

What Lies behind Graphicacy? Relating Students' Results on a Test of Graphically Represented Quantitative Information to Formal Academic Achievement

Lisbeth Åberg-Bengtsson,¹ Torgny Ottosson²

¹*Department of Education, Göteborg University, Box 300, SE 405 30 Göteborg, Sweden*

²*Department of Behavioural Sciences, Kristianstad University, Kristianstad, Sweden*

Received 27 December 2003; Accepted 22 December 2004

Abstract: Based on studies carried out on qualitative data an instrument was constructed for investigating how larger numbers of students handle graphics. This test, consisting of 18 pages, each with its own graphic display(s) and a set of tasks, was distributed to 363 students, 15–16 years of age, from five different schools. The format of the questions varied, as did the format of the graphics. As students' performance was expected to be multidimensional, confirmatory factor analysis was carried out with a structural equation modeling technique. In addition to the identification of a general graphicacy-test factor (*Gen*) and an end-of-test effect (*End'*), a narrative dimension (*Narr'*) was vaguely indicated. This model was then related to a six-factor model of students' formal academic achievement measured by their leaving certificates from compulsory education. The strongest correlation obtained was between the general graphicacy-test dimension (*Gen*) and a mathematic/science factor (*MathSc'*) in the grades model. In addition, substantial relationships were detected between the *Gen* factor and both an overall school achievement factor (*SchAch*) and a language factor (*Lang'*) in the grades model. © 2005 Wiley Periodicals, Inc. *J Res Sci Teach* 43: 43–62, 2006

Graphs, charts, cartograms, thematic maps, etc., are common tools for handling and communicating quantitative information in contemporary society. Successively, through their years of schooling, students encounter increasingly more advanced forms of graphic displays, both in direct educational situations and as illustrations and/or complementary facts on other subject matters. Students' access to modern computers equipped with graphic application software makes possible not only an abundance of graphics (sometimes unnecessarily elaborate) in magazines, newspapers, television, and electronic media—encountered in their everyday life both in and out of school—but also allows them to create these images. Thus, being “graphicate”

Contract grant sponsor: Swedish Council for Research in the Humanities and Social Sciences (HSFR).

Correspondence to: L. Åberg-Bengtsson; E-mail: lisbeth.aberg@ped.gu.se

DOI 10.1002/tea.20087

Published online 28 November 2005 in Wiley InterScience (www.interscience.wiley.com).

is becoming an important part of everyday knowledge, equal in status to being literate and numerate. In Sweden, the importance of this field in education is emphasized by the fact that the national syllabus in mathematics points out statistics (with graphics skills specifically mentioned) as one of the four main fields to be covered in education.

Of course, a substantial body of research on comprehension of graphics is available within educational psychology. Such studies, however, often deal with students' problems and misconceptions in their grappling with graphs, charts, and maps of different types or with the categorization of ways of making sense of such displays. Thus, although graphicacy is in no way an unexplored field of research, relatively few researchers (e.g., see Åberg-Bengtsson, 1999; Guthrie, Weber, & Kimmerly, 1993; Winn & Holliday, 1982) have dealt with issues as to which demands or abilities may be involved in the solving of graphic tasks. The investigation presented herein originates from a research project aimed at constructing a measurement device for the testing of graphicacy in compulsory school and for carrying out factor-analytic analyses by modeling data from students' performance on the test.

Some Previous Research

In the 1780s some of the most frequently used graphs and charts for presenting quantitative information were developed by a Scottish economist, William Playfair (Tufté, 1983). Although the inventions were useful for recording, analyzing, and communicating data, applied scientists at that time continued their obsession with tabular data and disregarded plotting and graphic analyses. A change in attitudes toward the new tools did not take place until the first half of the next century, when the usefulness of graphic displays for the social sciences was demonstrated and scientific journals began to record graphing and use of statistical cartography began to expand (Beniger & Robyn, 1978). Thus, graphic representations of numerical data are, compared with other types of graphic representations, such as maps for navigation, relatively modern. Possibly, interaction within a society must have reached a certain developmental level and the use of particular cultural and mathematical tools (such as the Cartesian grid) is needed for the invention and application of such devices for storing, communicating, and analyzing data.

The immediate reactions to Playfair's inventions were mixed (Wainer, 1980), as were the somewhat contradictory suggestions of our contemporary researchers concerning how easy or difficult graphic displays are to handle. On the one hand, it was suggested by Lewandowsky and Spence (1989) that most graphs are simple. Some research has focused on this simplicity. Wainer (1980), for example, having administered a test concerning some commonly used graphs and charts to 360 pupils, argued that children, by the age of 9 years, had, on average, reached the minimum acceptable level of an adult. Ainley (2000) reported intuitive reading of graphs among 6-year-olds as an example of transparency of certain aspects of graphing, which comes quite close to some of our own previous research (e.g., Åberg-Bengtsson, 1998; Ottosson, 1987, 1988; Ottosson & Åberg-Bengtsson, 1995), wherein the youngest elementary school pupils could handle particular features of commonly used graphs, charts, and maps quite adequately. Jones et al. (2000) have, from a neo-Piagetian perspective, identified four "thinking levels" for statistical thinking among pupils in grades 1–5.

A substantial body of research, however, has investigated students' difficulties (sometimes labeled "misconceptions") and maintained that understanding and aspects that go beyond the most obvious proportional relationships and a simple reading-off of values may be difficult even for older pupils (e.g., Preece, 1983) and university students (e.g., Bowen, Roth, & McGinn, 1999; Goldberg & Anderson, 1989; Lindwall, 1998). Even scientists expert in using graphs in their own research may have difficulties correctly interpreting similar graphical displays from other domains

(Roth, 2003). Researchers have pointed out the stumbling block of making so-called “iconic” interpretations or “reading the graph as a picture,” which occurs when the wrong spatial content is assigned to the display (e.g., Clement, 1989; Kerslake, 1981; Preece, 1983). However, the validity of some of these investigations has been called into question. For instance, Berg and Smith (1994) argued that the frequent occurrence of iconic misconceptions may, to a certain extent, be a function of the multiple-choice question format often used in such studies. A related and equally well-known phenomenon is the “height for slope confusion” (e.g., Preece, 1983; Roth, 2003) when, for example, the highest value is confounded with the steepest gradient. Other reported difficulties are problems with seeing the curve in Cartesian graphing as continuous, and other problems with scales and axes (e.g., Åberg-Bengtsson, 1998; Nemirovsky & Tierney, 2001).

A relatively large number of studies have focused on computer-aided, or computer-mediated, learning of graphics (e.g., Linn, Layman, & Nachmias, 1987; Lindwall, 1998; Lindwall & Ivarsson, 2004; Nemirovsky & Noble, 1997; Nemirovsky, Tierney, & Wright, 1998), whereas others deal with if—or to what extent—different types of graphic displays may enhance learning and recollection of subject matter (e.g., Kealy & Webb, 1995; Verdi, Kulhavy, Stock, Rittschof, & Johnson, 1996; Winn, 1991).

Research on graphics has been carried out from a number of quite different theoretical standpoints. The reading of graphs is, in the cognitivist tradition, regarded as information processing (Leinhardt, Zaslavsky, & Stein, 1990), or as the use of mental representations/models (e.g., Lowe, 2003; Schnotz & Bannert, 2003). Veriki (2002), in a review of studies aiming to explain the role of graphical displays in learning, maintained that this research “suggests that graphics are effective learning tools only when they integrate information with minimum cognitive processing” (p. 261).

