

How analogies are generated: The roles of structural and superficial similarity

ISABELLE BLANCHETTE and KEVIN DUNBAR
McGill University, Montreal, Quebec, Canada

Laboratory studies of analogical reasoning have shown that subjects are mostly influenced by superficial similarity in the retrieval of source analogs. However, real-world investigations have demonstrated that people generate analogies using deep structural features. We conducted three experiments to determine why laboratory and real-world studies have yielded different results. In the first two experiments, we used a "production paradigm" in which subjects were asked to generate sources for a given target. Results show that the majority of the analogies that were generated displayed low levels of superficial similarity with the target problem. Moreover, most of the analogies were based on complex underlying structures. The third experiment used a "reception paradigm" methodology. The subjects had to retrieve predetermined sources instead of generate their own. In this case, retrieval was largely constrained by surface similarity. We conclude that people can use structural relations when given an appropriate task and that this ability has been underestimated in previous research on analogy.

An important feature of human cognition is the ability of people to use analogical reasoning (Holyoak & Thagard, 1997). Analogy involves using knowledge from one relatively well-known domain (the source) and applying it to another domain (the target) on which the reasoner is working (Clement & Gentner, 1991; Spellman & Holyoak, 1996). Analogy involves two levels of relations between the representations of the source and the target. *Superficial similarity* refers to the resemblance between the objects in the source and target and their properties (Keane, Ledgeway, & Duff, 1994). *Structural similarity* refers to the resemblance in the underlying systems of relations between the elements of the sources and the elements of the target. Structural similarity exists if the relations between the objects in the source are similar to the relations between the objects in the target, independently of the similarity between the objects themselves (Forbus, Gentner, & Law, 1994). Many authors have argued that structural similarity is the crucial defining feature of analogy (Gentner & Markman, 1997).

Over the past 20 years, many experiments have been conducted to investigate the roles of superficial and structural similarity in analogical reasoning. These experiments have demonstrated that the two types of similarity constrain analogical reasoning in different ways. Superficial similarity has been found to be the main determinant of source access and retrieval, whereas structural similarity influences mostly the mapping and evaluation stages

(Catrambone, 1997; Gentner, Ratterman, & Forbus, 1993; Gick & Holyoak, 1980; Keane, 1987).

Most of the research on analogy has shown that subjects in laboratory experiments tend to rely on superficial features when accessing source analogs. For example, Gentner et al. (1993), using a story-memory task, examined the contributions of the different types of similarity in retrieval. In their experiment, subjects were presented with a set of stories and were later asked to mention any reminding that they experienced when presented with similar stories. Three types of correspondence between pairs of stories were manipulated: similarity between the objects in the source and target, similarity in the first-order relations (relations between the objects) in the source and target, and similarity in higher order relational structure (relations between relations). The results showed that surface similarity and similarity in first-order relations facilitate retrieval. Superficially similar pairs of stories, those sharing object attributes and first-order relations, were more frequently accessed than those with only first- and higher order relations in common. Catrambone (1997), using a similar paradigm but manipulating the degree to which stories shared superficial and structural features, also demonstrated a strong effect of superficial similarity on the retrieval of sources. His results also show that similarity in first-order relational matches increases the frequency of source access.

The facilitating effect of surface similarity on retrieval has also been demonstrated in the context of problem-solving tasks. Gick and Holyoak (1980) used Dunker's radiation problem¹ to investigate analogical problem solving. Their subjects were presented with a source story that suggested a correct convergence solution before they were given the problem-solving task. The results revealed that the vast majority of subjects failed to notice the relevance of the source story when they attempted to solve

This research was supported by FCAR and SSHRC graduate fellowships to the first author and Grant OGP0037356 to the second author from the National Sciences and Engineering Council of Canada. Correspondence concerning this article should be addressed to I. Blanchette or K. Dunbar, Department of Psychology, McGill University, 1205 Docteur Penfield, Montreal, PQ, H3A 1B1 Canada (e-mail: isablan@psych.mcgill.ca or dunbar@psych.mcgill.ca).

the target problem. Only 20% of the subjects retrieved the source analog spontaneously. However, when the surface features of the source were more similar to those of the target problem, 90% of the subjects spontaneously retrieved the source analog to produce the convergence solution (Holyoak & Thagard, 1995). Keane (1987) also looked at the effect of surface similarity in the context of a problem-solving task. In his experiment, surface similarity and structural similarity were manipulated. Keane's results indicate that source analogs that share only structural similarity with the target are seldom retrieved (12%) whereas analogs with common structural and surface features are retrieved often (88%). Overall, these results indicate that people often fail to retrieve a relevant source analog in memory when the analog lacks surface similarity with the target problem.

