Explanation Reconstruction through Reinterpretation of Key Facts

Hitoshi Terai (terai@is.nagoya-u.ac.jp)
Kazuhisa Miwa (miwa@is.nagoya-u.ac.jp)
Shota Matsubayashi (s.matubayashi@cog.human.nagoya-u.ac.jp)
Graduate School of Information Science, Nagoya University, Furo-cho, Chikusa-ku,
Nagoya, Aichi, 464-8601, Japan

Abstract

Reconstructing explanations is crucial for the progress of science. We focused on the transition of interest in a key fact that contradicts the preceding explanation and has a central role in its reconstruction. We used a short story as an experimental material in which the participants first constructed a naïve explanation and reconstructed it. First, when the naïve explanation was rejected, a new explanation was required, after interest in the key fact was inhibited. Second, hypothesized premises not inconsistent with the naïve explanation were sought to protect the naïve explanation. Third, interest in the key fact was recovered through the process of the explanation reconstruction. Last, we facilitated the explanation reconstruction by having the participants focus on the key fact.

Keywords: eye movement analysis; scientific explanation; key fact; insight; naïve concept.

Introduction

Scientific activities aim to understand the world by two ways: descriptive and explanative (Simon, 2000). Descriptive understanding grasps the nature and characteristics of phenomena by observations and experiments; explanative understanding grasp the mechanisms behind the phenomena and the reasons why such phenomena appear.

Through the history of science, descriptive understanding is usually established first and then explanative understanding is investigated. For example, Keplar's law described the orbit of the planets, and then Newton's law explained why they moved in such orbits. Science has developed while pursuing such explanative understanding about phenomena. The construction of explanations is crucial for science.

As historical facts, we can confirm many cases where the explanation for a certain phenomenon was completely changed because the structures of the explanation and the concepts of objects were essentially shifted. Such cases are generally observed in the history of science: e.g., the shift from the caloric theory to the oxygen theory and the transition from Newton's traditional theory to Einstein's relative theory.

As an example, consider the change of the caloric theory to the oxygen theory. Initially, in the caloric theory, burning was explained as the release of caloric. After an inconsistency was observed about the caloric theory, the weight increase after burning, a new explanation was required. In the current study, we call such an instance that contradicts a preceding explanation and must be interpreted by a new explanation a "key fact." The change of the explanation from the caloric theory, i.e., burning released caloric, to the oxygen theory, i.e., burning is connected with oxygen, was established by reinterpreting the key fact, i.e., the increase of weight by burning.

Note that there are two ways of understanding a key fact. One is by local modification and slight expansion of a previous explanation. The other is understanding by an essential change of a previous explanation. An interpretation about key facts completely changes between the two types of understanding: a completely different interpretation about a key fact is given in each of the two theories. In the oxygen theory, the key fact is explained by the connection with oxygen, but in the caloric theory, it is explained by the release of phlogiston that has negative weight.

Difficulties exist in such essential reconstruction of explanations. Interesting processes are often observed that prevent such reconstruction. One is the stubborn refusal to abandon explanations. People prefer to protect an established explanation by modifying and adding new reservations than shifting to a new one. In such a case, to protect the old explanation, people may focus on other irrelevant facts and arbitrarily proposed premises that are not inconsistent with the previous explanation. They sometimes add secondary explanations as protection. For example, in the caloric theory, a premise, phlogiston might have negative weight, was hypothesized and investigated.

Now we summarize the problems we address in this paper. We investigate a situation in which fact F that cannot be interpreted by explanation A is observed, and new explanation B is required. For the transition from explanation A to B, a mental leap is needed, meaning that fact F must be reinterpreted. We call fact F a key fact and investigate how it is processed through a reconstruction of the explanation.

The research questions and hypotheses are drawn in the following. We hypothesize that reinterpretation of key facts is crucial for the transition to a new explanation. However, people tend to pursue unrelated facts or arbitrarily hypothesized premises to protect old explanations, and such reinterpretation of key facts may be postponed. As a result, the interest in key facts is inhibited, and the reconstruction of explanation is impeded.

We propose two hypotheses:

Hypothesis 1 When a previous explanation is rejected and a new explanation is required, interest in a key fact may be temporarily inhibited.

