# COMPARING NOVICES & EXPERTS IN THEIR EXPLORATION OF DATA IN LINE GRAPHS

Bruce H. Tsuji* and Gitte Lindgaard**
**Department of Psychology, Carleton University, Ottawa, Ontario, Canada*
***Faculty of Design, Swinburne University of Technology, Prahran Victoria 3181 Australia*

**ABSTRACT**

This research compared undergraduate Novices and PhD Experts in psychology and business in their exploration of psychology and business domain graphs. An overall expertise effect in graph explanation was found. Results indicated that Novices paused longer than Experts before beginning their explanations. Qualitative analyses showed that Experts were generally more complete in their explanations, generating more inferences, more quantitative statements, and more conceptual messages. Psychology Experts tended to generate more complete explanations for psychology-domain graphs whereas Business Experts generate less complete explanations for business-domain graphs. The results suggest that Experts have superior strategies to Novices in graph exploration that may be accommodated by the graph comprehension model of Pinker (1990). An implication of these results is that simple instructions may greatly enhance the data literacy of students and might be embodied in data visualization tools for adults and researchers as well.

**KEYWORDS**

Graphs, expertise, experts, novices, data

## 1. INTRODUCTION

Graphs are among the most effective ways for people to understand data (Tufte, 1983). Often the purpose is straightforward communication of data as might be found in school textbooks or newspapers or internet sites (Roth, Bowen, & McGinn, 1999). *Exploration* is a special and very interesting case of graph use (Behrens, 1997), often representing a person's attempt to understand, interpret or communicate data. While a common application of graph exploration is in scientific reasoning, it is difficult to imagine many domains where graphs intended for exploratory purposes are *not* found (e.g., Bertin, 1983; Kosslyn, 2006).

Curiously, unlike many other domains such as chess and physics, (Eriksson, 2005) graph exploration does not appear to demonstrate a consistent difference between experts and novices. This is unfortunate because confronted with a graph that requires people to utilize complex inferential processes, a number of interesting theoretical and practical questions arise: Do experts apply qualitatively different strategies than novices (e.g., Gick & Holyoak, 1983)? Do novices focus on the graph's syntactic structure at the expense of an analysis of the deeper semantic components (e.g., Preece & Janvier, 1993)? Are experts able to recognize patterns in graphs in ways that may be similar to how expert chess players recognize chess positions (e.g., Newell and Simon, 1972)? Can graphical visualization tools be designed to better facilitate novice understanding (e.g. Konold, 2007)? More generally, what differences do experts and novices exhibit in graph exploration?

Relatively few studies have addressed the issue of expertise in graph exploration directly. One instance is the ethnographic research of Roth and Bowen (2003) who examined how domain experts in biology, physics, and forest sciences interpreted familiar and unfamiliar graphs. Roth and Bowen found that experts had significant difficulty interpreting graphs taken from undergraduate textbooks from their respective domains but they had little difficulty with familiar graphs taken from their own personal research.

In a different domain, Trafton et al (2002) described how expert meteorologists create spatial transformations of meteorological data when the information requested of them is not explicitly present. For example, in determining the air pressure over Pittsburgh, Trafton et al.'s eye movement data suggested that participants were identifying nearby isobars, calculating the distance between them, and then using the proportional distance to calculate the atmospheric pressure.

However, neither the Roth and Bowen nor the Trafton et al. studies directly compared experts against novices in their respective domains. Thus, it is difficult to know if the strategies inferred by these authors were attributable to expertise per se or were idiosyncratic to the domains selected.

Freedman and Shah (2002) conducted one of the few studies to explicitly compare domain-specific expert (psychology graduate students) and novice (undergraduate students) graph exploration. Freedman and Shah's domain-specific graphs included graphs on cognitive studies of aging whereas domain non-specific graphs were concerned with non-cognitive aging data. Freedman and Shah reported that novices tended to describe main effects while experts were more likely to describe the underlying mathematical functions in the graph stimuli. However, the domain manipulation had no effect. Freedman and Shah interpreted these results as supporting the notion that novices attend to lower-level perceptual features of a graph whereas experts enrich and elaborate the visual features of a graph with their domain knowledge. However, it is difficult to reconcile Freedman and Shah's results with those of Roth and Bowen (2003). Is expertise in graph exploration a general skill (as suggested by Freedman & Shah) or one that is very specific to a given expert's domain (like those of Roth & Bowen)?