Some researchers, however, have found some effects of structural similarity in retrieval of analogical sources (Catrambone, 1997; Gentner et al., 1993; Wharton et al., 1994). The role of structural similarity seems to be limited, mostly complementary to the role of surface similarity. In certain cases, greater structural similarity can facilitate retrieval for equal levels of superficial similarity.

Overall, empirical findings from memory and problem-solving tasks indicate that surface similarity is the most important factor in determining retrieval of a source. However, research on the use of real-world analogy has shown that the analogies that people use tend to be based on deep structural features rather than superficial features. For example, Dunbar (1995, 1997) has investigated the use of analogy at weekly meetings in molecular biology laboratories. He has found that over 50% of the analogies that the scientists generate are based on deep structural features rather than superficial features. Similarly, Blanchette and Dunbar (1997) have found that over 60% of the analogies used in newspapers during a political debate in Canada showed low levels of superficial similarity between the source and the target.

Why is it that subjects in standard psychology experiments infrequently use structural features in their retrieval of source analogs, whereas analogies based on structure are frequently used in real-world contexts? One possibility is that research on analogy has tended to use what we call a *reception paradigm*,² in which subjects are given source and target analogs and must see the relationship between them, rather than generate their own analogies. The reception paradigm may constrain the search for structural relations and provide a picture of analogical reasoning that underestimates the subjects' abilities to use deep structural features in the retrieval of source analogs. In real-world contexts, people generate their own analogies. This suggests a different approach to investigating analogy: Rather than a reception paradigm, in which subjects are given sources and targets, one can use a *production paradigm*, in which subjects are given a target problem and are asked to generate source analogs. The production paradigm makes it possible to investigate why people in real-world contexts can generate analogies based on un-

derlying structure, as well as the types of features that they attend to when they generate analogies and the types of search strategies that they use.

We designed a hypothetical situation in which subjects had to produce persuasive political analogies. The topic for which the subjects were asked to generate analogies concerned the debate over the importance of achieving a zero deficit. *Zero deficit* is a term frequently used to describe a government's goal of eliminating the yearly differences between earnings and expenses that lead to national debt. We chose this issue because it is pervasive and important. Many governments in Western societies face substantial problems with public finances. Particularly in Canada, where this study was conducted, the provincial and federal governments are all concerned with this problem and have dealt, or are currently dealing, with the issue in different ways. Although most people agree that public debts are an important problem, the debate resides mainly in how urgent, or how important, the attainment of a zero deficit should be. To simplify the debate greatly, advocates of the pro-zero-deficit side argue that reducing the deficit must be the most urgent priority, and call for drastic cuts in government spending. The anti-zero-deficit side, although not opposed to the reduction of the deficit, most frequently argues in favor of the preservation of social programs.

Our goal in Experiment 1 was twofold. The first goal was to determine whether the use of a production paradigm would make it possible for subjects to retrieve source analogs based on structural features rather than superficial features. If the use of a reception paradigm is the reason why subjects in standard psychology experiments on analogy do not show the same ability to use structural features in the retrieval of source as do subjects in real-world contexts, we would expect subjects in a production paradigm to generate analogies based on structural similarity. The subjects were therefore asked to generate analogies on a given target problem: the zero-deficit issue.

A coding scheme was developed to assess the influence of superficial and structural constraints on the production of analogies in this task. Because *superficial similarity* is defined as the similarity between the objects and their properties in the source and the target (Keane et al., 1994), we categorized the different generated sources according to their semantic domain of origin. Sources from the same categories share similarities in features, and these categories could in turn be evaluated in terms of their level of similarity with the target problem. The underlying structure of each of the analogies was also assessed by identifying both the first-order relational statements (relations between objects) and the higher order relational structures (relations between relations) common to both the source and the target (Gentner et al., 1993).

Our second goal was to further understand the features of the problem that people use to generate and evaluate analogies. In this experiment, we wanted the subjects to give as much information as possible about the ways in which they retrieve and evaluate source analogs. Our

previous research has shown that groups of scientists have little difficulty in generating structural analogs, and that discussion of the features of the sources and targets by members of the group provided insights into the mechanisms that the scientists used in generating analogies (Dunbar, 1995, 1997). Similarly, we used groups of subjects to generate analogies. By using groups, we expected the subjects to discuss features of the sources and targets that were relevant to them. Furthermore, we could get the subjects to evaluate their analogies and gain further information on the dimensions that are important in the evaluation of political analogies.

