Learning and Instruction 20 (2010) 155–166

Learning and Instruction

www.elsevier.com/locate/learninstruc

Effects of knowledge and display design on comprehension of complex graphics

Matt Canham, Mary Hegarty*

Department of Psychology, University of California, Santa Barbara, Santa Barbara, CA 93106, USA

Abstract

In two experiments, participants made inferences from weather maps, before and after they received instruction about relevant meteorological principles. Different versions of the maps showed either task-relevant information alone, or both task-relevant and task-irrelevant information. Participants improved on the inference task after instruction, indicating that they could apply newly acquired declarative knowledge to make inferences from graphics. In Experiment 1, participants spent more time viewing task-relevant information and less time viewing task-irrelevant information after instruction, and in Experiment 2, the presence of task-irrelevant information impaired performance. These results show that domain knowledge can affect information selection and encoding from complex graphics as well as processes of interpreting and making inferences from the encoded information. They also provide validation of one principle for the design of effective graphical displays, namely that graphics should not display more information than is required for the task at hand.

Keywords: Graphical comprehension; Eye fixations; Meteorology instruction; Display design

1. Introduction

Cognitive models of graphics comprehension have an important influence on theories of how to use media in instruction and how to develop students' graphical literacy (Ainsworth, 2006; Schnotz & Bannert, 2003). These models propose the following three component processes in understanding a graphical display (Bertin, 1983; Carpenter & Shah, 1998; Pinker, 1990). First, users must encode the visual features of the display (e.g., lines of different slopes in a line graph). Next they must map these onto the conceptual relationships that they convey (e.g., an upwardly sloping line shows an increasing quantity). Finally, they need to relate these conceptual relationships to the referents of the graphs (e.g., an upwardly sloping line represents an increase in the value of some stock). These models propose that graphics comprehension involves interaction between bottom-up perceptual processes of encoding information from the graphic and top-down processes of applying graph schemas and domain knowledge. They are supported by studies showing that graphics comprehension is significantly affected by the display format (Shah, Mayer, & Hegarty, 1999; Simkin & Hastie, 1986), knowledge of graphics conventions (Körner, 2005; Shah, Freedman, & Vekiri, 2005) and domain knowledge (Freedman & Shah, 2002; Lowe, 1993).

Existing models of graphic comprehension are limited in a number of ways (Trafton & Trickett, 2001; Trickett & Trafton, 2006). First, they have been applied primarily to comprehension of relatively simple displays, such as bar or line graphs showing two or three variables and less than a dozen data points (Carpenter & Shah, 1998; Simkin & Hastie, 1986). For these displays, it is plausible that viewers attend to all the information in the graphic. However, using more complex graphics often involves selecting task-relevant information from a much larger amount of displayed information. Second, current models focus on simple tasks in which the users read off values from the graph (Lohse, 1993; Peebles & Cheng, 2003) or describe trends in the displayed data (Carpenter & Shah, 1998; Shah & Carpenter, 1995), and rarely address situations in which new information must be inferred

* Corresponding author. Tel.: +1 805 893 3750.
E-mail address: hegarty@psych.ucsb.edu (M. Hegarty).

0959-4752/$ - see front matter
doi:10.1016/j.learninstruc.2009.02.014

from the information that is shown in the display, as pointed out by Trafton and Trickett (2001). Third, existing models do not specify "at what stage during the process of graph comprehension does content knowledge influence comprehension" (Shah et al., 2005, p. 461). Finally, current models of graph comprehension do not specify how characteristics of graphical displays and knowledge effects interact.

The present study examined comprehension of relatively complex graphical displays, that is, weather maps, in a task that requires not just encoding of information from displays, but also making inferences from this information. The effects of manipulating both the complexity of the display and the viewer's domain knowledge were examined, in order to reveal possible interactions between display format and knowledge. Finally, eye fixations and measures of other aspects of task performance were recorded, in order to study how these factors affect both the selection of information for encoding from a graphic and the inferences made from that information.

1.1. Effects of knowledge

At what stage in the process of graphics comprehension does knowledge affect comprehension? A possible answer to this question is that perception and encoding of the information in graphical displays is a purely bottom-up process and knowledge comes into play only in interpreting the meaning of the visual relationships and making inferences after the information in the external display has been encoded. But knowledge also has the potential to affect which locations and visual features are fixated and consequently encoded. Recent research on viewing of other complex visual displays, such as pictures of natural scenes, has emphasized the top-down effects of knowledge on scene perception (Henderson, 2003; Henderson & Ferreira, 2004). This research suggests that eye fixations on meaningful scenes are primarily directed by knowledge, including short-term episodic knowledge built up while viewing a scene, scene schema knowledge such as the typical locations of objects in that scene, and knowledge relevant to gaze control in the service of a specific task.

Studies with more abstract representations, such as X-ray images (Myles-Worsley, Johnston, & Simons, 1988) and chess diagrams (Chase & Simon, 1973; Reingold, Charness, Pomplun, & Stampe, 2001) also suggest that there are top-down effects of knowledge on processing of visual displays in showing that experts and novices attend to different aspects of visual displays and extract different information from these displays. In meteorology, the domain of interest here, novices tend to focus primary on a weather map's superficial features, whereas experts focus on elements that are thematically relevant (Lowe, 1993, 1994, 1996, 2004). For example, when experts and novices are asked to sort features from weather maps into clusters, experts group features that are causally related, whereas novices cluster features in terms of surface similarity (Lowe, 1996). Research on expertise in graphical comprehension has mostly been based on indirect evidence, such as the sorting task used by Lowe (1996). In the present study eye fixations were examined to show that knowledge affects where people actually look in a graphical display, following research of Reingold et al. (2001) in the domain of chess.

The tendency for experts to focus more on thematically relevant aspects of visual displays, and less on salient but less relevant features can be seen as an example of the information reduction hypothesis (Haider & Frensch, 1996, 1999; Jarodzka, Scheiter, Gerjets, & Van Gog, 2010). Haider and Frensch (1996, 1999) showed that with extensive practice on a laboratory task (making judgments about alphanumeric character strings), participants learned to ignore task-irrelevant information, and this occurred at a perceptual level of processing such that participants initially made equal numbers of eye fixations on task-relevant and task-redundant information, but with repeated practice, they made significantly fewer fixations on task-redundant information. In the present study the information reduction hypothesis was examined to determine whether it can be generalized to comprehension of weather maps and to situations in which people are explicitly taught domain-relevant information rather than gradually developing expertise.

