reflects, I think, the true aim of this entire book, which is not ICT in curriculum, but rather "literacy and ICT", or as Cook says "child-friendly software to support appropriate literacy work" (p. 106) and "computer-writing mirrors" (p. 107). I was hoping that Cook would have more deeply investigated either the social aspects of children using computers (as Sherry Turkle has done in her work, for example). I would have like to read more about the developmental complications of integrating technology into "the typical learning environment" (p. 103).

Regarding the chapter "Capability and initial teacher training" by Dore and Wickens, I would have suggested adding a subtitle, "the role of student prior knowledge on learner control." If I were an editor, I would also have asked the authors to conclude the chapter with statements that

linked their observations more closely to those of Martyn Wild and Diana Laurillard, paraphrased and cited earlier in this chapter.

I don't have much to say about John Chatterton's chapter "ICT as a specialist subject in initial teacher training" other than it appears to be Chatterton's own reflections on his class activities, using his own terminology (e.g., "phased cognitive acts," p. 132), without a rationale, a research design, or even informed opinion from peer-reviewed sources. His reference list simply shows a series of government policy documents and unpublished material on a personal website.

From the title of Tony Fisher's chapter "Information society, situatedness, and social construction: Student teachers learning on a PGCE geography course," I had expected to read about best-practice teaching of geography education in the UK, maybe something about "problem-based learning" used to great effect elsewhere. Instead, Fisher offered three instructional activities somewhat disconnected from educational theory. The first activity described in the chapter was a WebQuest, the standard fare of teacher web education since the 1980's, somewhat old fashioned by today's standards. The second student activity described in the chapter was that these student geography teachers were assigned the GeoSkills program "to reinforce geographical skills in relation to map reading." I found the "reinforcement" term troubling amidst so much cognitive terminology as "situatedness", "authentic tasks" and "authentic assessment"—a little confusing. Specifically, I would've liked to know more about the how features in the GeoSkills program were implemented to enable students to better learn map reading. Or was it "map reading"? Fisher's description said "geographical skills in relation to map reading," a description that was a little ambiguous. A statement about "instructional design" as an objectivist activity really dated the literature for me, given recent data-driven theories that have informed contemporary instructional design of the past fifteen years, such as Mayer's "Cognitive Theory of Multimedia Learning" (Mayer, 2001), Sweller's "Cognitive Load Theory" (Paas, Renkl & Sweller, 2003), van Merriënboer's "4C-ID Model" (van Merriënboer, Kirschner, & Kester, 2003), Reigeluth's "Elaboration Theory" (Reigeluth, 1999), Jonassen's "Constructivist Learning Environments" (CLE's), (Jonassen, 1999), and Hannafin's "Open Learning

Environments" (OLE's) (Hannafin, Land & Oliver, 1999), to mention only a few. In sum, this treatment left me wanting more. Instead of situating his instructional approach within geography education methods, Fisher simply outlined recent policy statements from Annex B, and described his classroom practice.

Vanhegan and Wallace's chapter, "IT as a skill for teachers: The delivery of IT as an integral component of a fulltime PGCE programme" provided a description of plans decreed by the FENTO since 2000 (Further Education National Training Organisation). The FENTO has been responsible for producing occupational standards for everyone involved in delivering further education. In this chapter, Vanhegan and Wallace followed a cohort of 41 prospective lecturers through a Further Education course to gauge "their motivation and associated learning and teaching styles" (p. 154). The authors used student journals as data to determine how these students used the IT component of their professional development in ICT. The information of FENTO in their tables is very general for immediate application. For example, "Produce appropriate learning support materials using information learning technology" and "Use information technology and learning technology as appropriate" (p. 154). Their approach to bringing students up to the FENTO standard was to enlist peer tutoring as a strategy. Teacher training in ICT in the United Kingdom, as in other countries, has become standardized by government-funded, non-profit organizations (EdGov, 2002; e-Learning Action Plan, 2005; European Teacher's Portal, 2001). Similarly, in the United States there are the National Standards for Technology in Teacher Preparation (the NETS-T) developed by the International Society for Technology in Education (ISTE) and The National Council for Accreditation of Teacher Education (NCATE). The interesting part of the chapter is the description of the FENTO in the UK, use of questionnaire has become standard fare to gauge student opinion. I would have liked to have seen an integration of the results with previous studies under the FENTO regime as well as what could be said about it in the light of the students' comments?

From its outset, Monteith's book *ICT for Curriculum Enhancement* looked like it had the potential to be a winner, a real resource for teachers and researchers in the United Kingdom and elsewhere. As I

read each of its chapters, however, I was frequently reminded of chapters in other publications with the same "literacy" focus. In this light, a more appropriate title for this work would have been "screen-based literacies," "literacy and ICT," "screen-based communications," or even "the changing structure and nature of grammar." Every chapter could have been much improved with an updated review of literature. Given so many concerns in its current form, I cannot recommend this book.