Quite typically, people’s understandings of graphical representations are implicitly or explicitly regarded as relying on one or another cognitive ability (which often is, but does not necessarily have to be, related to a traditional cognitivist perspective). Obviously, it may then easily be assumed that solving tasks related to visual displays, such as graphs, charts, maps, cartograms, and diagrams of different types, relies mainly, or at least to a great extent, on visual or spatial abilities—a perspective supported by some researchers. Veriki (2002), for example, stated that “learners’ characteristics, such as prior subject-matter knowledge, visuospatial ability, and strategy, influence graphic processing and interact with graphical design to mediate its effects” (p. 261). Kozhevnikov, Hegarty, and Mayer (2002), in a study involving 60 undergraduate psychology students, compared, among other things, a group of “visualizers” (i.e., individuals relying primarily on imagery processes when performing cognitive tasks) with high spatial ability to a group with low spatial ability and found that this difference also reflected the dissociation between visual (iconic) and spatial imagery when solving time–position graphs. The difference remained when controlling for other factors such as mathematical background, general intelligence, and the use of metacognitive strategies.

Carrying out a study with a somewhat different focus, Winn and Holliday (1982) investigated individual differences in learning from texts provided with complementary flowcharts. They suggested that benefiting from diagrams and charts goes beyond the use of visual (or spatial) and verbal abilities and relies on more general and abstract powers of reasoning. In a later study, Winn (1993) maintained that there seem to be differences in strategies when searching for information in diagrammatic displays, compared with searching for information in texts that do not involve visual relationships. He indicated that, in diagrams, the initial stages are guided perceptually, because symbol systems and conventions are primarily spatial. On the other hand, he also noted that previous knowledge of the content of a diagram and knowledge of the symbol systems used in this type of display are supposedly the most important factors when directing

the search. Guthrie, Weber, and Kimmerly (1993) investigated cognitive processes in undergraduate students' understanding of graphic representations, such as graphs, tables, and illustrations, and found two main factors influencing performance. One factor had to do with students' abilities to locate specific information, whereas the other concerned perception of trends and patterns or the extraction of global information. As performance on this second factor was significantly lower than performance on the first factor, it was suggested that a substantial number of students had not learned abstraction processes related to the reading of information at an overall level.

Research on the Swedish Scholastic Aptitude Test (an entrance test to higher education), in which one of the subtests deals with diagrams, tables, and maps (DTM) exclusively, has directly or indirectly addressed the issue of factors involved in the interpreting of graphics. Gustafsson, Wedman, and Westerlund (1992) argued for two factors underlying performance on the SweSAT. Their results show that, although what was labeled a "knowledge" factor to a great extent influenced performance on tests like reading comprehension, study technique, and general information, the DTM subtest and a test measuring mathematical reasoning (the data sufficiency or DS subtest) were related mainly to what was interpreted to be a general analytic ability. However, there seemed to be one problem with the Gustafsson et al. results. In general, the SweSAT, particularly the DTM and DS subtests is characterized by a gender difference in performance in favor of males. The crux is that, in tests measuring general ability, no such difference is presumed (Halpern, 1992). Obviously, the involvement of a spatial factor, where boys usually achieve better than girls, would be one such plausible source for the deviation (e.g., Halpern, 1992). Focusing on the DTM subtest, Åberg-Bengtsson (1999) developed further the dimensionality of the internal structure of the SweSAT identified by Gustafsson et al. (1992). In the first of these studies a "quantitative" factor was identified in addition to a more general DTM dimension. Tasks that demanded not only the reading off of values or interpreting trends, etc., but also more or less complicated calculations, were related to this factor. Furthermore, the "quantitative" factor was found to be a "substantial contributor to the gender differences on the DTM test, if the selection effects...are taken into account" (Åberg-Bengtsson, 1999, p. 578). Åberg-Bengtsson maintained that this result makes sense, because males are often reported to perform better than females on mathematical factors, and because it may also explain the seemingly strange gender difference reported by Gustafsson et al. (1992). With the exception of an end-of-test effect, no other factors were identified despite the fact that a large number of hypotheses were elaborated and tested. However, because the DTM subtest involves a number of different graphic formats, as well as a number of different ways of posing the multiple-choice questions allowing for a range of strategies for the solving of the tasks, this complexity doubtlessly means that a number of intertwined factors are at play. Consequently, although not identified, the existence of a spatial ability dimension cannot be excluded (Åberg-Bengtsson, 1999).

Studies on graphing carried out within a sociocultural framework (e.g., Åberg-Bengtsson, 1998; Ainley, 2000; Roth, 2003) have focused on the interaction within the activity system or the context. Thus, the view on construing and interpreting graphic representations—or "inscriptions" if using a concept that has recently gained acceptance in science and technology—shifts from being something going on mainly in individual minds to taking place in the social arena (Roth & McGinn, 1998). When investigating university teachers' and experienced scientists' grappling with graphs from partly unfamiliar domains, Roth (2003) found that, among other things, familiarity with the content of the graph and conventions in the current discipline are important for interpretations in accordance with the normative expectations. Ainley (2000), introducing active computer-based graphing to elementary school education, maintained that important features for

facilitating transparency of graphs to children are related to the settings in which the graphs are presented rather than to the graphs themselves. One such feature was a “familiar and/or meaningful context” (p. 376).

Study Aims

From the findings just presented it seems obvious that identifying and investigating factors or abilities that may be involved in students’ interpreting graphic representations, such as graphs, charts, tables, and maps at different school levels, is, to some extent, a still unexploited area in the research on graphics and an important area for further investigations.

One purpose of the present study was to distinguish among different factors underlying students’ performance on an instrument designed for the measuring of graphicacy in secondary education. An additional aim was to gain a better understanding of the nature of the factors identified. Although not a main focus of the project, gender differences, if any, were also searched for.

Data Collection and Analysis

Construction and Testing of the Instrument

With the purpose of obtaining an overview of what kind of graphic displays the students had met, a number of textbooks in different school subjects were scrutinized. Thereafter, 21 graphic displays were designed covering a rather wide variety of formats, such as line graphs, bar charts, scatterplots, and cartograms. Some of these data thus represented were derived from two collections of facts for school use (Eriksson, 1998; Köhler, 1994) and concerned, for example, speediness of animals, rainfall in different parts of Africa, and the fluctuation of temperature in northern and southern parts of the continent. Other data were invented, such as those about life and circumstances in an imaginary school class. A graph, chart, or map was presented at the top of every page (see Table 1 for an overview of how the structures were designed), followed by two to eight tasks to be solved (see Appendixes A and B for examples). For the questions a number of different formats were used. Ordinary multiple-choice questions were mixed with tasks in which the students were supposed to read off a value and write it down. In some cases, they should just tick off “right” or “wrong.” Such questions were grouped by fours and treated as indices in the analysis. In a small number of more open-ended questions (seven in the main study; see Appendix B for an example), an account of a process or an explanation of a pattern was asked for. (Two of these tasks, 6b and 6c, were jointly scored and two were not scored at all as they were not meant to be included in the present analysis.) Three tasks involved coordinates on a Cartesian grid.