The purpose of the current research is to identify differences (if any) between novices and experts in their exploration of graphs drawn from familiar and unfamiliar domains. If experts are superior to novices regardless of domain, then graph expertise may be a more general ability. The contribution of this research speaks not only to our understanding of expertise but also to the application of data visualization tools and to the education of students from different disciplines in terms of their understanding of data.

Shah and Carpenter (1995) compared psychology graduate- and undergraduate students using graphs from common-knowledge domains. They found no effect of expertise. Using business and psychology domain-specific graphs and PhD faculty Experts compared to undergraduate Novices, the present study was designed as a more sensitive test of graph expertise, leading to Hypothesis 1: Experts would generate more causal inferences about graphs than Novices.

Expertise tends to be domain-specific. However, the role of domain-specificity as a function of expertise has not been investigated in graph exploration studies before and this formed Hypothesis 2: Experts would provide more complete explanations of graphs in familiar than in unfamiliar domains.

Carpenter and Shah (1998) found the proportion of nominal, ordinal, and metric descriptions of graphs varied across different graph types. Nominal utterances were defined as the names of *z*-variables without any ordinal or metric information about the *z-y* relation; ordinal utterances mentioned the explicit relationships between *z*-variables; and metric utterances included descriptions of the interval or ratio relationship between *z*-variables. Equating Carpenter and Shah's nominal, ordinal, and metric descriptions with the different types of conceptual messages proposed by Pinker (1990) we may be able to extend Pinker's model to include expertise and which leads to Hypothesis 3: Experts would generate more conceptual messages (nominal, ordinal, and metric combined) than Novices.

In order to understand how expertise might exert its effects on graph exploration and to better control for potential floor- and ceiling effects, both simple and complex graphs were employed. Somewhat more complex graphs might allow Experts to demonstrate superiority over Novices, as predicted by Hypothesis 4: Experts will provide more complete graph explanations than Novices.

# 2. METHOD

## 2.1 Participants

Twenty-six participants were recruited from the Carleton University community. Out of ten (seven female) undergraduate Novices, six were majoring in psychology and four in business. The Expert sample comprised eight psychology (seven female) and eight business (six female) PhD faculty. Five Novices, six business Experts, and seven psychology Experts reported that they had to create graphs and all reported that line graphs were the graphs most familiar to them. Novice undergraduate students were granted 1.0% course credit, and Experts were given a $10 coffee shop gift certificate for their participation. All had normal or corrected-to-normal vision. Participants were tested individually in sessions lasting a mean of 75 minutes.

## 2.2 Apparatus & Materials

Ten, three-point, two *z*-variable line graphs were used, five simple and five more complex as determined through pilot testing. Each graph was assigned two sets of titles, labels and axes; one drawn from psychology and one from business. The business labels were selected from an undergraduate textbook on international business (Griffin & Pustay, 2007), and psychology labels were drawn from an undergraduate textbook on psychology (Weiten & McCann, 2007). The 10 business graphs were the mirror images of the 10 psychology graphs as shown in the typical examples in Figure 1 below.