1.2. Effects of display format

In addition to top-down effects of knowledge, there is evidence that the format of the display can affect comprehension of various kinds of graphical displays (Ainsworth, 2006; Cheng, 1999; Shah et al., 1999; Simkin & Hastie, 1986), suggesting that there are also bottom-up effects of display format on comprehension. In this regard, many theorists have prescribed how graphics should be designed for efficient performance (Bertin, 1983; Kosslyn, 1989; Tufte, 1983), but few of these prescriptions have received empirical validation. One cardinal rule, proposed by Kosslyn (1989) is that "no more or less information should be provided than is needed by the user" (p. 211). Tufte (1983) also cautions against including extraneous information in visual displays, calling this information "chartjunk". In instructional situations, extraneous information in a display, especially when it is highly salient or interesting, may attract students' attention away from the task-relevant information, and may need to be consciously suppressed for good comprehension (Sanchez & Wiley, 2006). This process can be seen as a source of extraneous cognitive load (Sweller & Chandler, 1994), which may be particularly high when viewers have limited domain knowledge to discern what information is task-relevant. Another aim of the present study was to provide empirical support that extraneous information in a graphical display impairs performance of students and examine how effects of extraneous information interact with effects of domain knowledge.

1.3. The experimental task

A graphics comprehension task in the domain of meteorology was used in the present study. Meteorology is an ideal domain in which to examine interactions between knowledge and display design. Weather maps vary in complexity, sometimes showing only a single variable (e.g., temperature) but more typically displaying several variables simultaneously (Hoffman, Detweiler, Conway, & Lipton, 1993). In addition, weather map comprehension typically involves going beyond the information that is explicitly displayed, by applying domain knowledge to make inferences. For example, when a meteorologist inspects a weather map, he or she is not just interested in reading off the current weather conditions, but in predicting how the weather will change in the future, which involves making inferences from the displayed information.

In the experiments reported in the present article, participants were shown a weather map displaying either pressure information alone or pressure-plus-temperature and place (see the sample trials in Fig. 1). An arrow on the map indicated a possible direction of wind in one region of the map. The task was to judge whether this showed the actual direction in which the wind would be blowing in that region. Wind direction can be inferred from knowledge of the pattern of surface pressure in an area, but is unrelated to temperature, so the pressure information on the map is *relevant* information for this inference task and temperature information is *irrelevant*. Moreover, the wind direction in an area is most influenced by the closest pressure system, so that the most relevant place to look on a weather map when inferring the wind direction in a region is the closest pressure system to that region.

Inferring wind direction from pressure is based on two meteorological principles, the pressure gradient principle and the Coriolis phenomenon (Ahrens, 2000). The pressure gradient principle refers to the tendency for air (wind) to flow from areas of high pressure toward areas of low pressure. The Coriolis phenomenon is due to the rotation of the earth, and causes air to circulate in a clockwise direction around high pressure systems, and counterclockwise around low pressure systems, in the Northern hemisphere. Because of the combination of these principles, and because friction at the surface of the earth reduces the Coriolis phenomenon, air tends to move clockwise and outward around areas of high pressure and counterclockwise and inward around areas of low pressure in the Northern hemisphere. So in the examples in Fig. 1, the arrow shows the correct direction of wind because it is pointing in a counterclockwise direction relative to the adjacent low pressure system and is pointing slightly inward.

2. Experiment 1

In Experiment 1, participants' eye fixations and their performance on this wind inference task were examined before and after they were taught the relevant meteorological principles. This allowed assessment of how well students were able to apply their newly acquired knowledge of the pressure differential principle and Coriolis phenomenon to infer wind direction; their accuracy on this inference task was expected to improve (Hypothesis 1). To examine possible interactions between graphical displays and knowledge, performance with displays presenting only the relevant information (pressure) and displays presenting additional irrelevant information (temperature and place) were also contrasted. The main prediction was that the irrelevant information in the more complex displays would impair performance, leading to either longer response times or lower accuracy rates (Hypothesis 2). A secondary prediction was that these effects would be reduced after instruction, when participants had more domain knowledge to discern which information is relevant to the task (Hypothesis 3).

The most critical question was whether acquiring meteorological knowledge would affect participants' eye fixations as well as their task performance. If domain knowledge affects where people look on graphical displays as well as how they interpret what they see on these displays, then participants should spend relatively more time inspecting the task critical information (i.e., the closest pressure system to the target arrow) and less time viewing task-irrelevant information (e.g., the temperature scale) after instruction compared to before instruction (Hypothesis 4). In contrast, if domain knowledge comes into play only after the information in a graphical display has been encoded (and is not involved in selecting where to look on the display), then we should not observe differences in eye fixations from before to after instruction (null hypothesis). Participants were expected to look at the arrow both before and after instruction, because their task was to verify the direction of the arrow.

2.1. Method

2.1.1. Participants

Participants in Experiment 1 were 16 students (nine females and seven males) from an introductory psychology

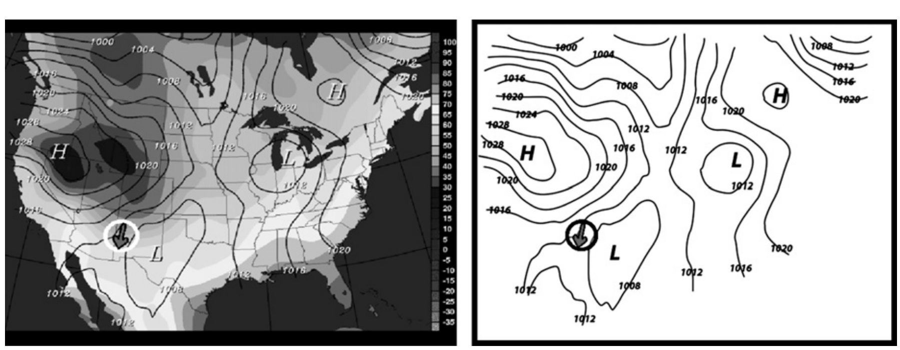


Fig. 1. Examples of pressure-plus-temperature and pressure-only maps in Experiment 1.

class (age ranged from 18 to 22 years). None had any formal knowledge of meteorology. One participant's data were not analyzed, due to poor calibration of the eye tracking apparatus.

2.1.2. Materials

2.1.2.1. Weather maps. The weather maps were obtained online from Weather World 2010 Project (2000) and showed the actual weather conditions in North America on ten different dates in the previous 10 years. One version of each of the 10 maps, labeled "pressure-plus-temperature maps", showed an outline of the United States, with information about temperature (represented by color), and pressure (represented by isobars, i.e., lines of equal pressure). The other map for each date, labeled "pressure-only maps" depicted only the isobars and pressure systems (see Fig. 1).

Each trial showed one of these maps with a target area indicated by a circle and an arrow inside the circle, indicating a hypothetical wind direction. On one third of trials ("correct arrow" trials), the arrow showed the correct direction of surface winds (angling 20° inward and counterclockwise around low pressure systems and 20° outward and clockwise around high pressure systems). On one third ("opposite arrow" trials), the arrow was in the opposite direction (i.e., different by 180°). On the final one third ("pressure gradient" trials), the arrow pointed directly into a low pressure system (or directly out of a high-pressure system), that is, the direction of wind if only the pressure gradient principle applied (approximately 70° different from correct). Type of arrow was crossed with type of map for the 10 different dates, giving a total of 60 trials. Half of the trials in each condition of the design (map by arrow by date) were shown before instruction and half were shown after instruction. The before- and after-instruction trials were presented in a different random order to each participant.[1]

2.1.2.2. Tutorial. A tutorial presented in Powerpoint, was adapted from the description of the pressure gradient and Coriolis phenomena in meteorology textbooks and the Weather World 2010 Project (2000) online tutorial, explained how these factors influence wind movement (see Appendix A) and provided three worked examples (see example in Appendix B).