The questions were tested in two pilot studies, the first with six students from grades 5, 7, and 9 (i.e., 12, 14, and 16 years of age, respectively). Thereafter, the instrument was revised and the decision taken to carry out the study merely with the oldest group, as we had clear indications that it would be extremely difficult to construct an instrument yielding sufficient dispersion among both fifth and ninth graders. A second pilot study was carried out with approximately 150 students from two schools. A few questions were excluded because of, for example, too small a variance or for being frequently misunderstood, and some smaller adjustments were made to some of the remaining tasks. The final version consisted of 18 pages, and thus 18 graphic displays. A testing time of 90 minutes in the pilot studies was shortened to 70 minutes in the main study. This was mainly due to the fact that the final version included fewer pages in addition to us having noticed

Table 1
An overview of the graphicacy test

Page	Theme	Type of Graphic	Question Formats ^a (Task No)
1	The 9-graders' Aquarium	Pie chart	shoe (1a, 1c); mc (1b, 1d)
2	A Junior Football Team	Clustered bar graph	shoe (2a, 2b, 2d, 2e); mc (2c)
3	The Price of "The Rag"	Pie chart	shoe (3a, 3b, 3c)
4	The Profit of "The Rag"	Bar graph	shoe (4a, 4b, 4c); mc (4d)
5	Sports in Grade 5	Bar graph	shoe (5a, 5c, 5e); narr (5b, 5e)
6	Elin's Bike Ride	Line graph	mc (6a); narr (6b, 6c)
7	Trips by Car and Motorbike	Multiple line graph	shoe (6a, 6b), mc (6g) rw (6c–d)
8	On Paper Production	Two pie charts	shoe (8f); mc (8e), rw (8a–d)
9	A Lot of Trees	Scatterplot	narr (9a, 9b)
10	How Fast are Animals?	Bar graph (horizontal)	shoe (10a, 10b, 10e); mc (10c); rw (10e–h)
11	The Population of Africa 1	Cartogram	mc (11a, 11b)
12	The Population of Africa 2	Pie chart	shoe (12f); mc (12e); rw (12a–d)
13	Rainfall in Africa	Cartogram	shoe (13a, 13b); rw (13c–f)
14	Temperatures in Africa	Multiple line graph	shoe (14a, 14b, 14c); rw (14d–g); narr (14h)
15	A Party Game	Coordinate grid	shoe (15a, 15b, 15c, 15d)
16	Large Lakes in the World	Table	shoe (16a, 16b, 16d); mc (16c)
17	Large Lakes in the World	Coordinate grid	shoe (17e, 17f)
18	Natural Resources in Africa	Cartogram	shoe (18b); mc (18a); rw (18c–f)

^ashoe, short open-ended question (e.g., the writing down a value); mc, multiple choice; rw, a "right/wrong" cluster forming an index; narr, an open-ended question demanding a longer answer.

that 90 minutes was too long a session for a relatively large number of students to maintain concentration. The students, both in the pilot studies and in the main study, took the test during regular school hours, although not necessarily in math or science class. In the very first tryouts, however, one, two, or three students were solving the pilot tasks in a group room outside the ordinary classroom, and in one case two girls solved the tasks in a home setting. At the testing sessions, at least one of the researchers and (in most cases) one of the ordinary teachers were present.

When used in the analysis, most of the answers were assessed on a scale ranging from 0 to 6. The number of steps used between these extremes was, however, dependent on the characteristics of the question posed. A small number of questions were dichotomously scored (i.e., ascribed either 0 point if wrong or 6 points if correct). Four open-ended questions of a narrative kind were both quantitatively graded on the scale presented and qualitatively categorized. The indexes comprised four "right or wrong" questions, each scored 0 or 2, giving a possible maximal of 8 points. The estimated reliability for the measurement comprising 60 items used in the subsequent analyses yielded a value of 0.93 for Cronbach's alpha.

Schools and Students in the Main Study

The main study was carried out at five schools (some urban, some rural) situated in western Sweden. Of the 494 students in grade 9 in these schools, 363 (participation was voluntary) took the test. Because some of the test-takers did not write their names on the booklet a coupling to formal academic achievement was possible for only for 355 students, which comprises the number of subjects for most of the present analyses. A comparison of grades from the leaving certificate from grade 9 showed that the students who actually took the test had

significantly better grades, on average, than those who did not. However, using the facilities for modeling on incomplete data built into the software program used (Gustafsson & Stahl, 2000) and presented later in this study, no differences in patterns of relationships between manifest and latent variables or other important discrepancies between models with 355, 363, or 494 students were indicated.

Students' Grades Used in the Analysis

The students' final grades awarded at the end of Year 9 (i.e., leaving certificate from compulsory school) were collected at the end of the school year and used in the analysis. The grades included in the Swedish grading system, namely *Pass*, *Pass with Distinction*, and *Pass with Special Distinction* (Skolverket, 2003), were given values of 10, 15, and 20, respectively. The value 0 was assigned in cases where students had failed to achieve a passing grade in a subject. (This is a system of evaluating the grades often used in Swedish educational practice, such as when ordering students in preference for admission to upper secondary education.)

Although the majority of students in the study were awarded separate grades for each subject by their teachers, one small group of 16 individuals was given overall grades for the subject block of science (biology, physics, and chemistry). Because this group formed a relatively small portion of the entire number of students in the study, the grades missing for each of these students were replaced with their block grade. After this, replacement grades from 16 different subjects were available for analysis (for an overview, see Fig. 2 in the Results section).

Strategies for Analyses

The analyses were carried out with confirmatory factor analysis (CFA) using STREAMS (Gustafsson & Stahl, 2000), which is a front-end system for model-testing programs such as LISREL (Jöreskog & Sörbom, 1996) and Amos (Arbuckle, 1997). In CFA, a hypothesized model subjected to a structural equation modeling (or SEM) technique is statistically tested against data in that the model-implied covariance matrix is re-created from the relations estimated; for example, a maximum-likelihood procedure is compared with an empirical covariance matrix. A good fit is seen when the difference is low and statistically insignificant. A number of fit indexes, each with its particular characteristics, may be obtained from the estimation program used. In the present analyses, the root-mean-square error of approximation (RMSEA) index, which takes into account the complexity of the model and sample size, was used to judge how well the models fitted the data. Usually, a value of 0.05 or lower is considered to indicate a good fit (Browne & Cudeck, 1993), but when taking a confidence interval into account the limit for an acceptable fit may be extended to around 0.07. The traditional χ^2 measure is also provided with the results in addition to the degrees of freedom for the model. As the χ^2 value is, among other things, sensitive to sample size, insignificant values seldom occur with larger samples. Sometimes, as a rule of thumb, a χ^2/df ratio of 2.5 times the degrees of freedom or less is considered an acceptable fit (e.g., Stage, 1990).

Following a well-established tradition of cognitive achievement being hierarchically organized in broad and narrow factors (see, e.g., Carroll, 1993; Gustafsson, 1988a, 1988b; Horn & Cattell, 1966), a general factor as well as intermediate and more specific dimensions were hypothesized for the test of graphicacy as well as for the grades from the leaving certificate, and nested factor (NF) modeling was used. In an NF model, less general and narrower factors are nested under broader and more general ones (Gustafsson & Balke, 1993; Gustafsson & Undheim,

1996). Such models, which are assumed to be orthogonal, can be reformulated from more traditional higher-order models, as shown by Gustafsson and Balke (1993). They are, however, often easier to formulate and interpret as the complexity of a manifest variable is directly observable when contemporaneously related to several factors (Figs. 1 and 2 show NF models).