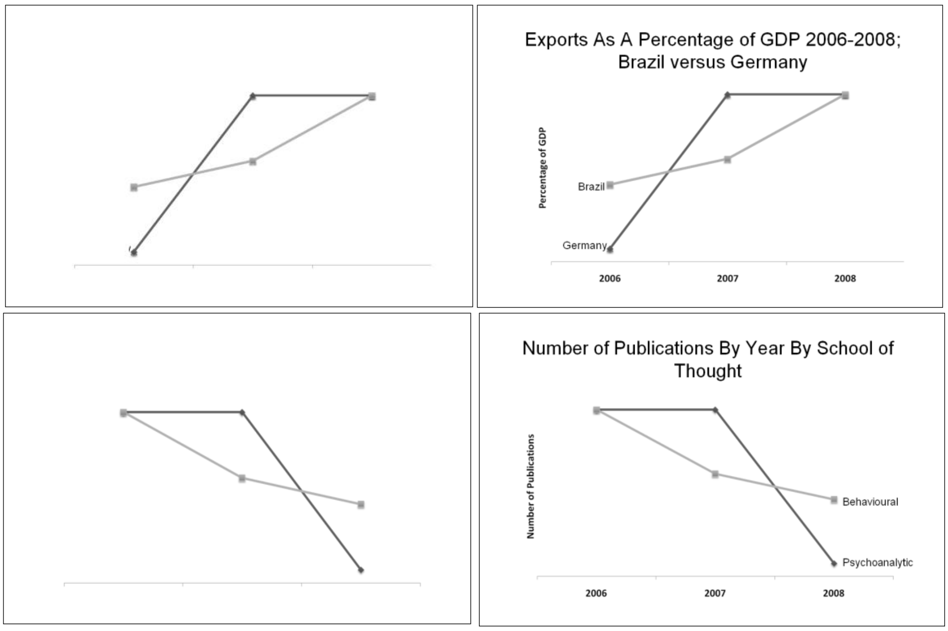


Figure 1. Example graph stimuli: business domain (top) and psychology domain (bottom); "Describe" (left) and "Explain" (right)

Stimulus presentation was randomized, controlled by DirectRT™ on a Dell Latitude D610 laptop computer with 1280 X 800 pixel screen resolution. Participant verbalizations were recorded on a Panasonic RR-US500 digital voice recorder.

## 2.3 Procedure

After Preliminary instructions and Informed Consent, detailed experimental instructions were provided. Four practice trials were followed by 20 experimental trials, each initiated by pressing the spacebar. On each trial a randomly selected graph without labels or titles was displayed with the word "describe" played over the computer speakers as well as appearing at the bottom of the display. (Pilot testing had indicated that alerting participants to the visual characteristics of a graph was important in order to prime their subsequent explanations.) When done, participants were instructed to press the spacebar whereupon the graph was re-displayed with the corresponding business or psychology labels and titles accompanied by the instruction "explain" played over the computer speakers and displayed on the screen. At the end of the experiment participants were debriefed, thanked, and paid (if applicable).

ISBN: 978-989-8533-23-4 © 2014 IADIS

## 2.4 Data Analysis

Verbal protocols were transcribed ad verbatim, coded, and analyzed with NVIVO™ Version 8.0. Frequency of utterance-type was calculated, as was the presence/absence and completeness of explanations and the frequency of conceptual messages (sum of nominal, ordinal, and metric utterances). Interrater reliability was assessed by an independent rater coding a randomly selected 15% of the verbal protocols and percentage agreement was 90.0%.

# 3. RESULTS

Coding of the graph explanation protocols resulted in nine themes, shown for each expertise group in Table 1. Values are proportions of the total number of trials per expertise group to enable direct comparison of the different groups. Because themes are not mutually exclusive, they do not sum to 1.0.

Novices and Experts differed in the frequency with which they voiced most themes. These differences will be reviewed in the context of the four Hypotheses followed by an unanticipated result related to response time.