Hypothesis 2 With the inhibition of interest in a key fact, other facts and hypothesized premises that are not inconsistent with a previous explanation are searched for, and

Introduction	Taro was driving to Las Vegas by rental car. His car broke down in a small town. He decided to get a haircut while the car was repaired. There are only two barbershops in the town: Alf's shop and Bally's shop. He is considering which to select.
Filler (Place)	Alf's shop is on the ground floor of a building located in the east area of town. In the building, there is a stationary shop. Bally's shop is along a street running in the west area of town. There is a supermarket near it.
Key Fact	Alf's hair is unkept, and the nape of his neck is messy. Bally's hair is beautifully cut, and the nape of his neck is neat.
Filler (Barbershop)	The windows of Alf's shop are light blue, and natural scene are pictured on the cover pages of the books in the shop. The windows of Bally's shop are light green, and various letters are written on the cover pages of the books in the shop.
Filler (Time)	Alf's shop is open until late. He often eats dinner at his favorite restaurant near the shop. Bally's shop is open early. He usually walks around the shop in the morning.

Figure 1: Barber task (used in Experiment 1).

the old explanation is protected.

We are also interested in the transition process from a former to a new explanation. We confirm that the reinterpretation of a key fact has a central role in the transition of the explanation. Two additional hypotheses are drawn.

Hypothesis 3 The recovery of interest in a key fact is observed through the process of the explanation reconstruction.

Hypothesis 4 We may promote explanation reconstruction by having participants focus on a key fact.

Concern for a key fact

We used a short story as an experimental material because people understand a text by unifying meanings while adding implicit information and inferences about omitted and unpresented sentences (Rumelhart & Ortony, 1977; Seifert & Robertson, 1985).

We used a text that was modified from a barber task (Gardner, 1978). Figure 1 shows the material. In this material, (1) a naïve explanation is initially constructed by a key fact; (2) a new explanation is required where a shift of perspective is needed; (3) a rational explanation is constructed by reinterpreting the key fact.

In the story, in a town with only two barbershops, a character is looking for a barber and must to select either barber A whose staff has unkept hair or barber B whose staff's hair is beautifully cut. Initially, participants may select barber B where the following naïve explanation is given: "a barber with beautiful hair is very skilled." However, a new explanation is required after being informed that the character selected barber A. The fact, "barber A's hair is messy and barber B's hair is neat," contradicts the initial explanation. Therefore, in the story, the key fact that must be reinterpreted is: "barber A's staff has unkept hair and barber B's staff has neat hair." The reconstructed explanation from Gardner (1978) is: "each does the other's hair because there are only two shops in town; therefore, barber A's staff who did barber B's staff's hair is more skilled." In the reconstructed explanation, the key fact becomes evidence for selecting barber A but in the naïve explanation, it is evidence for selecting barber B. The meaning of the key fact has completely shifted with the transition from the initial to the reconstructed explanation.

In the text, other unrelated facts than the key fact are described; therefore, other secondary additional explanations may be possible to protect the initial naïve explanation. However, if participants construct such an explanation based on other facts than the key facts, the contradiction remains unsolved: the character selects a messy barber. The shift to the reconstructed explanation by reinterpreting the key fact is required for consistently understanding the story structure.

Experiment 1

We confirmed the validity of the barber's story as an experimental task to examine our hypotheses. Experiment 1 confirmed whether most participants initially constructed the naïve explanation. Additionally, to confirm Hypothesis 2 preliminary, participants reconstructed their explanations after the naïve explanation was rejected.

The definitions of the naïve and reconstructed explanations are described below.

Naïve explanation A barber who has beautiful hair is very skilled.

Reconstructed explanation Each does the other's hair because there are only two shops in the town; therefore, barber A's staff who did barber B's staff's neat hair is more skilled.

Subjects

Fifty-three undergraduate students participated in Experiment 1.

Task

Figure 1 shows the barber task used in Experiment 1.

Procedure

Experiment 1 was constructed of two phases: the initial explanation phase and the reconstruction phase. In the initial explanation phase, the participants read the story while thinking about which barbershop to select and their task was to construct an explanation for their decision. The initial explanation phase was followed by the reconstruction phase in which the naïve explanation was rejected, and they were required to reconstruct their explanation of the story.

Results

Forty four of the 53 participants initially constructed the naïve explanation. A binomial test revealed that they primarily constructed naïve explanations (two-sided: $p < .01$). Moreover, the reconstructed explanations by the 44 participants were classified into four types (Table 1). Three other explanations than the reconstructed explanation were based on such facts about place as "shop B is located near the repair shop," and about time such as "shop B is open until late," and hypothetical information not included in the story. A chi-square test revealed a significant difference in the numbers of these explanations ($\chi^2(3) = 10.8, p < .05$), and a multiple comparison using Ryan's method showed that the explanations based on place and time were constructed significantly more than the reconstructed explanation ($p < .01, p < .01$).