Results

Initially, a nested factor (NF) model was identified for the test of graphicacy and, thereafter, another NF model was identified for the leaving certificate from compulsory school. Next, these two measurement models were combined into one structural model whereby the latent variables in

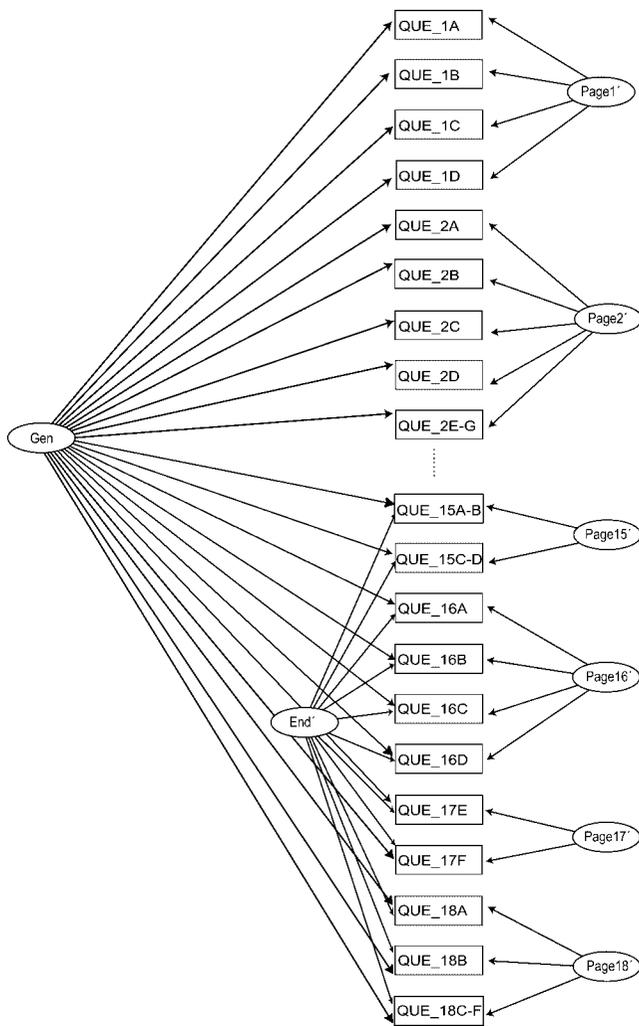


Figure 1. The measurement model for the graphicacy test with the general-factor, end-of-test effect, and the page-specific factors. (The narrative factor is not included.) For simplification, the residuals related to the dependent variables are omitted from the diagram.

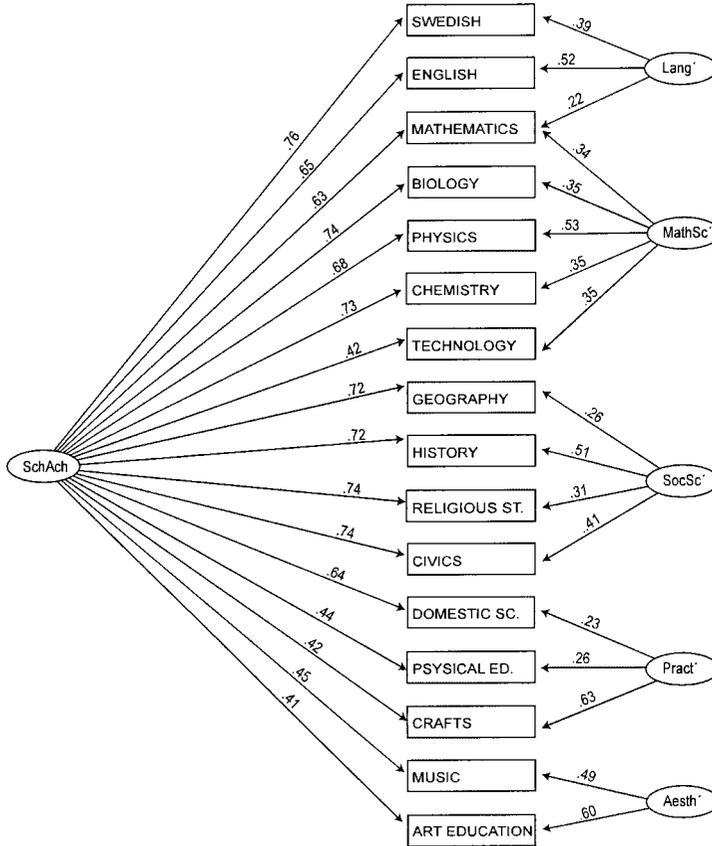


Figure 2. The grades model. Relationships and standardized factor loadings (residuals omitted).

the previous models were stipulated to co-vary between the models to a greater or lesser extent. The presentation of the results is sequenced accordingly in the following subsections and concluded with a brief summary.

Identification of a Measurement Model for the Graphicacy Test

In a first step all items (manifest variables) were regressed on a single latent variable (a “general” test factor). The factor loadings thus obtained were statistically significant and, in most cases, quite substantial. Although this “general” factor accounted for almost 93% of the total variance, the amounts of explained variance in the single items were in many cases quite low or, in some cases, very low, particularly for items located early in the test, whereas a more substantial part of the variance was explained for items toward the end of the test. Furthermore, the fit of the model was rather poor, $\chi^2(1710) = 4082.1$, RMSEA = 0.081. Consequently, a more elaborate model was built.

As we had observed, although the students were taking the test, many students did not solve the whole body of tasks making the hypothesizing of an end-of-test effect a possible next step. However, strong indications of correlations between items on the same page of the booklet, and thus related to the same graphic display, were obtained from the LISREL and

Amos outputs. Therefore, 18 specific factors related to each of the 18 graphic displays were stipulated and brought into the model. To keep the number of parameters reasonably low for the handling of the joint model to be tested later, two factors with insignificant relationships were removed. The fit was now considerably better, $\chi^2(1661) = 2392.3$, $RMSEA = 0.040$. In addition, the amount of explained variance showed a substantial increase for most of the items.

Next, the hypothesized “end” factor was introduced into the model as a latent variable at an intermediate level—that is, it was nested under the general factor. A successive inclusion of items starting from the end of the test showed that items from page 12 onward significantly loaded on this factor. Thereafter, a large number of hypothesized factors were tested, one after the other. However, with one exception, no further dimensions influencing performance were found—neither with respect to type of graphic display (such as pie charts, line graphs, scatterplots, etc.) nor with respect to the format of questions (e.g., multiple choice or simple reading off of value). Although at one stage of the model testing there were some vague indications of a quantitative factor related to tasks involving comparisons of proportions and other calculations, no such factor could be properly identified.

There was one case where an additional factor was (vaguely) pointed out. The estimation program successfully converged on a solution when a narrative factor ($Narr'$) was postulated. Four items, which involved the writing of longer answers to open-ended questions, loaded on this factor. It should be mentioned, however, that only two of the relationships were statistically significant and that two correlations were relatively weak, which are discussed further in the final section. Figure 1 shows the factor structure of this model. The fit indexes were good, namely $\chi^2(1636) = 2009.40$, $RMSEA = 0.023$. The identified factors accounted for 95% of the total variance. The estimated factor loadings and amount of variance for each item explained by the factors are given in Table 2.

Finally, the gender of the test-takers was introduced into the model as an independent manifest variable. Regressing the Gen , End' , and $Narr'$ factors on “gender” yielded no differences between boys and girls on either of these factors.