Table 1.Utterance themes, examples, and mean proportions by Novices, Business Experts, and Psychology Experts

| Theme | Examples | Novice | BusExp | PsyExp |
|---|---|---|---|---|
| BECAUSE: Inferences of causality | "don't know whether they've had a change in government or if officials have just gotten a lot more corrupt but…" | .19 | .41 | .54 |
| BETWEEN *Z*: Comparisons between *z*-variables | "In 2008 the big 10 and the emerging economies have an equal amount of annual average growth in GDP" | .61 | .83 | .97 |
| DIRECTION: Within a single *z*-variable | "Azerbaijan is expected to remain stable …over 2010 to 2011, but then is predicted to decrease their instability" | .52 | .58 | .46 |
| QUANTITATIVE: Interval or ratio relationship | "difference increases dramatically in 2007. It is maybe 5 times or 4 times greater in 2007…" | .03 | .25 | .39 |
| TITLE: Repeat title of the graph | "hypnotic susceptibility by field dependence by gender" | .32 | .91 | .92 |
| TREND: Overall direction | "over a 3-year span, both groups seem to be decreasing the number of publications" | .11 | .18 | .35 |
| *X*-AXIS: References to abscissa | "*x*-axis shows Day 1, Day 2, Day 3" | .25 | .45 | .36 |
| *Y*-AXIS: References to ordinate | "The *y*-axis shows GDP—adjusted GDP—in billions of US dollars." | .23 | .40 | .28 |
| *Z*-Variable: Number or name of *z*- | "The two lines represent…, respectively, the scores for males and for females…" | .18 | .20 | .25 |

## 3.1 Proportion of "Because" Inferences

Although all participants were asked to "explain the graph as if you were the author and you were explaining the results to another person", utterances of the form "variable *a* causes variable *b*" were observed infrequently in Novices. A repeated measures 3 (Expertise: novice, business expert, psychology expert) x 2 (Difficulty: simple, complex) x 2 (Domain: business, psychology) ANOVA revealed only a significant main effect of expertise, $F(2, 23) = 4.73$, $p = .019$, $\eta_p^2 = .29$. Independent post hoc Tukey tests confirmed that psychology experts ($M = .57$) attempted more inferences than novices ($M = .18$), $p = .015$; the difference between business- and psychology experts was not significant ($p = .440$), and nor was the difference between business experts and novices ($p = .214$). Hypothesis 1 stating that Experts would provide more inferences than Novices was thus supported.

## 3.2 Familiar and Unfamiliar Domains

Hypothesis 2 stated that Experts would generate more complete explanations of graphs in familiar than unfamiliar domains. Excluding Novices, a repeated measures 2 (Expertise, business, psychology) x 2 (Domain: familiar, unfamiliar) x 2 (Difficulty: simple, complex) ANOVA resulted in only one significant effect, the Expertise x Domain interaction, $F(1, 14) = 6.56$, $p = .023$, $\eta p2 = .56$. Post hoc t-tests for independent samples confirmed that the interaction was due to higher completeness scores for psychology Experts on familiar domain graphs ($M = .53$) compared to the unfamiliar domain ($M = .48$), $t(7) = 3.30$, $p = .013$, and business Experts exhibited the opposite effect of significantly lower completeness scores on familiar domain graphs ($M = .45$) compared to the unfamiliar domain ($M = .49$), $t(7) = -2.71$, $p = .030$. Thus Hypothesis 2 was partially confirmed by psychology Experts but refuted by business Experts.

## 3.3 Nominal, Ordinal, and Metric Conceptual Messages

The proportion of conceptual messages is the sum of z-variable, Between z-variable, and Quantitative proportions (refer to Table 1). A 3 (Expertise: novice, business expert, psychology expert) x 2 (Difficulty: simple, complex) ANOVA of the conceptual messages resulted in a significant expertise main effect, $F(2, 23) = 7.80$, $p = .003$, $\eta p2 = .40$. Planned comparisons indicated that business Experts ($M = 1.64$) generated more conceptual messages than Novices ($M = 1.11$), $p = .012$, psychology Experts ($M = 1.55$) generated more than Novices, $p = .013$ but business and psychology Experts did not differ from each other, $p = .59$ confirming Hypothesis 3.