Table 1: Produced second explanations.

Explanation	#	Example of description
Target (reconstructed explanation)	3	Each does the other's hair because there are only two shops in the town; therefore, Bally's staff who did barber Alf's beautiful staff's hair is more skilled.
Place	15	Alf's shop was near the car repair shop.
Time	18	Taro wanted to get a haircut late in the evening.
Misc	11	-

Some descriptions were classified into multiple categories because they included multiple facts.

These results confirmed the validity of the barber task as an experimental task for our study. Additionally, we confirmed that the participants tended to add secondary explanations based on the facts about place and time to protect the naïve explanation, preliminarily supporting Hypothesis 2.

Experiment 2

We confirmed both the inhibition of interest in the key fact (Hypothesis 1) and resumption of interest (Hypothesis 3) using eye movement analysis to capture the transition of interest.

Subjects

Twenty-one undergraduate students participated in the experiment.

Task

The story was displayed on a computer screen. The filler (barbershop) part of the text (Figure 1) was removed due to limitations of the display size.

Procedure

The experiment was conducted individually, and participant eye movements were recorded using a Tobii T60 eye tracker.

As in Experiment 1, in the initial explanation phase, the participants were required to explain the story. Their fixation ratios of the key fact during the initial explanation phase were used as the baseline for analysis of the subsequent reconstruction phase. The fixation ratio of each fact was normalized by the number of letters that were included in each part.

After the initial explanation phase, the participants were reconstructed their explanations as answers to a quiz. When the participants found an idea, they reported it. They did the experiment at their own pace. When they gave another explanation than the reconstructed explanation as the target, they were told that it was not correct and were told to reconsider. This phase was continued for 30 minutes; when each participant constructed the reconstructed explanation, it was terminated.

Results

Inhibition of fixation on key fact To examine the inhibition of interest in the key fact after the naïve explanation was rejected, we analyzed the fixation ratio of the key fact part in the initial stage of the reconstruction phase. First, we examined whether the fixation ratio of the key fact was less than the baseline obtained in the initial explanation phase and the other facts (place and time).

Figure 2 shows the fixation ratio of each of the facts (place, key, and time) during the first minute and the subsequent minute in the reconstruction phase.

In the first minute (0–60 sec), a t-test indicated no significant differences between the fixation ratio of the key fact and the baseline ($t(20) = 1.62, n.s.$). A one-way ANOVA showed a significant main effect of the three facts (place, key, and time) ($F(2,40) = 8.97, p < .001$), and a multiple comparison using Ryan's method showed that the fixation ratio of the fact about place was significantly higher than those of the key fact and the fact about time ($p < .05, p < .05$).

Next, we conducted the same analysis on the subsequent minute (60–120 sec). A t-test indicated a significant difference between the fixation ratio of the key fact and the baseline ($t(20) = 2.86, p < .01$). A one-way ANOVA showed a significant main effect of the three facts (place, key, and time) ($F(2,40) = 4.36, p < .05$), and a multiple comparison

using Ryan's method showed that the fixation ratio of the key fact was significantly lower than the place and time facts ($p < .05, p < .05$).

In the subsequent minute, we confirmed that the fixation ratio of the key fact was significantly lower than the baseline and the place and time facts. These results support that interest in the key fact was inhibited when the naïve explanation was rejected.

On the other hand, in the first minute, our prediction was not observed; the fixation ratio of the fact about place was substantially higher. This result might be affected by the order of the three facts, which were arranged as place, key, and time (see Figure 1). The participants probably read the story in this order, reflecting the result in the first minute.

The overall results of Experiment 2 supported Hypothesis 1.

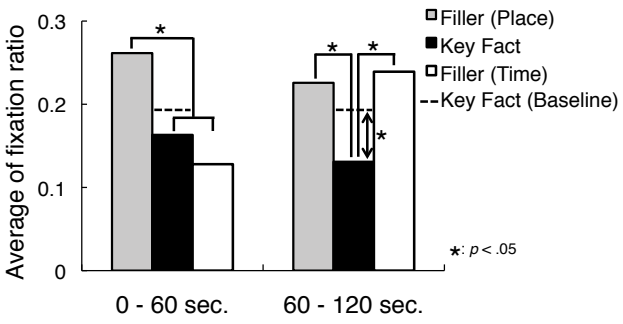

Figure 2: Transition of fixation ratio of each fact.