A Model for Grades From the Leaving Certificate

Previous research has maintained that grades from the leaving certificate in compulsory school are multidimensional. Gustafsson and Balke (1993) fitted a hierarchical model, including a general factor labeled “school achievement,” as well as orthogonal domain-specific factors for areas as practical studies (such as child studies, art education, and domestic science), science, social sciences, and language. A spatial/practical factor (related to art education, technology, and crafts) was also identified. In addition to the general dimension, Andersson (1998) argued for two domain-specific achievement factors: mathematics/science and language, respectively, as well as for two broad factors, one of them nonverbal and the other aesthetic/domestic. Gustafsson and Balke’s as well as Andersson’s studies refer to the previous national norm-referenced 5-point grading system used in Sweden from the beginning of the 1960s until the end of the 1990s, when this system was replaced with the goal- and criterion-related grading system presented earlier. Some changes with respect to subject matter were also made. Thus, our modeling on the leaving certificate for the group of test-takers applies to a somewhat different grading system.

In the hypothesizing and testing of models we drew heavily on the results referred to previously (Andersson, 1998; Gustafsson & Balke, 1993). In the same vein as in those studies, an overall school achievement dimension ($SchAch$) related to all subjects was first introduced.

Table 2

Graphicacy test model: standardized factor loadings and proportion of explained variance for test items

Manifest Variable (Item) ^a	<i>Gen</i>	<i>End'</i>	Page Specific Factor (Page1', etc.)	<i>Narr'</i>	Expl. var. (%)
QUE_01A	0.41		0.61		54.35
QUE_01B	0.40		0.40		31.89
QUE_01C	0.43		0.51		44.13
QUE_01D	0.43		0.52		45.81
QUE_02A	0.21		0.29		12.95
QUE_02B	0.37		0.27		20.78
QUE_02C	0.35		0.16		15.27
QUE_02D	0.21		0.14		6.28
QUE_02E-G	0.34		0.53		39.83
QUE_03A	0.38		0.55		43.87
QUE_03B	0.27		0.57		39.47
QUE_03C	0.25		0.79		69.27
QUE_03D	0.25		0.62		44.85
QUE_04A	0.24		0.11		6.86
QUE_04B	0.30		0.45		29.33
QUE_04C	0.29		0.58		41.55
QUE_04D	0.25		0.26		12.71
QUE_05A	0.27		0.29		15.73
QUE_05C	0.25		0.37		20.00
QUE_05D	0.19		0.40		19.57
QUE_06A	0.44				19.36
QUE_06B	0.45			0.14	40.02
QUE_07A	0.49		0.27		31.17
QUE_07B	0.42		0.74		72.57
QUE_07C-F	0.46		0.09		22.57
QUE_07G	0.37		0.17		16.54
QUE_08A-D	0.59				34.22
QUE_08E	0.41				16.32
QUE_08F	0.52				26.69
QUE_09A	0.61		0.51	0.49	64.14
QUE_09B	0.59		0.43	0.45	53.09
QUE_10A	0.38		0.17		17.57
QUE_10B	0.42		0.34		28.80
QUE_10C	0.34		0.19		15.16
QUE_10D	0.49		0.56		54.25
QUE_10E-H	0.49		0.28		31.76
QUE_11A	0.31		0.53		37.54
QUE_11B	0.29		0.53		36.25
QUE_12A-D	0.55	0.18	0.48		57.02
QUE_12E	0.39	0.19	0.44		38.46
QUE_12F	0.32	0.15	0.25		18.95
QUE_13A	0.61	0.18	0.41		57.95
QUE_13B	0.33	0.17	0.31		23.42
QUE_13C-F	0.51	0.37	0.20		43.16
QUE_14A	0.39	0.35	0.24		33.22
QUE_14B	0.41	0.30	0.09		27.10
QUE_14C	0.32	0.15	0.27		20.00
QUE_14E-G	0.42	0.45	0.57		70.54
QUE_14H	0.43	0.54	0.24	0.10	70.15
QUE_15A-B	0.44	0.55	0.37		67.45
QUE_15C-D	0.49	0.48	0.43		76.28

(Continued)

Table 2
(Continued)

Manifest Variable (Item) ^a	<i>Gen</i>	<i>End'</i>	Page Specific Factor (Page1', etc.)	<i>Narr'</i>	Expl. var. (%)
QUE_16A	0.38	0.63	0.37		58.02
QUE_16B	0.40	0.61	0.48		31.26
QUE_16C	0.34	0.45	0.51		56.41
QUE_16D	0.30	0.42	0.22		58.68
QUE_17E	0.35	0.49	0.45		64.72
QUE_17F	0.34	0.52	0.45		71.02
QUE_18A	0.22	0.55	0.55		73.50
QUE_18B	0.33	0.54	0.56		63.50
QUE_18C-F	0.33	0.63	0.48		65.23

^aThe page number is indicated by the figures after the underscore in the names of the manifest variables, whereas the letters (A, B, etc.) after the figures refer to the different questions on each page.

Thereafter, five orthogonal factors representing domain-specific areas were added to the model (Fig. 2), which bears great similarity to those fitted by Gustafsson and Balke (1993). There are, however, some crucial differences. In our model, mathematics and science were related to one joint factor (*MathSc'*) and the spatial/practical factor was replaced by a narrow aesthetic factor. Furthermore, mathematics has a rather substantial and statistically significant loading on the language factor, a phenomenon observed by Andersson (1998), and also Åberg-Bengtsson and Erickson (2004), who investigated performance on National Tests for grade 9. As shown in Figure 2, both the factorial structure and the pattern of factor loadings make sense with respect to each factor. Subjects traditionally regarded as more heavily drawing on academic performance, such as mathematics, languages, and the natural and social sciences, loaded more strongly on the *SchAch* factor than the practical and aesthetic subjects, to give just one example. The model differs from Anderson's model first and foremost with respect to the absence of a nonverbal factor maintained by Andersson (1998). The fit was good, $\chi^2(87) = 155.8$, RMSEA = 0.046. The six factors accounted for 94% of the total variance. Figure 2 gives an overview of the identified model as well as the estimated standardized factor loadings.

The Joint Model

With the aim of gaining a better understanding of the nature of the general dimension of performance identified for the graphicacy test, the two measurement models (presented in Figs. 1 and 2, respectively) were brought together into one structural model. However, although the existence of a narrative factor was indicated when modeling on the graphicacy test, it was not regarded as substantial enough to be used in further modeling. Thus, only the *Gen* and *End'* factors of the test were allowed to co-vary with the factors of the grades model. The fit indexes obtained were $\chi^2(674) = 3269.10$, RMSEA = 0.021. Table 3 shows the relationships between the latent variables in the joint model and the strength of the correlations.

As seen in Table 3, the "general" dimension for the graphicacy test (*Gen*) proved to be significantly related to all grade-model factors, except for practical and aesthetic factors. The general dimension (*Gen*) correlated most strongly with the mathematics/science (*MathSc'*) factor. In addition, the overall school achievement factor (*SchAch'*), as well as the language factor (*Lang'*), were relatively strongly related to this general test dimension, whereas the correlation

Table 3
Correlations between latent variables in the combined model

	<i>GenAch</i>	<i>MathSc'</i>	<i>Lang'</i>	<i>SocSc'</i>	<i>Pract'</i>	<i>Aesth'</i>
<i>Gen</i>	0.48**	0.58**	0.47**	0.23**	0.07	0.08
<i>End'</i>	0.05	-0.16*	0.07	-0.01	-0.07	-0.33**

* $0.10 > p > 0.05$; ** $p < 0.01$.

with the social science (*SocSc'*) factor showed a somewhat weaker relation. The end-of-test effect (*End'*) had its only significant correlation with the aesthetic factor, a relation that was negative.