## 3.4 Completeness

A completeness score was calculated by determining the proportion of all nine themes present in each participant's explanation of each graph. The mean completeness scores are shown in Figure 2 for each expertise group and for each domain. The Figure suggests that the two Expert groups' explanations were more complete than those of Novices and that this was more pronounced for psychology than for business graphs. This was confirmed by a repeated measures 3 (Expertise: novice, business expert, psychology expert) x 2 (Difficulty: simple, complex) x 2 (Domain: business, psychology) ANOVA on completeness scores. The main effect of expertise was significant, $F(2, 23) = 8.02$, $p =. 002$, $\eta p2 = .41$ and independent Tukey post hoc comparisons confirmed that business Experts ($M = .47$) provided more complete explanations than Novices ($M = .27$), $p = .014$, and the same was also true for the psychology Experts ($M = .51$), $p = .004$, confirming Hypothesis 4.

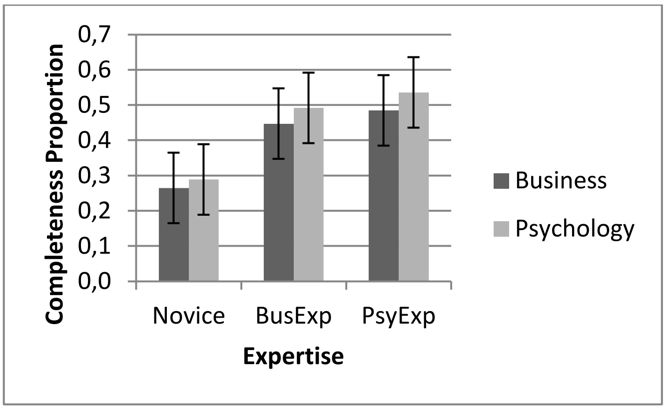


Figure 2. Mean completeness scores for expertise and graph domain. (95% confidence intervals were calculated using the procedure of Jarmasz & Hollands, 2009)

## 3.5 Total Response Time, Silent Time, Explain Time

Total Response Time (TRT) for the graph explanation task was composed of Silent Time (ST) plus Explain Time (ET). ST was the silent period before participants began their graph explanations and ET was the time during which participants voiced their explanations. A repeated measures 3 (Expertise: novice, business expert, psychology expert) x 2 (Difficulty: simple, complex) x 2 (Domain: business, psychology) ANOVA for TRT revealed no main effect for Expertise ($p = .817$) nor any significant interactions with Expertise, suggesting that the efficiency with which Experts completed the graph explanation task was no better than that of Novices.

A repeated measures 3 (Expertise: novice, business expert, psychology expert) x 2 (Difficulty: simple, complex) x 2 (Domain: business, psychology) ANOVA for ET revealed no main effect for Expertise ($p = .478$) nor any significant interactions with Expertise.

However, a repeated-measures 3 (Expertise: novice, business expert, psychology expert) x 2 (Difficulty: simple, complex) x 2 (Domain: business, psychology) ANOVA on ST revealed a significant main effect of expertise, $F(2, 23) = 7.71$, $p = .003$, $\eta p2 = .41$. Independent Tukey post hoc tests confirmed that novices had longer silent periods before beginning their explanations ($M = 8.95$ s) than business ($M = 2.17$ s,), $p = .003$, or psychology Experts, ($M = 3.37$ s), $p = .016$. If ST represents the time required to select and/or initiate a strategy then Experts required less time to select their graph explanation strategies than Novices. Novices appeared uncertain about what to say or perhaps how to start their graph explanations.

## 3.6 Results Summary

The current research demonstrated a difference between Novices and Experts in their graph exploration in terms of the proportion of time Experts attempted inferences in their interpretation of the graph data; and the completeness of their explanations. The greater Silent Time of Novices before initiating their explanations suggests that undergraduate students struggle with an appropriate strategy to attempt their efforts and the results suggest a parsimonious extension to the graph comprehension model of Pinker (1990). However, the lack of a consistent effect of familiar versus unfamiliar domain in the performance of Experts leaves some question as to the locus of these effects—whether they are evidence of a general expertise effect or one limited to a specific domain. These results are summarized in Table 2.