Resumption of fixation on key fact Next, we analyzed the process of recovering interest in the key fact to build reconstructed explanation (Hypothesis 3). We focused on the 11 of 21 participants who successfully constructed the reconstructed explanation and analyzed the transition process of their fixation ratio of the key fact part.

Figure 3 shows the transition with the progress of three phases: the first 60 seconds, the last 60 seconds before reaching the reconstructed explanation, and the residual between them.

A one-way ANOVA showed an significant main effect of the three phases ($F(2,20) = 7.62, p < .005$), and a multiple comparison using Ryan's method showed that the fixation ratio of the key fact in the last phase was significantly higher than those in the first and middle phases ($p < .05, p < .05$). The interest in the key fact gradually improved even though the ratio in the middle phase was not greater than in the first phase, partially supporting Hypothesis 3.

Experiment 3

In Experiment 2, we observed the inhibition of interest in the key fact after the naïve explanation was rejected (Hypothesis 1). This means that the participants searched for unimportant facts that did not contradict the naïve explanation.

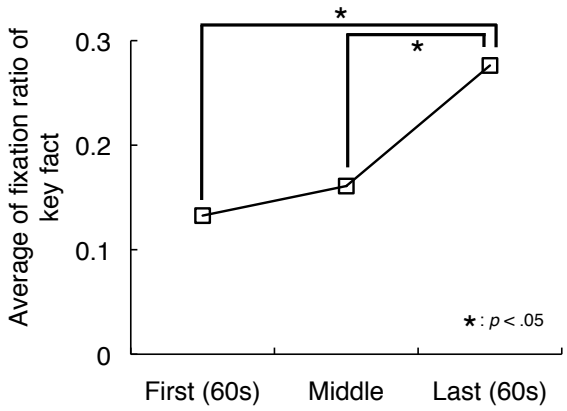

Figure 3: Transition of fixation ratio of key fact in successful group.

In Experiment 3, we examined whether forced recovery of interest in the key fact could facilitate the reconstruction of their explanations (Hypothesis 4) by externally controlling interest in the key fact.

Subjects

Forty-one undergraduate students participated in this experiment.

Procedure

This experiment was conducted in small groups on personal computers for the stimulus presentation and data acquisition.

The participants were divided into two experimental conditions: highlighted (21 participants) and non-highlighted (20 participants). In the highlighted condition, the key fact was colored red to facilitate interest in it; the participants were instructed that the highlighted sentences were crucial for finding the right explanation. There was no such highlight in the non-highlighted condition.

In the initial explanation phase, the participants constructed explanations about the story as in Experiments 1 and 2. Then, in the reconstruction phase, they were also required to reconstruct their explanations and report them by computer keyboard. After reporting their explanations, they received a message: "since there is another reasonable explanation acceptable to all, please reconsider." They were told that there was an evaluator in another room connected by the Internet, even thought no such evaluator existed, and the same message was always returned. The maximum time of the reconstruction phase was 15 minutes, and the data were analyzed until they reached reconstructed explanations.

Results

The explanations that were constructed in the initial explanation phase were mostly naïve explanations (19 of 21 in the

highlighted condition and 18 of 20 in the non-highlighted condition).

Next, we analyzed the facts to which the participants referred until they reached the reconstructed explanation in the reconstruction phase. The referenced facts were identified by their description about the explanations. The ratio of each of the referenced facts in the generated explanations is shown in Figure 4.

A two-way ANOVA was conducted with the experimental conditions (highlighted and non-highlighted) as a between-participant factor and the facts (place, key, and time) as a within-participant factor. There was neither a significant main effect of the experimental conditions nor of the facts ($F(1,33) = .16, n.s.; F(2,66) = 1.15, n.s.$), but there was significant interaction between the conditions and the facts ($F(2,66) = 8.55, p < .001$). A multiple comparison using Ryan's method showed differences between the highlighted and non-highlighted conditions of the key and place facts ($p < .05, p < .05$). There were also significant differences between the key and place facts in the highlighted condition ($p < .05$), and the key fact and the place and time facts in the non-highlighted condition ($p < .05, p < .05$).