Summary

For the graphicacy test, a general factor related to all items was successfully identified in addition to an end-of-test effect. In spite of positing and testing a considerable number of postulated factors, which could be assumed to underlie performance, no such factors were found. Items (in most cases, one question, or, alternatively, four “right or wrong” questions forming an index) on a particular page related to one graphic display were, as a rule, loaded on a particular “page factor.” There were some indications of a “narrative” factor of some importance to test performance. The attempt to identify such a factor (*Narr'*) from the four narrative questions was successful only to some extent. No gender differences were found, neither in the overall performance of the test (*Gen*) nor in the end-of-test factor (*End'*) or hypothesized narrative factor (*Narr'*).

The *Gen* and *End'* factors were allowed to co-vary with the factors of a model identified for the students' grades from the leaving certificate from compulsory school. The strongest (positive) correlation was obtained between the general graphicacy-test factor (*Gen*) and the mathematics–science factor (*MathSc'*), followed by correlations between the general test factor, on the one hand, and the general (academic) achievement and the language factors, on the other.

Discussion

The purpose of the study was to investigate factors that may underlie secondary students' performance on a test of graphicacy and to gain an understanding of the factors identified. Thereby, confirmatory factor analyses were carried out on the graphicacy test as well as on the leaving certificate from compulsory schooling. Thereafter, the factors identified in these analyses were related.

Having postulated and successfully tested an overall graphicacy-test factor (*Gen*), in addition to an end-of-test effect, no broad dimensions of performance could be identified. However, a “narrative” factor was vaguely indicated. An initial question would be: What interpretation could be given this overall factor? First, it seems reasonable to assume a more general ability as being involved in solving the diagrammatic tasks, which is indicated by the substantial correlation ($r = 0.48$) with the general school-achievement factor (*SchAch*). Second, the overall graphicacy-test factor (*Gen*) had its highest correlation (0.58) with the mathematics–science factor (*MathSc'*) of the grades model. This, of course, could have several explanations. Certainly, although applied in a range of school subjects and everyday situations, the gathering of numerical information

presented in diagrammatic displays belongs, to a great extent, to the realm of mathematical discourse and reasoning. Thus, a specific *MathSc'* ability factor may be influencing the overall test performance, which certainly makes sense. In addition, as the test situation was presumably bearing such mathematical connotations as mentioned earlier, there might be an additional effect from the students' expectations. Stated differently, students may have regarded the test as a mathematical test and thus approached the test situation accordingly. However, alternative or complementary explanations could be considered. Although not properly identified, a "quantitative" ability might still be involved, particularly in tasks that demanded calculations in addition to the reading off of values or estimation of proportions. Such a "hidden" quantitative factor nested within the overall dimension would certainly, if substantial, render high correlations with the *MathSc'* factor.

As noted earlier, the four narrative items loaded on a language factor (*Narr'*). It may be assumed that a verbal or linguistic dimension could be involved in performance on the graphicacy test. Quite obviously, a verbal ability is demanded when reading introductions and texts presenting the subjects of the pages as well as tasks to be carried out. Thereby, the "narrative" items may be thought of either as related to a (nested) subfactor on their own or intertwined with other verbal or linguistic aspects. Such an intertwining makes sense with respect to the difficulties of a stable identification of the narrative factor, which could thus be tentatively explained.

An alternative explanation, however, as to the absence of clearly identified, additional broad factors (except from the end-of-test effect) may be that performance of diagrammatic tasks is, on the whole, unidimensional, or involves merely a few dimensions, a result in line with previous research on the Swedish Scholastic Aptitude Test (Åberg-Bengtsson, 1999). Such an interpretation can also be compared with the arguments of Winn and Holliday (1982) that benefiting from diagrammatic representations goes beyond the verbal and visual and relies mainly on a more general ability or "intelligence." However, the fact that no spatial or visual factor was identified in the present study could not, and should not, be interpreted as the absence of such a dimension. The complexity of the test may very well obscure such results. Moreover, the grades model onto which the graphicacy test was related had no adequate indicator of spatial abilities.

The results also point out substantial relationships between the majority of graphicacy-test items and a page-specific factor. The relationships proved very persistent during the modeling processes and may be the reason why hypothesized factors were difficult to identify. Furthermore, identification of these page-specific factors should be considered in light of research referred to in the introductory sections. Roth (2003) argued that familiarity with the content of the graph and with conventions in particular domains was important for graphical competencies and "such competencies do not easily transfer" (p. 104). In addition, Winn (1993) argued that previous knowledge of the content of a diagram and knowledge of the symbol system used in the display were the most influential factors when searching for information in diagrammatic representations. The consistent occurrence of page-specific factors does not seem to be an artifact of the test construction as we initially feared, but rather a result in accordance with the suggestions made by Roth (2003) and Winn (1993). In fact, such an explanation seems most plausible considering the fact that each page of the graphicacy test had its own theme and its own graphic display. In addition, even though it has particular aspects in common with other displays of the test, each graphic representation was to a certain degree unique for the page and theme in question and, consequently, so were the conventions and symbols systems. Of course, alternative and/or complementary explanations can be suggested. For example, it seems plausible that having once "invested" in sense-making of the

information and the illustration on a particular page the students continued with the entire set of related tasks.

Some Reflections on Educational Implications

Identification of strong page (i.e., content-related) factors as well as the indicated characteristics of the overall graphical test factor certainly has some educational implications. On the one hand, it may be argued that, in authentic educational situations, graphical illustrations are often followed-up by or following up more or less extensive verbal explanations of one kind or another. Thus, text and illustration together provide a whole, where both parts are intertwined in the building of an understanding of the information given. On the other hand, and in line with a number of other researchers (e.g., Ainley, 2000; Roth, 2003), we argue that the importance of interest and familiarity with the content domain of the graph interpreted cannot be overemphasized. Quite obviously, a certain basic knowledge of the graph in question and its conventions is a prerequisite for understanding the provided information at all—a statement supported by the strong relationship identified for the overall graphical test factor (*Gen*) with the grades model *MathSc'* factor. However, it may, for example, *not* be assumed that having learned the handling a particular type of graph in math class provides sufficient background knowledge for an intended reading of the same type of graph in a new context (e.g., a particular situation in science class). Consequently, in addition to teachers making sure that their students have a basic knowing of the graph or chart *per se*, it seems to be of utmost importance that the “newness” of every graphic illustration should be attended to so that the graph may be “worth a thousand words” and, hopefully, serve as a structuring resource (Wenger, 1998) for learning and understanding.