EXPERIMENT 1

Method

Subjects

Forty-nine undergraduate students majoring in different disciplines participated in the study. They were recruited through sign-up sheets posted in the Department of Psychology at McGill University and through the subject pool of the Department of Psychology, in which case they received credit for their participation in the study. The subjects were compensated for their time.

Twelve groups were run in this study: 4 groups of 3 subjects, 3 groups of 4 subjects, and 5 groups of 5 subjects.

Task and Procedure

A brief introduction to the zero-deficit issue was first read to the subjects. This short text defined terms like *deficit* and *debt*, succinctly described the problem of large public deficits, gave an estimate of the proportions of the problem for Canada and its provinces, and presented two basic opposing viewpoints: pro-zero deficit and anti-zero deficit (see Appendix A). After the subjects had been introduced to the issue, they were read the instructions (see Appendix B). There were two conditions, pro-zero deficit and anti-zero deficit. In both conditions, the subjects were asked to pretend to be consultants who had been hired by either the National Association of Community Groups (NACG, in the anti-zero deficit condition) or the National Association of Responsible Citizens (NARC, in the pro-zero deficit condition). In each case, the group was described as a national coalition that was currently preparing a campaign to sensitize both the government and the citizens to their viewpoint. The subjects were told that this group was aware that analogies are often used in politics to either explain things or convince people to change their opinions. The analogy between Saddam Hussein and Hitler that was used by President Bush during the Persian Gulf crisis (see Spellman & Holyoak, 1992) was given as an example. The subjects were then informed that their "employers" had hired them to generate analogies that could be used in their campaigns to convince people to adopt their viewpoints. In the anti-zero deficit viewpoint, the basic position of the fictitious NACG was the following: The reduction of the deficit is an important goal, but it should not be pursued at any cost, especially if it implies eliminating social programs, healthcare, education, and support for the needy. In the pro-zero deficit viewpoint, the basic position of the NARC was the following: The deficit problem should be the number one priority. Given the state of the economy, governments do not have a choice; spending has to be drastically reduced and citizens have to make sacrifices.

After reading the instructions, the experimenter asked whether the subjects had any questions and then left the group to begin the task, leaving a written copy of the instructions and of the introductory text. The subjects were videotaped and audiotaped during the task. The subjects, working as a group, were allowed 20 min to produce as many analogies as possible. They were asked to write down any analogy that came to mind, regardless of their opinions on the

quality of the analogy. After completing the task, the subjects were asked to select the two best analogies.

Data Analysis

To analyze the analogies produced by the subjects, all elements explicitly stated in the source and all elements explicitly stated in the target were first listed. Information on explicit links established by the subject between elements in the source and in the target was also used to complete the description of each analogy. Codings were based on information both from the verbal descriptions of the analogies made by the subjects during the task and the written descriptions also produced during the task.

Coding Scheme

The analogies generated by the subjects were coded along different dimensions reflecting the use of superficial and structural features: semantic category of source, range, and underlying structure. In addition, the analogies were also coded for the degree of elaboration and explicitness.

Semantic category of source. The source analogs were classified into different semantic categories. The set of source categories used was not imposed on the data, but rather, semantically similar sources were grouped together. A minimum number of five sources had to be grouped together in order to create a source category. Sources not corresponding to any of the categories were classified in a category labeled *other*.

Range. Sources were also coded on their *range*, defined as the degree of semantic similarity between the source and the target. Range was created to provide a global estimate of the degree of superficial similarity between the source and the target. In this case, in which the target was always the deficit problem, analogies coded as *within-domain* were analogies for which the source was taken from the domains of politics, economics, or personal finances. Sources from these three domains have a high degree of superficial similarity with the target problem. Analogies with sources from any other domain (see Table 1) were coded as *other-domain* analogies.

Underlying structure. The underlying structures of the analogies were determined. For each analogy, the relational structure common to both the source and the target was identified. Because many of these relational structures were identical or highly similar (within each condition, pro-zero deficit and anti-zero deficit), they were also grouped into similar structure categories following the same procedure as that for the semantic categories, with a criterion of five similar structures as a minimum for establishing a category (see Table 4). Analogies using relational structures that did not correspond to any of the major categories identified through the grouping process were put in the category *other*.

Number of elements in source and number of elements in target. To provide an estimate of the degree of elaboration of the analogies, the number of elements explicitly stated in the source and the number of elements explicitly stated in the target were counted. This measure is designed to provide a partial indication of the structural overlap between source and target.