Table 2. Research hypotheses, results, and conclusions

| Hypotheses | Results | Conclusions |
|---|---|---|
| H1. Experts will generate more "because" inferences than novices | • More "because" inferences by BusExp & PsyExp than Novices<br>• Similar number of "because" inferences by BusExp & PsyExp | Expertise effect in graph exploration supported |
| H2. Experts will provide more complete graph explanations for familiar compared to unfamiliar domain graphs. | • PsyExp psych domain explanations more complete than business domain<br>• BusExp business domain explanations similar completeness scores to psych domain | Domain-specificity of graph expertise partially supported |
| H3. Experts will generate more conceptual messages than Novices. | • BusExp and PsyExp generated more conceptual messages than Novices | Supports extension of Pinker (1990) model |
| H4. Experts will generate more complete explanations than novices | • Higher completeness scores by Experts than Novices | Supports perspective on graph exploration where completeness=expertise |
| Unanticipated | • Silent Time greater for Novices<br>• Explain Time similar for all groups | Suggests that Expert/Novice differences may be due to conscious strategy |

## 4. CONCLUSIONS

The present research contributed to an understanding of graph exploration in three ways. First, the experiment is among the first to demonstrate an expertise "effect" in the domain of graph exploration. Although others have studied graph expertise (e.g., Roth, 2004; Roth & Bowen, 2003) they have not directly contrasted novice and expert performance. Previous attempts to distinguish novice and expert graph comprehension (Shah & Carpenter, 1995) found no differences between the two types of participants. However, since the effect of domain was inconclusive in the current research, it remains unknown whether this expertise effect is general or limited to specific domains.

Second, this experiment showed that Experts adopt a graph exploration strategy in which specific elements of a graph are explored. It is proposed that these elements represent a list of conceptual questions that is the embodiment of a graph exploration strategy. The addition of this top-down process adds clarity to Pinker's (1990) model of graph comprehension by introducing a mechanism for the operation of expertise. In contrast, novices' strategies were inconsistent. As a consequence, it took them longer to initiate their graph exploration, and their explanations were less complete than those of the experts.

Finally, the issues identified here in Expert/Novice differences in graph explanation lend themselves to intriguing ideas in education and data visualization. Perhaps it would be possible to address these to improve the data literacy of children or older students (e.g. Feldon et al., 2010), or in the teaching of statistics (e.g. Cleveland, 1987; Huff, 1954). In particular, it is reasonable to believe that an instantiation of the Expert graph exploration strategies determined here might be embodied in a training regimen to bootstrap the understanding of data by Novices. This is research that we have currently underway. It is also easy to imagine these reflected in computer-based data visualization tools (e.g. Heer, et al., 2010; Konold, 2007; http://datavisualization.ch/tools/).

Unfortunately, the data are insufficient to determine if the inconsistent effect of domain provides evidence of a global expertise effect or if they are limited to specific domains. Perhaps more complex graphs, in terms of either visual or semantic complexity would have resulted in more definitive evidence. A replication of the current research using interactive graphs might be particularly informative.

In conclusion, the importance of this line of research is underscored by regular national comparisons of student performance in mathematics (e.g. OECD, 2014). The OECD Programme for International Student Assessment asserts that the *application* of mathematics (including graph exploration) is a key attribute of "What is important for citizens to know and be able to do?" (OECD, 2014 p. 3). The current research may contribute to an improvement in what students can do with data.

## REFERENCES


Behrens, J. T., 1997. Principles and procedures of exploratory data analysis. *Psychological Methods, 2*(2), 131-160.

Bertin, J., 1983. *Semiology Of Graphics.* University of Wisconsin Press, Madison, WI.

Carpenter, P. A., & Shah, P., 1998. A model of the perceptual and conceptual processes in graph comprehension. *Journal of Experimental Psychology: Applied, 4*(2), 75-100.

Chambers, J. M., et al., 1983. *Graphical methods for data analysis.* Wadsworth International Group, Belmont, CA.

Cleveland, W. S., 1987. Research in statistical graphics. *Journal of the American Statistical Association, 82*(398), 419-423.