Future Analyses

The results presented herein maintain and support the significance of the content dimension as to how graphical tasks are handled and how graphs are understood. Thus, it is important to further pursue analyses of the open-ended questions. Therefore, a complementary analysis has been carried out on the “narrative” tasks of the graphicacy test. Using this test, the students’ written answers were categorized with respect to qualitatively different ways of making sense of the information illustrated (Åberg-Bengtsson & Ottosson, 2004). In addition, there was supplementary data collected in the pilot study for the investigation presented in this study, where 152 students were asked to draw 13 graphs and charts from sets of given data. This material will be subjected to qualitative analyses. Although the main investigation of these data remains to be done, initial scrutiny suggests some noteworthy potential outcomes. However, although the present project will hopefully contribute to the understanding of students’ grappling with information presented graphically, there are a number of questions left unanswered and worthy of future study. The further pursuit of a spatial or visual factor is one such example. To identify such a dimension, if any, the complexity of the instrument should presumably be reduced as much as possible with a minimum of question-and-answer formats in addition to well-distinguished graphic displays.

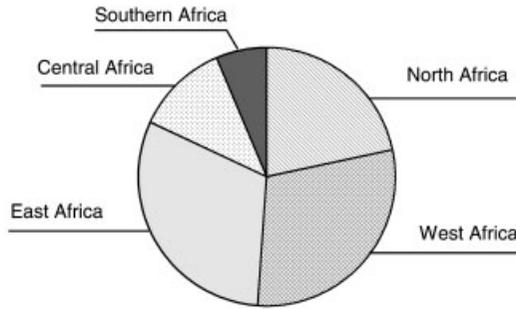
The authors thank Jan-Eric Gustafsson for help and support during the research process, Bob Kaill for helping to prepare this article for publication, and two anonymous reviewers for valuable comments on the manuscript.

Appendix A

English translation of page 12 with its theme and examples of right/wrong, multiple-choice, and short open-ended questions (cf. Table 1).

The Population of Africa 2

Almost 740 million people live in Africa. They make up the whole circle or 100 percent (100%). The pie chart below shows how this population is distributed over different parts of the continent.



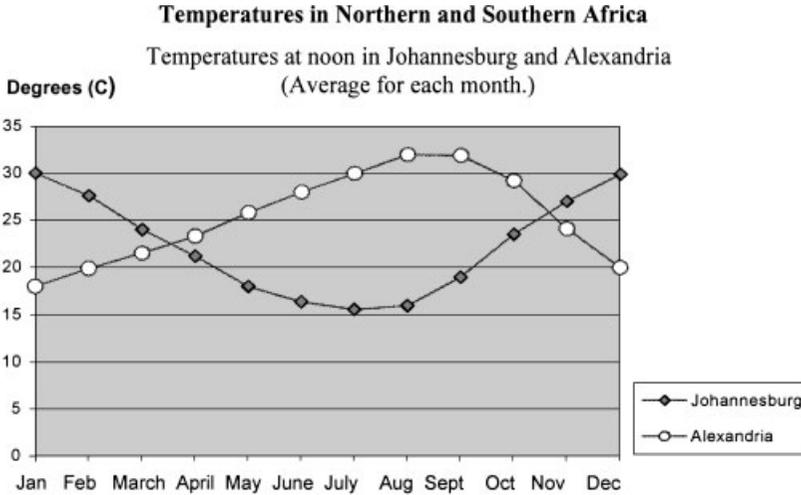
Mark right or wrong with an X for the following Questions:

Right Wrong

- (a) About one third of the population of Africa lives in East Africa. Right Wrong
- (b) More than one fourth of the population of Africa lives in West Africa. Right Wrong
- (c) Twice as many people live in North Africa as in Central Africa and Southern Africa together. Right Wrong
- (d) If all people in Southern Africa moved to Central Africa the population of Central Africa would increase by 50%. Right Wrong
- (e) In the pie chart above you can see that approximately 6% of the population of Africa live in Southern Africa. But how many percent live in Central Africa?
 - About 3%
 - About 9%
 - About 12%
 - About 18%
- (f) According to the pie chart, how large a part of the population of Africa lives in Northern Africa?

Appendix B

English translation of page 14 with its theme and examples of short open-ended, right/wrong, and narrative questions (cf. Table 1).



(All questions refer to temperature at noon, i.e., what the graph shows.)

- (a) How many degrees is it in August in Johannesburg? _____
- (b) How much colder is it in Johannesburg in July compared to Alexandria during the same month?

- (c) How much warmer is it in February than in September in Johannesburg? _____

Mark right or wrong with an X for the following Questions:

- | | Right | Wrong |
|---------------------------------------------------------------------------------------------|--------------------------|--------------------------|
| (d) In Alexandria it is warmest in July-October. | <input type="checkbox"/> | <input type="checkbox"/> |
| (e) In November it is 6° warmer in Johannesburg than in Alexandria | <input type="checkbox"/> | <input type="checkbox"/> |
| (f) It is roughly equally warm in November in Johannesburg as it is in March in Alexandria. | <input type="checkbox"/> | <input type="checkbox"/> |
| (g) From April to June the temperature rises approximately 5° in Alexandria. | <input type="checkbox"/> | <input type="checkbox"/> |

- (h) Relate how the temperature varies during the year in Johannesburg? _____

References

Åberg-Bengtsson, L. (1998). Entering a graphicacy society: Young children learning graphs and charts (Göteborg Studies in Educational Sciences, 127). Göteborg, Sweden: Acta Universitatis Gothoburgensis.

Åberg-Bengtsson, L. (1999). Dimensions of performance in the interpretation of diagrams, tables, and maps: Some gender differences in the Swedish Scholastic Aptitude Test. *Journal of Research in Science Teaching*, 36, 565–582.

Åberg-Bengtsson, L. (in press). Separating quantitative and analytic dimensions in the Swedish Scholastic Aptitude Test. *Scandinavian Journal of Educational Research*

Åberg-Bengtsson, L. & Erickson, G. (2004). Dimensions of national test performance in language and mathematics: A two-level approach. Manuscript submitted.

Åberg-Bengtsson, L. & Ottosson, T. (2004). Students' ways of making sense of line graphs and scatter plots in open-ended questions. Manuscript submitted.

Ainley, J. (2000). Transparency in graphs and graphing tasks: An iterative design process. *Journal of Mathematical Behavior*, 19, 365–384.

Andersson, A. (1998). The dimensionality of the leaving certificate in the Swedish compulsory school. *Scandinavian Journal of Educational Research*, 42, 25–40.

Arbuckle, J.L. (1997). *Amos users' guide*, version 3.6. Chicago, IL: Small Waters Corporation.

Beniger, J.R. & Robyn, D.L. (1978). Quantitative graphics in statistics: A brief history. *The American Statistician*, 32, 1–11.

Berg, C.A. & Smith, P. (1994). Assessing students' abilities to construct and interpret line graphs: Disparities between multiple choice and free response instruments. *Science Education*, 78, 527–554.

Bowen, G.M., Roth, W.-M., & McGinn, M.K. (1999). Interpretation of graphs by university biology students and practicing scientists: Toward a social practice view of scientific practices. *Journal of Research in Science Teaching*, 36, 1020–1043.

Browne, M.W. & Cudeck, R. (1993). Alternative ways of assessing model fit. In K.A. Bollen & J.S. Long (Eds.), *Testing structural equation models* (pp. 136–162). Thousand Oaks, CA: Sage.

Carroll, J.B. (1993). *Human cognitive abilities: A survey of factor analytic studies*. New York: Cambridge University Press.

Clement, J. (1989). The concept of variation and misconceptions in Cartesian graphing. *Focus on Learning Problems in Mathematics*, 11, 77–87.

Eriksson, L. (1998). *Världen i siffror: Upplaga 98 [The world in numbers: Year 1998.]* (6th ed.) Stockholm: Natur och Kultur.