2.1.2.3. Knowledge of principles questionnaire. A four-item questionnaire, presented in Appendix C, was used to assess participants' knowledge of the meteorological principles after they studied the tutorial. Scoring of the items is given at the end of Appendix C.

2.1.3. Apparatus

Eye movements were monitored by an SMI EyeLink head mounted eye tracking system which sampled eye position every 4 ms (250 Hz). The aggregation software was set to detect saccades with an amplitude of .05° or greater, an acceleration threshold of 9500° per second squared and a velocity threshold of 30° per second. Participants viewed images presented on a computer monitor screen while resting their chins on a chin rest, set 30 inches from a 15×11.5 viewing screen ($21.7° \times 28.08°$ visual angle). The resolution of the screen was 800×600, with a refresh rate of 75 Hz.

2.1.4. Procedure

Participants were tested individually. After calibration of the eye tracker, they were familiarized with the graphical conventions of the maps and were instructed that on each trial they would see an arrow indicating the possible wind direction at a particular location. Their task was to judge whether the arrow indicated the correct direction (true) or incorrect direction (false) of wind and to indicate their answer by pushing one of two buttons on a button box. After six practice trials (with the same format as the experimental trials), the participants performed the first block of 30 trials.

Then participants were given the tutorial, which they studied at their own pace, and had the opportunity to ask questions. Next, they completed the Knowledge of Principles questionnaire to assess how well they had learned the meteorological principles and were given feedback on their answers. If participants missed any of the questions, they were asked to review the presentation to find the correct answer. Next participants were given the three worked examples (see example in Appendix B). Finally, the eye tracker was recalibrated and the participants completed the second set of 30 trials.

2.1.5. Coding of eye fixations

To analyze the eye fixations, areas of interest on the maps that were either relevant or irrelevant to inferring the wind direction were defined. The first relevant area on each map was the circle containing the arrow to be verified. This had a diameter of 50 pixels (about 2° of visual angle) and corresponded to 0.4% of the display. The second relevant area was the closest pressure system to the arrow, which was a circular region with a diameter of 200 pixels (about 7° of visual angle) and took up 6.5% of the display area. Note that these regions partially overlap, as the closest pressure system is by definition next to the arrow. The number of fixations and the total amount of time spent fixating each of these areas of interest (looking time) were computed for each participant on each trial of the experiment. Number of fixations and total time spent on all other areas of the map (outside the relevant areas of interest) were also computed. Finally, a rectangular region of 100×600 pixels defined the temperature scale on the pressure-plus-temperature maps, and represented 12.5% of the display area of the screen. This region is irrelevant, as wind direction does not depend on temperature.

Measures of the number of fixations and the total looking time in the areas of interest were highly correlated (median correlation = .98 across all variables computed), so only the

[1] Due to experimenter error, an incorrect map was shown on one temperature-plus-pressure after-instruction trial, so this trial was not included in computing the variables for this condition.

analyses for total looking time are reported.[2] Because response time was somewhat variable across participants, looking time on the different areas of interest for each participant was expressed as a proportion of his or her total looking time on the map, in order to reduce inter-subject variability (Cohen, Cohen, West, & Aiken, 2003).

2.2. Results and discussion

2.2.1. Effects of knowledge and display on accuracy[3]

The predictions were that performance on the wind inference task would improve as a result of instruction (Hypothesis 1) would be better for the pressure-only maps (Hypothesis 2) and that the effects of map display would be reduced after instruction (Hypothesis 3). Accurate performance was defined as responding "true" for the correct arrows (i.e., hits) and "false" for opposite and pressure gradient arrows (i.e., correct rejections) and analyzed in a 2 (before vs. after instruction) × 2 (map type) × 3 (type of arrow) ANOVA. Descriptive statistics for the different conditions of the design are shown in Table 1. Consistent with Hypothesis 1, performance improved from before instruction ($M = .51$, $SD = .13$) to after instruction ($M = .76$, $SD = .10$), $F(1, 14) = 28.30$, $p < .001$, partial $\eta^2 = .67$. A signal detection analysis indicated that sensitivity (d')[4] also increased from before to after instruction (from 36 to 1.71 for pressure-plus-temperature maps; from .02 to 1.51 for pressure-only maps). Contrary to Hypothesis 2, there was no significant difference between performance on the pressure-only maps ($M = .64$, $SD = .06$) and the pressure-plus-temperature maps ($M = .62$, $SD = .07$), $F(1, 14) < 1$. Furthermore the interaction of instruction with map type was not significant, $F(1, 14) < 1$, that is, Hypothesis 3 was not supported. In summary, the accuracy data indicated that participants learned from the tutorial and were able to apply their knowledge to the inference task, but map type had no significant effect on accuracy.

A secondary result was that there was a significant main effect of the arrow to be verified, $F(2, 28) = 14.71$, $p < .001$, partial $\eta^2 = .51$, such that participants had poorer performance on the pressure gradient arrows than on the other two arrows. Failure to reject this arrow (but good performance otherwise) is predicted if participants apply the pressure differential principle but do not take the Coriolis phenomenon into account (this arrow is consistent with the pressure differential principle but inconsistent with the Coriolis phenomenon). Participants were classified as showing this pattern if they answered correctly for the majority of "correct" and "opposite" arrows, but incorrectly for the majority of "pressure gradient" arrows. Six of the 15 participants showed this pattern, five answered correctly on the majority of trials for all three types of arrows, and four did not show either of these patterns.

Table 1
Means (and SD) of accuracy (proportion correct) in Experiment 1 for the two maps.

Arrow	Before instruction	After instruction
Pressure-plus-temperature map		
Correct	.64 (.22)	.85 (.21)
Opposite	.56 (.24)	.96 (.11)
Pressure gradient	.40 (.39)	.44 (.29)
Pressure-only map		
Correct	.53 (.29)	.80 (.23)
Opposite	.56 (.25)	.96 (.11)
Pressure gradient	.36 (.34)	.53 (.28)

Mean performance on the knowledge of principles questionnaire was 3.07 out of 4 ($SD = 1.03$) and was not significantly correlated with accuracy on any of the after-instruction arrows ($r = .15$ with correct arrow, $r = -.21$ with opposite arrow, $r = .24$ with pressure gradient arrow), indicating that knowledge of the meteorological principles does not imply ability to apply these principles to the inference task.