Datavisualizations's *A carefully selected list of recommended tools*. May 7, 2012. Retrieved April 22, 2014 from http://datavisualization.ch/tools/

Ericsson, K. A., 2005. Recent advances in expertise research: A commentary on the contributions to the special issue. *Applied Cognitive Psychology, 19*, 231-241.

Feldon, D.F., et al., 2010. Translating expertise into effective instruction: The impacts of cognitive task analysis (CTA) on lab report quality and student retention in the biological sciences. *Journal of Research in Science Teaching, 47(10),* 1165-1185.

Freedman, E. G., & Shah, P., 2002. Toward a model of knowledge-based graph comprehension. In M. Hegarty, B. Meyer & N. H. Narayanan (Eds.), *Diagrams 2002, LNAI 2317* (pp. 18-30). Springer-Verlag, Berlin, Germany.

Gick, M. L., & Holyoak, K. J., 1983. Schema induction and analogical transfer. *Cognitive Psychology, 15*, 1-38.

Griffin, R.W., & Pustay, M.W., 2007. *International Business: A Managerial Perspective* (5th ed.). Pearson Education, Upper Saddle River, NJ.

Heer, J., et al., 2010. A tour through the visualization zoo: A survey of powerful visualization techniques from the obvious to the obscure. *ACMQueue, 8*(5), 1-22.

Huff, D., 1954. *How To Lie With Statistics.* W. W. Norton & Company Inc, New York, NY.

Jarmasz, J., & Hollands, J.G., 2009. Confidence intervals in repeated-measures designs: The number of observations principle. *Canadian Journal of Experimental Psychology, 63*(3), 124-138.

Konold, C. 2007. Designing a data analysis tool for learners. In M. C. Lovett & P. Shah (Eds.) *Thinking With Data*. (pp. 267-291). Lawrence Erlbaum, New York, NY.

Kosslyn, S.M., 2006. *Graph Design For The Eye And Mind*. Oxford University Press, New York, NY.

Newell, A., & Simon, H. A., 1972. *Human Problem Solving.* Prentice-Hall, Englewood Cliffs, NJ.

OECD, 2014. *PISA 2012 Results: What Students Know and Can Do – Student Performance in Mathematics,*

*Reading and Science* (Volume I, Revised edition, February 2014), PISA, OECD Publishing. Retrieved April 22, 2014 from http://www.oecd.org/pisa/keyfindings/pisa-2012-results-overview.pdf

Pinker, S., 1990. A theory of graph comprehension. In R. Freedle (Ed.), *Artificial Intelligence And The Future Of Testing.* (pp. 73-126). Lawrence Erlbaum, Hillsdale NJ.

Preece, J., & Janvier, C., 1993. Interpreting trends in multiple-curve graphs of ecological situations: The role of context. *International Journal of Science Education, 15*(2), 199-212.

Roth, W., 2004. Emergence of graphing practices in scientific research. *Journal of Cognition and Culture, 4*(3-4), 595-627.

Roth, W. M., & Bowen, G. M., 2003. When are graphs worth ten thousand words? an expert-expert study. *Cognition and Instruction, 21*(4), 429-473.

Roth, W., et al., 1999. Differences in graph-related practices between high school biology textbooks and scientific ecology journals. *Journal of Research in Science Teaching, 36*(9), 977-1019.

Shah, P., & Carpenter, P. A., 1995. Conceptual limitations in comprehending line graphs. *Journal of Experimental Psychology: General, 124*(1), 43-61.

Trafton, J. G., et al., 2002. Extracting explicit and implicit information from complex visualizations. In M. Hegarty, B. Meyer, N. Narayanan (Eds.) *Diagrams 2002, Lecture Notes in Computer Science*, 2317/2002, 206-220.

Tufte, E. R., 1983. *The Visual Display Of Quantitative Information.* Graphics Press, Cheshire, CT.

Weiten, W., & McCann, D., 2007. *Psychology: Themes and Variations*. (1st Canadian ed.). Nelson Education, Toronto, ON.