Explicitness of the mapping. Because the analogies varied widely on how explicit the relations between elements in the source and target were, a coding dimension regarding the explicitness of the mapping was added. Analogies were coded as *totally explicit* if all the elements stated in the source had an explicit counterpart in the target. *Totally implicit* analogies were those for which no elements of the target were mentioned, keeping the correspondence to the target problem totally implicit. Analogies in which some of the source elements were explicitly linked to elements from the target but others were not were coded as *partially explicit*.

Production measures. Because of the nature of the task, three different ways of computing the number of analogies produced by each subject were used. A first measure was the number of analogies initiated by a subject. A second production measure, the total

Table 1
Semantic Source Categories: Experiment 1

Source Category	Example	<i>n</i>	%
Animals/farming	"If you have a farm, and you decide that it's too expensive to buy pesticide for your apples, but then all your apples die."	5	3.6
Domestic tasks	"The deficit is like lint in your dryer. If you don't pull it out, it just accumulates and blocks, and it becomes inefficient. (...) - It will grow and render the machine inoperable."	6	4.3
Eating	"Cutting a little of social programs today will ensure their continuance in the future, like an apple pie, without immediate restraint on eating, the pie will disappear and there will be none left for the future eaters the next day."	9	6.5
Fables/characters	"There's the custard pot and there's the girl and she leaves it unattended and the custard pot, it's a magic one, it keeps making custard. So you can think debt here. It just keeps bubbling over because she's not paying attention to it."	7	5.1
Illness/medicine	"The deficit is like cancer, you have to get rid of it. - Also, it multiplies if you don't take care of it. (...) the whole idea is that it gets worse."	12	8.7
Economics/politics	"There was a government in England in the 70's, a labor government (...) that had a lot of social programs, and in the end, they just got too far into debt and they had to go to the IMF and say we need a loan and the IMF said well, we won't give you the loan unless you change your programs (...) if we want to maintain our ability to make or own decisions about policies, we have to stop this now before it takes control of us."	28	20.3
Natural resources/disasters	"Throwing all the water on one flaming ember and forgetting the adjacent (roaring) fire."	19	13.8
Personal finances	"So it's like a baby and you want the baby to grow up well. So to do that, you don't sacrifice all the care for the baby until it's 18 years old just so you can have enough money then to send it to college (...) you have to care for it all along cause early damage could create more damage later on."	18	13.0
Sayings/expressions	"I'm thinking of 'A chain is only as strong as its weakest link' (...) the most vulnerable people have to be strengthened, not weakened."	11	8.0
Others	"The government is writing a great last page to a novel without including the rest of the story."	23	16.7
Total		138	100

number of analogies, included both the number of analogies that a subject initiated and the number of analogies that the subject did not initiate but in which the subject participated by adding a mapping or an element. The total number of analogies produced by each group was also recorded.

RESULTS

Production of Analogies

A total of 138 analogies were produced by the subjects. On average, subjects initiated 2.82 ($SD = 1.74$) analogies and participated in 2.82 ($SD = 2.29$) additional analogies for a total average production of 5.63 ($SD = 3.0$) analogies per subject. The average number of analogies produced by a group was 11.50 ($SD = 4.15$).

Semantic Category of Source

Other-domain analogies constituted 66.67% ($n = 92$) of the total number of analogies produced, whereas within-domain analogies represented 33.3% ($n = 46$). Nine semantic source categories were obtained. These results and the corresponding frequencies for each cate-

gory are presented in Table 1, along with examples from each of the categories. The source categories used most often were economics/politics ($n = 28$, 20.3%), natural resources/disasters ($n = 19$, 13.8%), and personal finances ($n = 18$, 13.0%). Sources in the *other* category represented 16.7% ($n = 23$) of the total number of analogies coded. This category included sources such as training wheels, hostages, or a leaky bucket, all drawn from domains which, overall, were infrequently used.

Underlying Structure of Analogies

Ten different categories of underlying structures were identified through the grouping process, 5 for each viewpoint (anti- and pro-zero deficit). These structure categories are presented along with their respective frequencies in Table 2. For each condition, specific first-order relational statements were very prevalent. In the anti-zero deficit condition, the first-order relational statement most often found was "If cut X, then negative consequence." In this case, the specific objects representing both "X" and "negative consequence" varied in the sources. In the targets, "X" and "negative consequence" typically corre-