2.2.2. Effects of knowledge and display on eye fixations

If domain knowledge affects where people look on graphical displays, participants should spend more time viewing task-relevant map areas (the closest pressure system to the arrow) after instruction compared to before instruction (Hypothesis 4). Consistent with this hypothesis, the 2 (before vs. after instruction) × 2 (map type) ANOVA showed that the proportion of looking time spent viewing the closest pressure system was greater after instruction ($M = .66$, $SD = .10$) than before instruction ($M = .49$, $SD = .11$), $F(1, 14) = 23.25$, $p < .001$, partial $\eta^2 = .62$. Neither map type, $F(1, 14) = 2.22$, $p = .16$, nor its interaction with instruction $F(1, 14) < 1$, had significant effects on this variable (see Table 2 for cell means).

Because the area of interest for the closest pressure system partially overlapped with that of the arrow, it is important to show that the increased time on this pressure system did not merely occur because people spent more time viewing the arrow after instruction. A 2 (before vs. after instruction) × 2 (map type) ANOVA showed that proportion of time spent viewing the arrow did not differ significantly as a function of map type, $F(1, 14) = 3.93$, $p = .07$, instruction, $F(1, 14) = 1.35$, $p = .27$, or their interaction, $F(1, 14) = 2.71$, $p = .12$ (see Table 2). Thus, the additional time spent viewing the pressure system after instruction was not due to additional time spent viewing the arrow.

[2] The patterns in the data are the same if amount of time rather than proportion is the dependent measure and all but one of the statistically significant effects for the proportion data were also statistically significant for absolute amount of looking time (in the other case the trend was marginally significant, $p = .07$).

[3] The analyses of performance focused on accuracy because response time was not significantly affected by instruction, map type, or their interaction, $F(1, 14) < 1$ in all cases. The mean response time across trials was 5.74 sec ($SD = 2.99$).

[4] In this context, the sensitivity or discriminability index, d', derived from signal detection theory (Wickens, 2002) is an estimate of the difference in participants' internal response to correct arrows versus incorrect arrows, independent of any bias they might have to respond "true" or false. Larger values of d' indicate that participants are more sensitive to the difference between correct and incorrect arrows.

Table 2
Means (and SD) of proportion of looking time on different regions of the maps in Experiment 1.

Region	Before instruction	After instruction
Pressure-plus-temperature map		
Closest pressure system	.47 (.14)	.64 (.12)
Arrow	.25 (.18)	.21 (.11)
All irrelevant regions	.36 (.14)	.23 (.11)
Temperature scale	.03 (.04)	.01 (.01)
Pressure-only map		
Closest pressure system	.51 (.13)	.68 (.11)
Arrow	.24 (.18)	.24 (.09)
All irrelevant regions	.36 (.16)	.21 (.09)
Temperature scale	n/a	n/a

Hypothesis 4 also predicts that participants should spend less time looking at task-irrelevant map areas of the display after instruction compared to before instruction. First, the proportion of time spent viewing all irrelevant areas of the map (i.e., areas besides the arrow and closest pressure system) was compared in a 2 (before vs. after instruction) × 2 (map type) ANOVA. This analysis showed that, consistent with Hypothesis 4, the proportion of looking time on these regions decreased significantly from .36 (SD = .14) before instruction to .22 (SD = .09) after instruction, $F(1, 14) = 16.54$, $p = .001$, partial $\eta^2 = .54$. Neither map type, nor the interaction of map type with instruction had significant effects, $F(1, 14) < 1$ on this variable (see Table 2).

Finally, we examined whether participants' inspection of the irrelevant temperature scale decreased from before to after instruction. The temperature scale was viewed on only a minority of trials (4.46, SD = 5.58 of the 30 pressure-plus-temperature map trials). Proportion of looking time spent on the temperature scale was very low ($M = .03$, SD = .04) before instruction, but even lower ($M = .007$, SD = .008) after instruction. Consistent with Hypothesis 4, participants viewed this scale more before instruction ($M = 3.26$ trials, SD = 4.50) than after instruction ($M = 1.20$ trials, SD = 1.37), $t(14) = 2.21$, $p < .04$, Cohen's $d = .60$.

In summary, as predicted (Hypothesis 4), the eye-fixation data indicated that after learning relevant meteorological principles, participants spent more time viewing the most task-relevant areas of a weather map and less time viewing task-irrelevant regions. In fact, after instruction they spent 22% of their time viewing the arrow and 66% of their time viewing the closest pressure system, although these features took up only 0.4% and 6.5% of the map area, respectively. This finding is strong evidence for top-down influences of task and domain knowledge on eye fixations in map comprehension.

Hypothesis 4 predicted that participants would spend more time viewing the closest pressure system to the arrow after instruction, but an unexpected result was that they also spent a large proportion of their time (49%) viewing this pressure system before instruction. In related research using the same paradigm with participants from the same population, debriefing revealed that about half of the participants knew that pressure was the relevant variable for predicting wind direction before instruction, although they did not know *how* to infer wind direction from pressure (Hegarty, Canham, & Kriz, 2006). The pressure system may also have been viewed more than chance due to its proximity to the arrow or its salience. Nevertheless, the comparison of time spent on the pressure system before and after instruction indicated a large effect of knowledge on viewing this relevant area (Cohen's $d = 1.24$).

3. Experiment 2

Contrary to our predictions (Hypotheses 2 and 3), map type did not have significant effects on performance of the inference task in Experiment 1. One possible reason for this null effect is that map type was manipulated within participants, so that viewing the pressure-only maps might have caused participants to consider only the pressure information on the more complex maps. Another possible reason is that although the pressure-only maps were simpler (presenting only task-relevant information), they were also less familiar and less ecologically valid, as lay people rarely see weather maps showing pressure alone. Lack of familiarity of the pressure-only maps might have cancelled out positive effects of their simplicity.

In Experiment 2 performance on the two types of weather maps used in Experiment 1 was contrasted in a between-participants design, such that each group of participants saw either pressure-only maps or pressure-plus-temperature maps. Assuming that the null effects of map type in Experiment 1 were due to the within-participants design in that experiment, Hypothesis 2 should be supported with the between participants design in Experiment 2. In contrast, if unfamiliarity of the pressure-only maps cancels out positive effects of their simplicity, we might expect equivalent performance on the pressure-only and temperature salient maps (Hypothesis 5). Finally, Experiment 2 offered an opportunity to replicate the effects of instruction seen in Experiment 1, such that people would perform better after instruction compared to before instruction (Hypothesis 1).

3.1. Method

3.1.1. Participants

The participants were 40 students (22 females and 18 males) from an introductory psychology class (age ranged from 18 to 22 years). None had any prior formal training in meteorology. Twenty students were assigned to each of the two experimental conditions (pressure-only maps condition and pressure-plus-temperature maps condition).

3.1.2. Materials and design

The weather maps were created using ESRI ArcMap from actual data from the National Oceanic and Atmospheric Administration (NOAA)'s hourly NCEP/NCAR reanalysis data composites for ten different dates in the last 10 years. They were designed to be similar to the pressure-plus-temperature and the pressure-only maps which were used in Experiment 1.