Goldberg, F.M. & Anderson, J.H. (1989). Student difficulties with graphical representations of negative values of velocity. *Physics Teacher*, 27, 254–260.

Guthrie, J.T., Weber, S., & Kimmerly, N. (1993). Searching documents: Cognitive processes and deficits in understanding graphs, tables, and illustrations. *Contemporary Educational Psychology*, 18, 186–221.

Gustafsson, J.-E. (1988a). Hierarchical models of individual differences and cognitive abilities. In R.J. Sternberg (Ed.), *Advances in the psychology of human intelligence* (vol. 4, pp. 35–71). Hillsdale, NJ: Erlbaum.

Gustafsson, J.-E. (1988b). Models of intelligence. In P.J. Keeves (Ed.), *Educational research, methodology, and measurement: An international handbook* (pp. 437–441). Oxford: Pergamon Press.

Gustafsson, J.-E. & Balke, G. (1993). General and specific abilities as predictors of school achievement. *Multivariate Behavioral Research*, 28, 407–434.

Gustafsson, J.-E. & Stahl, P.A. (2000). *STREAMS user's guide: Version 2.5 for Windows*. Mölndal, Sweden: MultivariateWare.

Gustafsson, J.-E. & Undheim, J.O. (1996). Individual differences in cognitive functions. In D.C. Berliner & R.C. Calfee (Eds.), *Handbook of educational psychology* (pp. 186–242). New York: Macmillan.

Gustafsson, J.-E. Wedman, I., & Westerlund, A. (1992). The dimensionality of the Swedish Scholastic Aptitude Test. *Scandinavian Journal of Educational Research*, 36, 21–39.

Halpern, D.K. (1992). *Sex differences in cognitive abilities* (2nd ed.) Hillsdale, NJ: Lawrence Erlbaum.

Horn, J.L. & Cattell, R.B. (1966). Refinement and test of the theory of fluid and crystallized general intelligences. *Journal of Educational Psychology*, 57, 253–270.

Jones, G.A., Thornton, C.A., Langrall, C.W., Mooney, E.S., Wares, A., Perry, B., & Putt, I. (2000). A framework for characterizing children's statistical thinking. *Mathematical Thinking and Learning*, 2, 269–307.

Jöreskog, K.G. & Sörbom, D. (1996). *LISREL 8 user's reference guide* (2nd ed.) Chicago: Scientific Software International.

Kealy, W.A. & Webb, J.M. (1995). Contextual influences of maps and diagrams on learning. *Contemporary Educational Psychology*, 20, 340–358.

Kerslake, D. (1981). Graphs. In K. Hart (Ed.), *Children's understanding of mathematics: 1–16* (pp. 120–136). London: John Murray.

Köhler, P.O. (1994). *Skolans tabeller: Tabeller och diagram för grundskolans orienteringsämnen* [School statistics: Tables and diagrams for social subjects in primary school] (2nd ed.) Stockholm: Natur och Kultur.

Kozhevnikov, M., Hegarty, M., & Mayer, R.E. (2002). Revising the visualizer–verbalizer dimension: Evidence for two types of visualizers. *Cognition and Instruction*, 20, 47–77.

Leinhardt, G., Zaslavsky, O., & Stein, M.K. (1990). Functions, graphs, and graphing: Tasks, learning, and teaching. *Review of Educational Research*, 60, 1–64.

Lewandowsky, S. & Spence, I. (1989). The perception of statistical graphs. *Sociological Methods and Research*, 18, 200–242.

Lindwall, O. (1998). *Samarbete och problemlösande i mikrodatorbaserade laborationer* [Collaboration and problem solving in micro computer based laboratory lessons] (Magisteruppsatser från Tema K, 1998:2). Linköping: Linköping University, Tema K.

Lindwall, O. & Ivarsson, J. (2004). What makes the subject matter matter? Contrasting probeware with graphs & tracks. Manuscript submitted.

Linn, M.C., Layman, J.W., & Nachmias, R. (1987). Cognitive consequences of microcomputer-based laboratories: Graphing skills development. *Contemporary Educational Psychology*, 12, 244–253.

Lowe, R.K. (2003). Animation and learning: Selective processing of information in dynamic graphics. *Learning and Instruction*, 13, 157–176.

Nemirovsky, R. & Noble, T. (1997). On mathematical visualization and the place where we live. *Educational Studies in Mathematics*, 33, 99–131.

Nemirovsky, R. & Tierney, C. (2001). Children creating ways to represent changing situations: On the development of homogenous space. *Educational Studies in Mathematics*, 45, 67–102.

Nemirovsky, R., Tierney, C., & Wright, T. (1998). Body motion and graphing. *Cognition and Instruction*, 16, 119–172.

Ottosson, T. (1987). *Map-reading and wayfinding* (Göteborg Studies in Educational Sciences, 65). Göteborg, Sweden: Acta Universitatis Gothoburgensis.

Ottosson, T. (1988). What does it take to read a map? *Cartographica*, 25, 28–35.

Ottosson, T. & Åberg-Bengtsson, L. (1995, August). Children's understanding of graphically represented quantitative information. Paper presented at the sixth EARLI conference, Nijmegen, The Netherlands.

Preece, J. (1983). Graphs are not straightforward. In T.R.G. Green, S.J. Payne, & G.C. van der Veer (Eds.), *The psychology of computer use* (pp. 41–56). London: Academic Press.

Roth, W.-M. (2003). *Toward an anthropology of graphing: Semiotic and activity theoretic perspectives*. Dordrecht, The Netherlands: Kluwer.

Roth, W.-M. & McGinn, M.K. (1998). Inscriptions: Toward a theory of representing as social practice. *Review of Educational Research*, 68, 35–59.

Schnotz, W. & Bannert, M. (2003). Construction and interference in learning from multiple representations. *Learning and Instruction*, 13, 141–156.

Skolverket [the Swedish National Agency for Education]. (2003). *The Swedish school system*. [From the official homepage of the Swedish National Agency for Education.] Available: <http://www.skolverket.se/english/system/index.shtml> (June 2003).

Stage, F.K. (1990). LISREL: An introduction and application in higher education research. In J. Smart (Ed.), *Higher education handbook of theory and research* (vol. 6). New York: Agathon Press.

Tufte, E.R. (1983). *The visual display of quantitative information*. Cheshire, CT: Graphics Press.

Verdi, M.P., Kulhavy, R.W., Stock, W.A., Rittschof, K.A., & Johnson, J.T. (1996). Text learning using scientific diagrams: Implications for classroom use. *Contemporary Educational Psychology*, 21, 487–499.

Veriki, I. (2002). What is the value of graphical displays in learning? *Educational Psychological Review*, 14, 261–312.

Wainer, H. (1980). A test of graphicacy in children. *Applied Psychological Measurement*, 4, 331–340.

Wenger, E. (1998). *Communities of practice: Learning, meaning and identity*. New York: Cambridge University Press.

Winn, W. (1991). Learning from maps and diagrams. *Educational Psychology Review*, 3, 211–247.

Winn, W. (1993). An account of how readers search for information in diagrams. *Contemporary Educational Psychology*, 18, 162–185.

Winn, W. & Holliday, W. (1982). Design principles for diagrams and charts. In P.H. Jonassen (Ed.), *The technology of texts* (vol. 1, pp. 277–299). Englewood Cliffs, NJ: Educational Technology Publications.