REFERENCES

- Bates, A. W. (1999). *Managing technological change: Strategies for academic leaders*. San Francisco: Jossey Bass.
- Bernard, R. M., Abrami, P. C., Lou, Y., Borokhovski, E., Wade, A., Wozney, L., Walseth, P., Fiset, M., & Huang, B. (2004). How does distance education compare with classroom instruction? A meta-analysis of the empirical literature. *Review of Educational Research, 74*(3), 379-439.
- Chandler, P., & Sweller, J. (1991). Cognitive load theory and the format of instruction. *Cognition and Instruction, 8*, 293-332.
- Clark, R. E. (Ed.). (2001). *Learning from media: Arguments, analysis and evidence*. Greenwich, CT: Information Age Publishing.
- Cohen, L., Manion, L., & Morrison, K. (2000). *Research methods in education* (5th ed.). London, UK: Taylor & Francis, Inc.
- e-Learning Action Plan. (2005). Designing tomorrow's education. European Commission. Retrieved 01 February, 2005, from http://europa.eu.int/comm/education/programmes/elearning/index_en.html
- EdGov. (2002). No Child Left Behind, US Department of Education. Retrieved 01 February, 2005, from <http://www.ed.gov/>
- European Teacher's Portal (2001). European Schoolnet. European Commission. Retrieved 01 February 2005 from <http://eschoolnet.eun.org/ww/en/pub/eschoolnet/index.htm>

- Hannafin, M., Land, S., & Oliver, K. (1999). Open learning environments: Foundations, methods, and models. In C. Reigeluth (Ed.), *Instructional design theories and models: A new paradigm of instructional theory*. Vol 2 (pp. 115-140). Mahwah, NJ: Lawrence Erlbaum Associates.
- Harris, A. (2000). Successful school improvement in the United Kingdom and Canada. *Canadian Journal of Educational Administration and Policy*, 15. Retrieved 12 September, 2005, from <http://www.umanitoba.ca/publications/cjeap/articles/harris.html>
- ISTE (2004). National educational technology standards for teachers (ISTE). Retrieved 01 February, 2005, from <http://cnets.iste.org/teachers>
- Jonassen, D. (1999). Designing constructivist learning environments. In C. Reigeluth (Ed.), *Instructional design theories and models: A new paradigm of instructional theory*. Vol 2 (pp. 215-240). Mahwah, NJ: Lawrence Erlbaum Associates.
- Mann, B. (2000). Phase theory: A teleological taxonomy of web course management. In Bruce L. Mann (Ed.), *Perspectives in web course management* (pp. 3-26). Toronto: Canadian Scholar's Press.
- Mann, B. (1999, Fall). Web course management in higher education. *Bulletin of The Canadian Society for the Study of Higher Education*, 14, 7 - 11.
- Mann, B. (1999b). Web course management. *Australian Educational Computing* 14(1), 15-20.
- Mann, B., Newhouse, P., Pagram, J., Campbell, A., & Schulz, H. (2002). A comparison of temporal speech and temporal text cueing in educational multimedia. *Journal of Computer-Assisted Learning*, 18(3), 296-308.
- Mann, B. (2005). Action research with Internet database tools. In Bruce L. Mann (Ed.), *Selected styles in web-based educational research* (pp. 179-187). Hershey, PA: Idea Group Publishing.
- Mayer, R. E. (2001). *Multimedia learning*. Cambridge, UK: Cambridge University Press.

- Merrill, M. D. (1999). Instructional transaction theory (ITT): Instructional design based on knowledge objects. In C. Reigeluth (Ed.), *Instructional design theories and models: A new paradigm of instructional theory*. Vol 2 (pp. 397-424). Mahwah, NJ: Lawrence Erlbaum Associates.
- Paas, F., Renkl, A., & Sweller, J. (2003). Cognitive load theory and instructional design: Recent developments. *Educational Psychologist* 38: 1-4.
- Reigeluth, C. M. (1999). Elaboration theory: Guidance for scope and sequence decisions. In C. Reigeluth (Ed.), *Instructional design theories and models: A new paradigm of instructional theory*. Vol 2 (pp. 424-454). Mahwah, NJ: Lawrence Erlbaum Associates.
- Sanderson, F., & Macdonald, F. (1989). *Marshall McLuhan: The man and his message*. Golden, Co: Fulcrum.
- Stake R. (1994). Case studies. In N. Denzin & Y. Lincoln (Eds.), *Handbook of Qualitative Research* (2nd ed., pp. 435-454). London, UK: Sage Publications.
- Van Merriënboer, J. J. G., Kirschner, P., & Kester, L. (2003) Taking the load off a learner's mind: Instructional design for complex learning. *Educational Psychologist*, 38: 5-13.
- Yin, R. (2003)). *Case study research: Design and methods*. Beverley Hills, CA: Sage Publications.