As in Experiment 1, each map contained a target arrow indicating a hypothetical wind direction. There were three versions of the maps (showing the correct arrow, opposite arrow, and pressure gradient arrow respectively) for each of the 10 dates for a total of 30 map-arrow trials, which were displayed both before and after instruction. The tutorial and Knowledge of Principles questionnaire were identical to those used in Experiment 1. In Experiment 2, a measure of background knowledge of meteorology, made up of 15 multiple-choice questions, was also included.

3.1.3. Procedure

Participants were tested in groups of up to ten, within one condition. They were first administered the measure of background knowledge of meteorology. Then they were familiarized with the weather maps and given instructions for the inference task. After six practice trials, the 30 experimental trials were displayed, one at a time, on a computer screen in the front of the room for 8 s each. On each trial, participants decided whether or not the arrow showed the correct wind direction and indicated their true–false judgment on paper using an answer sheet.

Then participants were shown the Powerpoint tutorial while the experimenter read the text aloud and participants had the opportunity to ask questions on each page. Next, they completed the Knowledge of Principles Questionnaire, and were given feedback on their answers. The tutorial was shown again, and all participants were asked to review the presentation to confirm their correct answers and to find the correct answer to any question they got wrong. Finally, participants were given three practice problems, received feedback on their answers, and completed the 30 wind inference trials.

3.2. Results and discussion

Participants in the two conditions did not differ significantly in background knowledge of meteorology ($M = 5.65$, $SD = 1.72$ out of 15 for the pressure-only condition; $M = 5.55$, $SD = 2.26$ for the pressure-plus-temperature condition), $t(38) = .16$, ns. After their first pass through the tutorial, students answered the four items of the knowledge of principles questionnaire. The mean number of correct answers was 3.35 ($SD = .81$) in the pressure-only condition and 2.85 ($SD = .88$) in the pressure-plus-temperature condition, $t(38) = 1.87$, $p = .07$.

3.2.1. Effects of knowledge and display on accuracy

Descriptive statistics for all the experimental conditions are shown in Table 3. The 2 (before vs. after instruction) × 2 (map type) × 3 (arrow) ANOVA showed that, as in Experiment 1, performance improved from before instruction ($M = .52$, $SD = .06$) to after instruction ($M = .71$, $SD = .11$), $F(1, 38) = 49.81$, $p < .001$, partial $\eta^2 = .57$ (supporting Hypothesis 1). A signal detection analysis indicated that sensitivity (d') increased from .01 before instruction to .95 after instruction for the pressure-plus-temperature maps, and from .36 to 1.62 for the pressure-only maps.

Table 3
Means (and SD) of accuracy (proportion correct) in Experiment 2 for the two maps.

Arrow	Before instruction	After instruction
Pressure-plus-temperature map		
Correct	.54 (.14)	.78 (.20)
Opposite	.58 (.21)	.85 (.20)
Pressure gradient	.35 (.22)	.29 (.27)
Pressure-only map		
Correct	.61 (.27)	.82 (.13)
Opposite	.59 (.28)	.86 (.20)
Pressure gradient	.45 (.26)	.65 (.29)

There was also a main effect of map type in this experiment. As predicted by Hypothesis 2, performance was better in general for participants who viewed the pressure-only maps ($M = .67$, $SD = .08$) than for those who viewed the pressure-plus-temperature maps ($M = .57$, $SD = .08$), $F(1, 38) = 11.40$, $p < .001$, partial $\eta^2 = .23$. This contrast remained significant when knowledge of the meteorological principles was entered as a covariate in the analysis, $F(1, 38) = 8.43$, $p < .01$, partial $\eta^2 = .19$, indicating that the maps affected participants' ability to apply their knowledge to the wind inference task, and not just how much knowledge they acquired from the instruction. The interaction of map type and instruction was not statistically significant, $F(1, 38) = 2.06, p = .16$. However, after-instruction accuracy was significantly higher for the pressure-only condition than for the pressure-plus-temperature condition when before-instruction accuracy was entered as a covariate in the analysis, $F(1, 37) = 6.14$, $p = .018$, partial $\eta^2 = .14$.

A secondary result indicated that (consistent with Experiment 1) there was a significant effect of the arrow to be verified, $F(2, 76) = 36.06$, $p < .001$, partial $\eta^2 = .49$, again indicating that people had difficulty rejecting the pressure gradient arrow. The type of arrow also interacted with map type, $F(2, 76) = 4.38$, $p = .01$, partial $\eta^2 = .11$. Analysis of simple effects indicated a significant effect of map type for the pressure gradient arrow, $F(1, 38) = 18.31$, $p < .001$, partial $\eta^2 = .33$, but not for the other two arrows, $F(1, 38) = 1.8$, $p > .19$ in both cases. A possible explanation for this finding is that there was less chance for map type to have a significant effect on performance for these arrows, because accuracy after instruction was very high for the correct and opposite errors arrows, even with the pressure-plus-temperature maps. Three of the 20 participants in the pressure-only condition and 12 of the 20 participants in the pressure-plus-temperature condition showed the pattern of responses that would be predicted if they applied only the pressure gradient principle, whereas 11 participants in the pressure-only condition and three participants in the pressure-plus-temperature condition showed good performance on all arrows. The remaining participants did not show either of these patterns.

In summary, the results of Experiment 2 indicate that when map type was manipulated between participants, Hypothesis 2 was supported, such that performance was relatively impaired for the maps that presented additional task-irrelevant information. In contrast, there was no evidence that the unfamiliarity of the pressure-only maps canceled out positive effects of their simplicity (Hypothesis 5). Finally, Experiment 2 replicated the result of Experiment 1 that performance of the wind inference task improves after instruction (again supporting Hypothesis 1).

4. General discussion

The two experiments presented in this article provide insight into how knowledge and display design affect the comprehension of complex graphics. In Experiment 1, participants spent most of their time viewing task-relevant information and spent a greater proportion of time viewing this information after a brief amount of instruction on relevant meteorological principles. These effects were accompanied by superior performance in making inferences from the weather maps, demonstrating that knowledge affects both processes of information selection from complex graphical displays, and processes of interpreting and making inferences from the selected information. In Experiment 2, eliminating task-irrelevant information in the display also improved performance, demonstrating that good display design can facilitate beginning students' comprehension of map displays.

The present research contributes to basic theories of graphics comprehension, which in turn can inform theories of how to develop graphic literacy and how to use graphical displays in instruction. Specifically, it demonstrates that declarative knowledge can influence what people attend to in a visual display, and consequently what information they encode. Of course, our results are not the first to show that domain knowledge can influence attention to visual-spatial displays. Several previous studies have shown that experts pay more attention to task-relevant aspects of such displays (Chase & Simon, 1973; Lowe, 1993, 1994, 1996; Reingold et al., 2001). However, expertise reflects 10 or more years of experience in a domain (Ericsson & Charness, 1994). What is novel about our results is the demonstration that attention to visual displays can change significantly with just a brief amount (10−15 min) of instruction that is conceptual rather than procedural in nature, that is, a type of instruction that is typical in geography and other science classrooms.