Table 2
Structure Categories: Experiment 1

Category No.	Target	Relational Structure Common to Source and Target	<i>n</i>
Anti-Zero Deficit Condition			
1	SP is needed If SP is cut Then negative consequence Therefore: don't cut SP	X is needed If X is eliminated Then negative consequence Therefore don't eliminate X	11
2	If SP is cut Then solve D in the short term But If SP is cut Then more problems in the long term	If cut X Then solve Y in short term But If cut X Then more problems in long term	7
3	SP is needed If only focus on D and neglect need for SP Then negative consequence	X is needed If only focus on Y and neglect X Then negative consequence	14
4	If want to help people Then solve D If want to solve D Then have to cut SP If cut SP Then negative consequences for people	If want to help people Then solve Y If want to solve Y Then have to cut X If cut X Then negative consequences for people	6
5	Present solution: cut SP to solve D Alternate solution: Do Z instead of cut SP to solve D	Present solution: Cut X to solve Y Alternate solution: Do Z instead of X to solve Y	8
Pro-Zero Deficit Condition			
6	D grows exponentially If solve D now Then less problems If solve D later Then more problems	Y grows exponentially If solve Y now Then less problems If solve Y later Then more problems	15
7	SP are not needed And cannot afford SP Therefore should cut SP To solve D	X is not needed And cannot afford X Therefore should cut X To solve Y	4
8	If D not solve Then negative consequences Therefore should solve D	If Y not solved Then negative consequences Therefore should solve Y	16
9	If do M (minimal effort) Then won't solve D If do L (large effort) Then will solve D	If do M Then won't solve Y If do L Then will solve Y	9
10	D makes you subservient to other Therefore eliminate D Therefore can be free	Y makes you subservient to other Therefore eliminate Y Therefore can be free	10
Others			28
Total			128

Note—SP, social programs (a specific one or social programs in general); D, debt problem; M, minimal effort; L, large effort.

sponded to social programs and any consequence of cutting the social programs, respectively. In the pro-zero deficit condition, the first-order relational statement most often found was "If Y not solved, then negative consequences." Again, the objects corresponding to "Y" and "negative consequence" varied in the sources. In the target, "Y" usually corresponded to the debt problem, and the negative consequences of having a debt were associated. In each condition, these first-order relational statements were typically accompanied by other statements, additional premises, or additional consequences. The way in which the different first-order relational statements were put together constituted the defining feature of each of the different structure categories.

Explicitness and Elaboration

A majority of the analogies ($n = 77, 56\%$) were partially explicit. Forty-four (32%) analogies were totally implicit, 14 analogies were totally explicit (12%). On average, analogies contained a greater number of elements explicit in the source ($M = 3.10, SD = 1.25$) than in the target [$M = 1.22, SD = 1.11; t(137) = 14.57, p < .001$].

Comparison of Anti- and Pro-Zero Deficit Conditions

Independent sample *t* tests revealed no differences between anti- and pro-zero deficit conditions on the total number of analogies produced [$t(47) = 1.49, p > .05$] and number of analogies initiated [$t(47) = 1.37, p > .05$].