These results suggest that the participants in the highlighted condition attempted to reinterpret the key fact. On the other hand, in the non-highlighted condition, there was little mention of the key fact. When there was no facilitation of interest in the key fact, the participants tended to modify their explanation based on other facts, supporting Hypothesis 2 that was preliminary supported in Experiment 1.

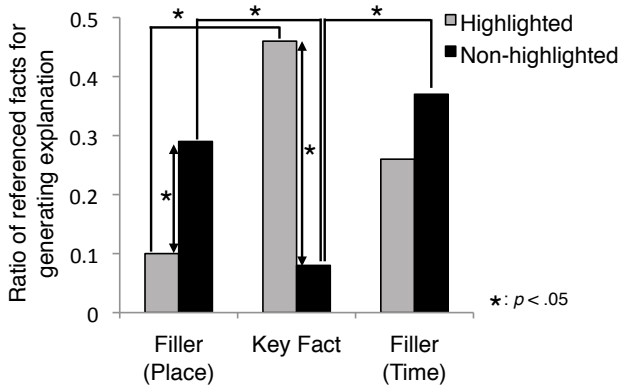

Figure 4: Ratio of referenced facts for generating explanation.

Finally, the ratios of the accumulated number of the participants who reached the reconstructed explanation in the reconstruction phase are shown in Figure 5. A chi-square test showed significant differences between the two experimental conditions at 10 and 15 minutes ($\chi^2(1) = 6.26, p < .05; \chi^2(1) = 7.79, p < .01$).

These results indicate that enhancing interest in the key fact facilitated the reconstruction of the naïve explanation, and shifting to the reconstructed explanation supporting Hypothesis 4.

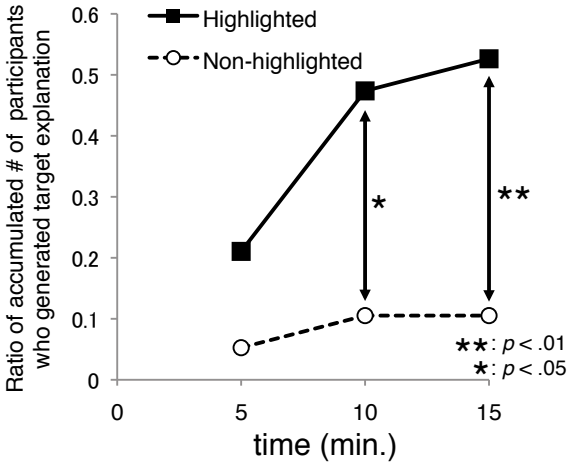

Figure 5: Ratio of accumulated number of participants who generated consistent explanations.

Discussions and Conclusions

In our study, we focused on reinterpretation of a key fact for explanation reconstructions and examined the inhibition of interest in it and the improvement of interest in it by eye-movement analysis. The following is a summary of our experiment.

- When the naïve explanation was rejected by contradictions with the key fact and a new explanation was required, interest in the key fact was inhibited (Hypothesis 1 was supported).
- In such a situation, other facts and hypothesized premises not inconsistent with the naïve explanation were searched for and focused on, and the naïve explanation was protected(Hypothesis 2 was supported).
- Interest in the key fact was recovered through the process of the explanation reconstruction, especially before reaching the solution (Hypothesis 3 was partially supported).
- We facilitated the explanation reconstruction by having the participants focus on the key fact (Hypothesis 4 was supported).

In studies of hypothesis generation and testing, some human biases have been confirmed. For example, confirmation bias gathers positive instances to confirm hypotheses (Wason, 1960), and established hypotheses are maintained against anomalous data (Chinn & Brewer, 1998; Mason, 2001).

In the educational psychology domain, it has also been confirmed that naïve concepts that are not supported by related evidences are held strongly (McCloskey, Washburn, & Felch, 1983; Watts & Zylberstan, 1981). Watts and Zylberstan (1981) studied junior high-school students' naïve concepts about the inertia law and concluded that even if they received accurate knowledge about the law from lectures, they repeated the naïve explanation for events related to it.

In our experiments, we set up an experimental situation in which the participants reconstructed the naïve explanation. Theories in scientific activities were established based on the accumulated results of experiments through history in which the fixation to an explanation may be much stronger. Note that similar phenomena were observed using a short story in the laboratory setting of our current study.