Previous research in the domain of meteorology has suggested that novices are drawn to perceptually salient features of weather maps whereas experts focus on what is thematically relevant in these displays (Lowe, 1993, 1994, 1996). These conclusions were based on indirect evidence from studies that used tasks such as copying weather maps and sorting of features from these maps. Experiment 1 provided more direct evidence from eye fixations that with knowledge, people attend more to thematically relevant information and less to irrelevant information in weather maps.

The increase in participants' attention to relevant map locations and the corresponding decrease in attention to irrelevant map areas are consistent with the information reduction hypothesis. Whereas Haider and Frensch (1996, 1999) originally proposed this hypothesis in the context of skill acquisition and tested it with a laboratory task, the present research indicated that this hypothesis can be extended to the more ecologically valid task of weather map comprehension and to situations that are more typical of classroom instruction.

Several theorists have proposed principles for the construction of effective graphics (Bertin, 1983; Kosslyn, 1989; Tufte, 1983), but there has been little empirical validation of these principles. The present research provided new evidence for the basic principle that graphics designed for use by beginning students should not provide more information than is needed by the student for the current task. In Experiment 2, performance was less accurate with maps showing temperature and geographic information in addition to the task-relevant pressure information. We have replicated these results in several experiments in which we varied the salience of irrelevant information across different maps rather than its presence (Hegarty et al., 2006) indicating that they are not isolated results.

The results of the present study prompt us to ask *how* extraneous information in a complex visual display like this impairs performance. There was a non-significant trend for the pressure-only group in Experiment 2 to have better understanding of the meteorological principles after instruction, despite equivalent background knowledge of meteorology at the beginning of the experiment. However, the effect of map type on the inference task in this experiment was independent of this trend, suggesting that the main locus of this effect is in the process of applying declarative knowledge to make an inference. One possibility is that the task-irrelevant information masked the relevant pressure information, making it harder to judge the angle of the arrow relative to the closest pressure system, which is the critical information for verifying wind direction. The irrelevant information might also have drawn attention away from the task-relevant information. Interestingly, in Experiment 1, in which participants saw both types of maps in a within-participants design, there were no measurable differences in eye fixations to task-relevant information or in performance. This suggests that the effects observed here and elsewhere (Hegarty et al., 2007) may be short-lived and remediable by showing students maps with only task-relevant information as well as maps with extraneous information.

There was no evidence in Experiment 2 that domain knowledge made participants more "immune" to the extraneous information in the display. In fact, the display effects were, if anything, greater after instruction (see Table 3). On first glance, this seems to contradict the idea that people with less knowledge should be more influenced by what is perceptually salient in a display (Lowe, 1993, 1994, 1996). However, it is important to realize that our participants were at a very early novice stage. In fact, even after instruction they had difficulty applying their knowledge to infer the correct wind direction, and the answer choices of several participants indicated that they did not fully understand the Coriolis phenomenon. The question of how to teach meteorological principles was not the focus of this current research, but our results point to this as an important issue for future research. Another important future goal of our research is to examine whether expert meteorologists are also affected by the presence of task-irrelevant variables in meteorological displays.

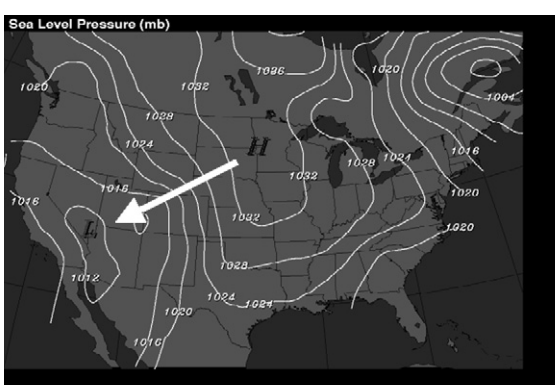
Previous research on graphics comprehension has focused on relatively simple displays (Carpenter & Shah, 1998; Körner, 2005; Lohse, 1993; Simkin & Hastie, 1986). One difference between the complex graphics studied here and the displays used in most previous research is that they present a lot of information, only some of which is relevant for a given comprehension task. By studying comprehension of such complex graphics, we have highlighted the importance of selection of task-relevant information as a critical step in graphics comprehension, and shown how both knowledge and good display design can facilitate this process.

Acknowledgements

This research was funded by grants N000140310119 and N000140610163 from the Office of Naval Research, USA. We thank Jerome Tietz for assistance with the eye tracking analysis, Shannon Thiele for help with data collection and coding, Joel Michaelson for providing meteorological expertise, and Priti Shah for comments on an earlier version of the manuscript.

Appendix A. Explanation of the meteorological principles from the instructional presentation (reprinted by permission of Weather World 2010 Project)

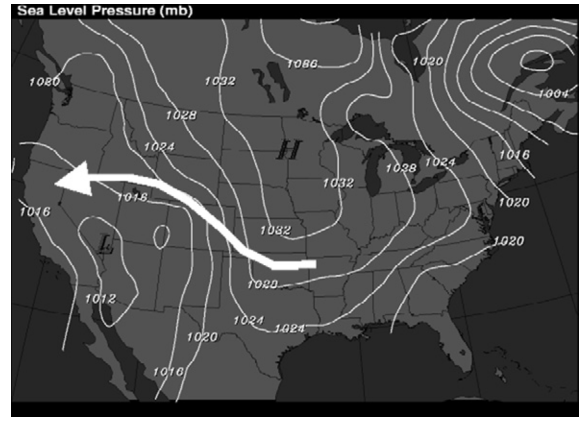
Air pressure is fundamental to how the winds blow. The pressure gradient force is the force that causes winds to blow. When there is a difference in the pressure between two different regions, the pressure gradient force pushes air from regions of high pressure to regions of low pressure in an attempt to balance the air pressure.

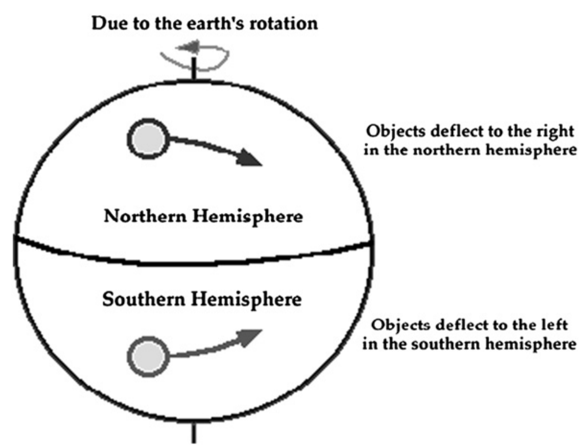


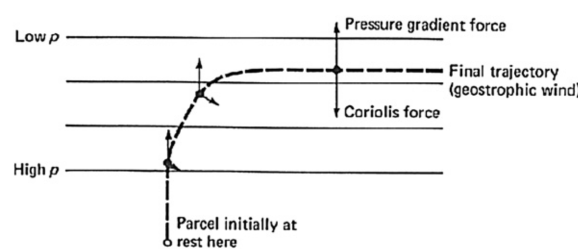
If the pressure gradient force were the only force acting upon air, winds would always blow directly from higher to lower pressure as indicated by the arrow in the map above.