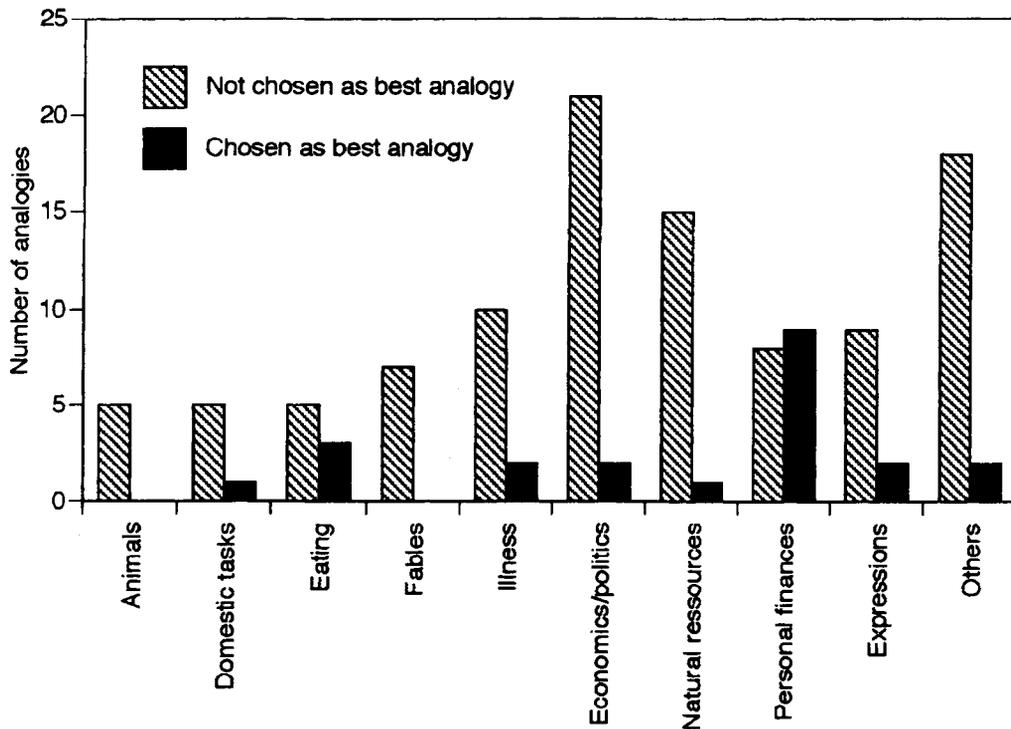


Figure 1. Choice of best analogy and source category, Experiment 1.

Analogies from the two conditions also did not differ on range [$\chi^2(1, N = 138) = .005, p > .05$], explicitness of the mapping [$\chi^2(9, N = 138) = .23, p > .05$], number of elements in the source [$t(136) = 1.18, p > .05$], and number of elements in the target [$t(136) = .90, p > .05$]. The only variable on which the *anti* and *pro* analogies differed was source category [$\chi^2(9, N = 138) = 18.89, p < .05$]. Pro-zero deficit analogies used sources from the natural resources/disasters category more frequently, and the sayings/expressions domain was used more frequently by the anti-zero deficit analogies.

Features Related to Choice of Analogy

A chi-square performed on source category \times choice of best analogy (chosen vs. not chosen) revealed a significant link between the two variables [$\chi^2(9, N = 125) = 22.87, p < .01$]. Specifically, sources from the "personal finances" category were more likely to be chosen as the best sources, as can be seen in Figure 1.

Structure category was also significantly related to choice of best analogies [$\chi^2(10, N = 125) = 18.56, p < .05$]. Analogies with relational structures from the category *others* were chosen proportionally less as the best analogies.

Another chi-square reveals the explicitness of the mapping (totally explicit, totally implicit, and partially explicit) was not related to the choice of best analogy [$\chi^2(2, N = 125) = .22, p > .05$]. Although explicitness was not related to the choice of best analogies, elaboration in the source was. A *t* test was performed to compare the analo-

gies chosen as best and the analogies not chosen on the number of elements contained in the source. On average, the analogies selected as best contained a greater number of elements in the source [$t(123) = 2.03, p < .05$]. A similar *t* test performed on the number of elements in the target showed no similar difference between chosen and not chosen analogies [$t(123) = 1.30, p > .05$].

DISCUSSION

The results of this first experiment support our hypothesis that a production paradigm would allow subjects to make more use of the structural features in the retrieval of source analogs. The majority of analogies generated by the subjects in this study did not exhibit high levels of superficial similarity with the target problem (67%). The subjects made use of different types of underlying structures in the generation of analogies. Specific first-order relational statements were used very frequently in each condition. Different combinations of these first-order relational statements produced the different structure categories that were used in the two conditions.

Although the influence of superficial constraints was not as important as might have been expected given the laboratory research findings, it did have a noticeable influence on retrieval and evaluation of analogies. Of the nine semantic categories of sources, the two categories exhibiting the greatest superficial similarity with the target deficit problem, economics/politics and personal finances, were the most and third most frequently used categories.

Table 3
Source Categories: Experiment 2

Source Category	<i>n</i>	%
Animals/farming	5	4.0
Domestic tasks	3	2.4
Eating	7	5.6
Fables/characters	4	3.2
Illness/medicine	14	11.3
Economics/politics	8	6.5
Natural resources/disasters	6	4.8
Personal finances	17	13.7
Sayings/expressions	6	4.8
Others	54	43.5
Total	124	100

Furthermore, the subjects' evaluations of analogies seem to have been affected by superficial similarity. Analogies using sources from the "personal finances" category were evaluated more positively than others.

Apart from the semantic source category, the only other feature significantly related to the evaluation of the analogies was the degree of elaboration in the source. The analogies selected as best contained a greater number of elements in the source.

The degree of explicitness of the analogies produced by the subjects in this experiment is also interesting. Most of the analogies were partially explicit. The objects in the source were not systematically related to ones in the target and many of the correspondences between source and target were left implicit. Moreover, the number of elements in the source was greater than the number of elements in the target.