The situation requiring a shift in explanations in our experiments seems similar to the settings dealt with in studies of insight problem solving. Here, mental constraints arising from the perceptual features of a problem and past experiences create an impasse and prevent problem solvers from finding the new relations required to solve the problem. These mental constraints are gradually relaxed unconsciously in some cases even if problem solvers meet an impasse, where they often ignore key evidence that leads to a solution. With activities that do not follow these mental constraints and acceptance of such crucial instances for solutions, problem solvers gradually reach a solution (Knoblich, Ohlsson, Haider, & Rhenius, 1999; Knoblich, Ohlsson, & Raney, 2001; Ohlsson, 1992; Terai & Miwa, 2003). In our experiment, we also observed such a recovery process that of focused on key facts.

It is difficult to manage interest in facts without the influence of hypotheses and concepts that were previously constructed (Kaplan & Simon, 1990; Luchins & Luchins, 1950; Wason, 1960). For example, Bilalic, McLeod, and Gobet (2008) used a rule discovery task that required participants to search for a better solution than the one they had already found. Their study indicated that the participants were unconsciously prevented from searching for facts that were unrelated to the existing solution, even if they were instructed to seek alternatives. This result also indicates that human behavior is largely constrained by constructed hypotheses and concepts.

In Experiment 3 of our study, even though we controlled the participant interest in the key fact that contradicted the naïve explanation by highlighting it, there was no significant difference between the highlighted and non-highlighted conditions during the first five minutes. Moreover, only half of the participants constructed the reconstructed explanation, even if such external stimuli were given. This suggests that the inhibition of interest in the key fact that contradicted the naïve explanation remained even after deep consideration for 15 minutes. We must study the interaction between conscious and unconscious activities by combining verbal protocols and eye movement analysis to understand the process of shifting explanations in more detail.

References

Bilalic, M., McLeod, P., & Gobet, F. (2008). Why good thoughts block better ones: The mechanism of the pernicious einstellung (set) effect. *Cognition*, *108*, 652–661.

Chinn, C. A., & Brewer, W. E. (1998). An empirical test of a taxonomy of responses to anomalous data in science. *JOURNAL OF RESEARCH IN SCIENCE TEACHING*, *35*, 623–654.

Gardner, M. (1978). *Aha! insight*. New York: W. H. Freeman & Co.

Kaplan, C. A., & Simon, H. A. (1990). In search of insight. *Cognitive Psychology*, *22*, 374–419.

Knoblich, G., Ohlsson, S., Haider, H., & Rhenius, D. (1999). Constraint relaxation and chunk decomposition in insight problem solving. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *25*(6), 1534–1555.

Knoblich, G., Ohlsson, S., & Raney, G. E. (2001). An eye movement study of insight problem solving. *Memory & Cognition*, *29*(7), 1000–1009.

Luchins, A. S., & Luchins, E. H. (1950). New experimental attempts at preventing mechanization in problem solving. *Journal of General Psychology*, *42*, 279–294.

Mason, L. (2001). Responses to anomalous data on controversial topics and theory change. *Learning and Instruction*, *11*, 453–483.

McCloskey, M., Washburn, A., & Felch, L. (1983). Intuitive physics: The straight-down belief and its origin. *Jornal of Experimental Psychology: Learning, Memory, and Cognition*, *9*, 436–648.

Ohlsson, S. (1992). Information-processing explanations of insight and related phenomena. In M. T. Keane & K. J. Gilhooley (Eds.), *Advances in the psychology of thinking* (pp. 1–44). Upper Saddle River, NJ: Prentice-Hall.

Rumelhart, D. E., & Ortony, A. (1977). The representation of knowledge in memory. In R. C. Anderson, R. J. Spiro, & W. E. Montague (Eds.), *Schooling and the acquisition of knowledge* (pp. 99–135). Hillsdale, NJ: Lawrence Erlbaum Associates.

Seifert, C. M., & Robertson, S. P. (1985). Types of inferences generated during reading. *Journal of Memory and Language*, *24*, 405–422.

Simon, H. A. (2000). Discovering explanations. In R. A. Keil F. C. & Wilson (Ed.), *Explanation and cognition* (pp. 21–59). Cambridge, MA: MIT Press.

Terai, H., & Miwa, K. (2003). Insight problem solving from the viewpoint of constraint relaxation using eye movement analysis. In *proceedings of the 4th international conference of cognitive science* (pp. 671–676).

Wason, P. C. (1960). On the failure to eliminate hypotheses in a conceptual task. *Quarterly Journal of Experimental Psychology*, *12*, 129–140.

Watts, D. M., & Zylberstan, A. (1981). A survey of some children's ideas about force. *Psysics Education*, *16*, 360–365.