However, pressure gradient force isn't the only factor to be considered. The Coriolis effect causes wind to be deflected to the right, as the arrow shows above.

The Coriolis effect occurs because of the rotation of the earth around its axis. As wind moves from high to low pressure regions, the rotation of the earth creates an apparent deflection of the wind. The Coriolis effect causes wind to



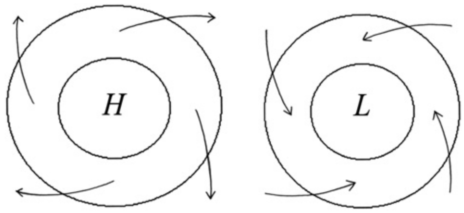




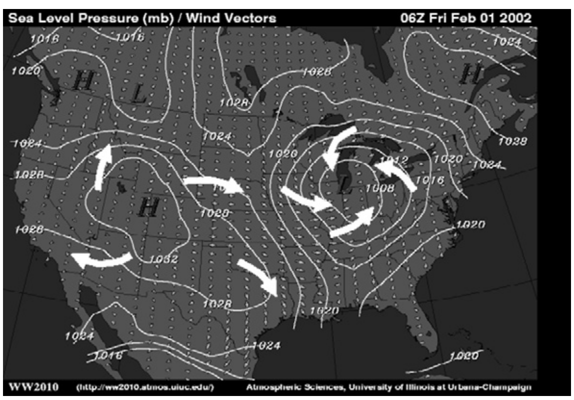
deflect to the right of its path in the Northern Hemisphere and to the left of its path in the Southern Hemisphere.

The Coriolis effect always acts at a right angle to the wind, changing the direction of the wind but not its speed. In the diagram above, the wind begins to move from high to low pressure areas, perpendicular to the lines of pressure (isobars). The Coriolis effect constantly pulls the wind direction at a right angle to the current wind direction causing the wind to bend right until the wind is moving parallel to the isobars. At this point, the pressure gradient force balances the Coriolis effect.

If pressure gradient force and the Coriolis effect were the only factors determining wind direction, wind would always flow parallel to isobars. However, near the earth's surface and



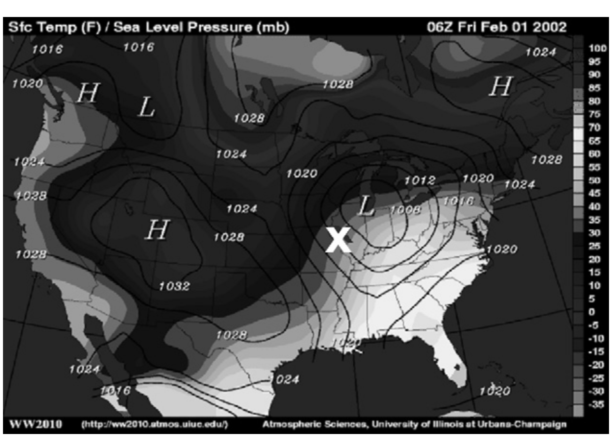
sea-level, friction reduces the wind speed, reducing the Coriolis effect but not pressure gradient force. Therefore, near the earth's surface, pressure gradient force will remain stronger than the Coriolis effect, resulting in winds blowing counterclockwise and into a low pressure area, called a cyclone, and clockwise and out of a high pressure area, called an anticyclone.



As indicated by the arrows, the wind is moving clockwise and out of the high pressure area (H) and counterclockwise and into the low pressure area (L). The large arrows have been added to make the wind direction more visible.

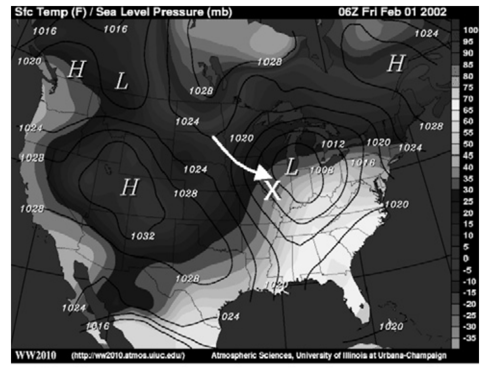
Appendix B. Example of a worked example from the instructional presentation

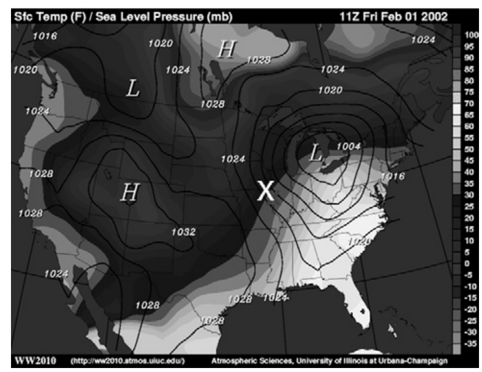
Predicting weather change: example 1



Will the area indicated by the X get *COLDER* or *WARMER*?

Answer and explanation





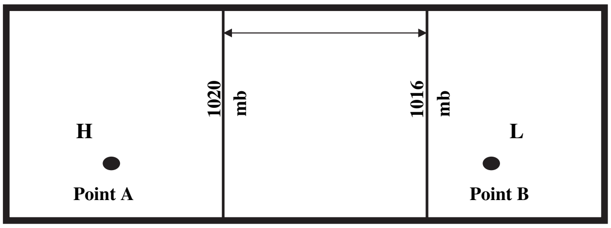
The correct answer was: COLDER. The map on the right shows the weather 5 h after the map on the left. You can see that the temperature at the area marked with the X got colder as the wind moved counterclockwise and into the low pressure area (as indicated by the white arrow).

Appendix C. The knowledge of principles questionnaire

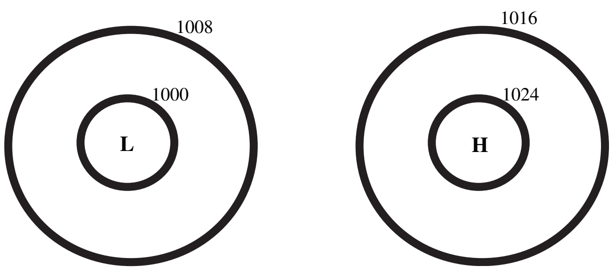
Please answer the following

1. In which direction does wind blow?
   a. From an area of low pressure to an area of high pressure.
   b. From an area of high pressure to an area of low pressure.
   c. In both directions.
   d. It depends on other weather conditions, so it can be either way.
2. How does air circulate around pressure systems in the Northern hemisphere?
   a. Counterclockwise around areas of high pressure, clockwise around areas of low pressure.
   b. Counterclockwise around areas of low pressure, clockwise around areas of high pressure.
   c. Counterclockwise around areas of high and low pressure.
   d. It depends on other weather conditions, so it can be any way.