Overall, the generation of analogies in this experiment appears not to have been highly constrained by superficial similarity. There was great diversity in the analogies generated. The domains from which the subjects retrieved sources were not limited to domains highly similar to the target problem such as economics, politics, and personal finances. Rather, the sources were drawn from domains as varied as natural resources, eating, illness, and domestic tasks. It is possible that the enhanced use of structural properties in the generation of analogies seen in these results is an artifact of using a group task. Groups might exhibit greater use of structural features because the subjects have different backgrounds. It is also possible that group reasoning leads to an increase in processing capacity, which would in turn increase the ability to use structural features in the retrieval of sources. To test this hypothesis, we conducted a second experiment, in which we used the same task but the subjects worked individually.

EXPERIMENT 2

Method

Subjects

Twelve undergraduate students participated in the study. They were recruited through sign-up sheets posted in the Department of Psychology at McGill University. Subjects were compensated \$6 for their time.

Task and Procedure

The subjects in Experiment 2 were given the same task and instructions as in Experiment 1, but they worked individually instead of in groups. The procedure was also the same as in Experiment 1. The subjects were read the same introduction before hearing their specific instructions, which again assigned them to either the anti- or the pro-zero deficit condition. They were also asked to come up with as many analogies as possible and to write them down. After 20 min, the experimenter came back and asked the subject to select the best two analogies. The only difference in procedure with Experiment 1 was that the experimenter asked the subject for a brief description of each of the analogies after the task was completed.

Data Analysis

The data were analyzed in exactly the same way as in Experiment 1. The information written down by the subject during the task and the subsequent description of the analogies made to the experimenter after the task were used. The coding of the analogies mirrors the coding in Experiment 1 (source category, range, structure category, explicitness of the mapping, number of elements explicit in the source, and number of elements explicit in the target).

RESULTS

Production of Analogies

A total of 124 analogies were produced. On average, the subjects produced 10.73 analogies during the task ($SD = 5.71$).

Semantic Category of Source

The majority of analogies produced were other-domain analogies ($n = 100$, 80.6%). The same semantic categories of sources were used as in Experiment 1. The frequency of use of the different semantic categories can be found in Table 3. The two categories used most frequently were personal finances ($n = 17$, 14%) and illness/medicine ($n = 14$, 11%). Sources in the category *other* constituted 44% ($n = 54$) of the total number of sources generated by the subjects.

Underlying Structures

The same 10 structure categories identified in Experiment 1 were also used by the subjects in this study. The frequency of use of the different structure categories are presented in Table 4.

Elaboration and Explicitness

For 52 analogies (42%), the mappings were totally implicit. In 31 analogies (25%), all elements in the source were explicitly linked to elements in the target and as such were coded as totally explicit. The remaining 41 analogies (33%) were partially explicit. Overall, analogies contained a greater number of elements in the source ($M = 2.65$, $SD = 1.41$) than in the target [$M = 1.55$, $SD = 1.44$; $t(123) = 7.73$, $p < .001$].

Comparison Between Anti- and Pro-Zero Deficit Conditions

A one-way analysis of variance (ANOVA) of the number of analogies produced as a function of condition showed that there was no significant difference between

Table 4
Structure Categories: Experiment 2

Category No.	Target	Relational Structure Common to Source and Target	<i>n</i>
Anti-Zero Deficit Condition			
1	SP is needed If SP is cut Then negative consequence Therefore: don't cut SP	X is needed If X is eliminated Then negative consequence Therefore don't eliminate X	13
2	If SP is cut Then solve D in the short term But If SP is cut Then more problems in the long term	If cut X Then solve Y in short term But If cut X Then more problems in long term	8
3	SP is needed If only focus on D and neglect need for SP Then negative consequence	X is needed If only focus on Y and neglect X Then negative consequence	8
4	If want to help people Then solve D If want to solve D Then have to cut SP If cut SP Then negative consequences for people	If want to help people Then solve Y If want to solve Y Then have to cut X If cut X Then negative consequences for people	21
5	Present solution: Cut SP to solve D Alternate solution: Do Z instead of cut SP to solve D	Present solution: Cut X to solve Y Alternate solution: Do Z instead of X to solve Y	2
Pro-Zero Deficit Condition			
6	D grows exponentially If solve D now Then less problems If solve D later Then more problems	Y grows exponentially If solve Y now Then less problems If solve Y later Then more problems	17
7	SP are not needed And cannot afford SP Therefore should cut SP To solve D	X is not needed And cannot afford X Therefore should cut X To solve Y	3
8	If D not solve Then negative consequences Therefore should solve D	If Y not solved Then negative consequences Therefore should solve Y	8
9	If do M (minimal effort) Then wont solve D If do L (large effort) Then will solve D	If do M Then won't solve Y If do L Then will solve Y	14
10	D makes you subservient to other Therefore eliminate D Therefore can be free	Y makes you subservient to other Therefore eliminate Y Therefore can be free	2
Others			11
Total			107