3. Please draw arrows to indicate the wind direction between Point A and Point B.



4. Please draw arrows to indicate wind directions.



Scoring

Question 1. The correct answer is b.
Question 2. The correct answer is b.
Question 3. This question was scored as correct if the participant drew an arrow from high pressure system to low pressure. A participant who drew a straight arrow from the high to the low pressure was given the following feedback: ‘‘It is correct that the wind flows from high to low pressure, but because of the Coriolis effect the wind doesn’t travel in a straight line’’. The experimenter then motioned the correct, curved direction of the wind.
Question 4. This question was scored as correct if the participant drew arrows that were counterclockwise and inward for the low pressure system and clockwise and outward for the high pressure system.

References

Ahrens, C. D. (2000). *Meteorology today: An introduction to weather, climate, and the environment.* Pacific Grove, CA: Brooks/Cole.

Ainsworth, S. (2006). DeFT: a conceptual framework for considering learning with multiple representations. *Learning and Instruction, 16*, 183–198.

Bertin, J. (1983). *Semiology of graphics (W. Berg, Trans.).* Madison, WI: University of Wisconsin Press.

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

Chase, W. G., & Simon, H. A. (1973). Perception in chess. *Cognitive Psychology, 4*, 55–81.

Cheng, P. C.-H. (1999). Interactive law encoding diagrams for learning and instruction. *Learning and Instruction, 9*. 309–309.

Cohen, J., Cohen, P., West, S. G., & Aiken, L. S. (2003). *Applied multiple regression/correlation analysis for the behavioral sciences* (3rd ed.). Mahwah, NJ: Erlbaum.

Ericsson, K. A., & Charness, N. (1994). Expert performance: its structure and acquisition: reply. *American Psychologist, 49*, 725–747.

Freedman, E. G., & Shah, P. (2002). Toward a model of knowledge-based graph comprehension. In M. Hegarty, B. Meyer, & N. H. Narayanan (Eds.), *Diagrammatic representation and inference* (pp. 8–31). Berlin: Springer.

Haider, H., & Frensch, P. A. (1996). The role of information reduction in skill acquisition. *Cognitive Psychology, 30*, 304–337.

Haider, H., & Frensch, P. A. (1999). Eye movement during skill acquisition: more evidence for the information-reduction hypothesis. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 2*, 172–190.

Hegarty, M., Canham, M.S., Kriz, S. (2006, November). Interactions between knowledge and display design in interpretation of complex graphics. Paper presented at the Annual Meeting of the Psychonomic Society, Houston, TX.

Henderson, J. M. (2003). Human gaze control during real-world scene perception. *Trends in Cognitive Sciences, 7*, 498–504.

Henderson, J. M., & Ferreira, F. (2004). Scene perception for psycholinguists. In J. M. Henderson, & F. Ferreira (Eds.), *The interface of language, vision and action: Eye movements and the visual world* (pp. 1–58). New York: Psychology Press.

Hoffman, R., Detweiler, M., Conway, J., & Lipton, K. (1993). Some considerations in using color in meteorological displays. *Weather and Forecasting, 8*, 505–518.

Jarodzka, H., Scheiter, K., Gerjets, P., & Van Gog, T. (2010). In the eyes of the beholder: how experts and novices interpret dynamic stimuli. *Learning and Instruction, 20*(2), 146–154.

Körner, C. (2005). Concepts and misconceptions in comprehension of hierarchical graphs. *Learning and Instruction, 15*, 281–296.

Kosslyn, S. (1989). Understanding charts and graphs. *Applied Cognitive Psychology, 3*, 185–225.

Lohse, G. L. (1993). A cognitive model for understanding graphical perception. *Human Computer Interaction, 8*, 353–388.

Lowe, R. K. (1993). Constructing a mental representation from an abstract technical diagram. *Learning and Instruction, 3*, 157–179.

Lowe, R. K. (1994). Selectivity in diagrams: reading beyond the lines. *Educational Psychology, 14*, 467–491.

Lowe, R. K. (1996). Background knowledge and the construction of a situational representation from a diagram. *European Journal of Psychology of Education, 11*, 377–397.

Lowe, R. (2004). Interrogation of a dynamic visualization during learning. *Learning and Instruction, 14*. 257–257.

Myles-Worsley, M., Johnston, W. A., & Simons, M. A. (1988). The influence of expertise on X-ray image processing. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 14*, 553–557.

Peebles, D., & Cheng, P. C.-H. (2003). Modeling the effect of task and graphical representation on response latency in a graph-reading task. *Human Factors, 45*, 28–46.

Pinker, S. (1990). A theory of graph comprehension. In R. Freedle (Ed.), *Artificial intelligence and the future of testing* (pp. 73–126). Hillsdale, NJ: Erlbaum.

Reingold, E. M., Charness, N., Pomplun, M., & Stampe, D. M. (2001). Visual span in expert chess players: evidence from eye movement. *Psychological Science, 12*, 48–55.

Sanchez, C. A., & Wiley, J. (2006). An examination of the seductive details effect in terms of working memory capacity. *Memory and Cognition, 34*, 344–355.

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

Shah, P., & Carpenter, P. (1995). Conceptual limitations in comprehending line graphs. *Journal of Experimental Psychology: General, 124*, 337–370.

Shah, P., Freedman, E. G., & Vekiri, I. (2005). The comprehension of quantitative information in graphical displays. In P. Shah, & A. Miyake (Eds.), *The Cambridge handbook of visual spatial thinking* (pp. 426–476). New York: Cambridge University Press.

Shah, P., Mayer, R. E., & Hegarty, M. (1999). Graphs as aids to knowledge construction. *Journal of Educational Psychology, 91*, 690–702.

Simkin, D. K., & Hastie, R. (1986). An information-processing analysis of graph perception. *Journal of the American Statistical Association, 82*, 454–465.

Sweller, J., & Chandler, P. (1994). Why some material is difficult to learn. *Cognition and Instruction, 12*. 185–185.

Trafton, J. G., & Trickett, S. (2001). A new model of graph and visualization usage. In J. D. Moore, & K. Stenning (Eds.), *Proceedings of the 23rd Annual Conference of the Cognitive Science Society* (pp. 1048–1053). Mahwah, NJ: Erlbaum.

Trickett, S. B., & Trafton, J. G. (2006). Toward a comprehensive model of graph comprehension: making the case for spatial cognition. In D. Barker-Plummer, R. Cox, & N. Swaboda (Eds.), *Diagrammatic representation and inference* (pp. 286–300). Berlin: Springer.

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

Weather World 2010 Project. (2000). Images retrieved October 2004 from http://ww2010.atmos.uiuc.edu/(Gh)/wx/surface.rxml.

Wickens, T. D. (2002). *Elementary signal detection theory*. New York, NY: Oxford University Press.