Note—SP, social programs (a specific one or social programs in general); D, debt problem; M, minimal effort; L, large effort.

the pro- and anti-zero deficit conditions [$F(1,10) = .922, p > .05$]. Analogies from the two conditions again did not differ on range, explicitness of the mapping, or number of elements in the target. As in Experiment 1, the analogies from the two conditions differed on the source categories used [source category \times condition, $\chi^2(9, N = 124) = 20.86, p < .05$]. In addition, there was also a significant difference on the number of elements contained in the source. A *t* test revealed that analogies produced in the pro-zero deficit condition contained on average a greater number of elements in the source than analogies from the anti-zero deficit condition.

Evaluation of Analogies

A chi-square on source category \times choice of best analogy (chosen vs. not chosen) indicated that the choice of best analogies was not related to the source category of the analogies [$\chi^2(9, N = 124) = 7.07, p > .05$]. Structure category also was not related to choice of best analogy [$\chi^2(10, N = 124) = 9.01, p > .05$].

Another chi-square was performed on explicitness of the mapping \times choice of best analogy. There was a significant relation between the choice of best analogies and the explicitness of the mapping [$\chi^2(2, N = 124) = 17.87, p < .001$]. Analogies that were totally implicit were se-

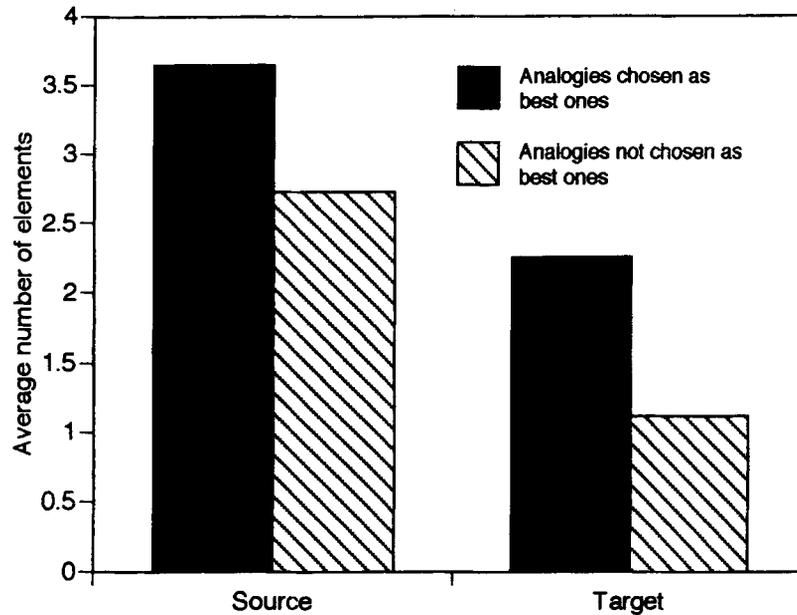


Figure 2. Choice of best analogies and number of elements explicit in the source and target, Experiment 2.

lected less than the analogies that were either partially or totally explicit.

The number of elements in both the source and the target was also consistently related to the evaluation of the analogies. ANOVAs performed on the number of elements in the source and target of chosen and not-chosen analogies revealed analogies chosen as best had a greater number of elements explicit in both the source [$F(1,122) = 17.30, p < .001$] and the target [$F(1,122) = 45.21, p < .001$], as can be seen in Figure 2.

DISCUSSION

The results of Experiment 2 replicate those of Experiment 1 and were even more pronounced than those of Experiment 1. The subjects who worked individually appear to have been even less influenced by superficial similarity than their counterparts who worked in groups. A great majority (81%) of the analogies produced in this second experiment were not superficially similar to the target deficit problem. A surprising finding was the proportion of source analogs that could not be classified into any of the semantic source categories (44%). The subjects in the individual condition exhibited a greater use of original, not frequently used, sources than did the groups. In addition, in this individual condition, the subjects' choice of best analogy was not influenced by the semantic category of the source. Sources from the personal finances category were not selected more frequently as best analogy than sources from the other categories. Overall, superficial similarity seems to have constrained subjects working individually even less than the subjects working in groups.

The subjects' choice of best analogies was again related to the degree of elaboration in the source. In addition, the degree of elaboration in the target was also related to the choice of best analogies. For both the source and the target, analogies chosen as best contained a greater number of elements.

Overall, it appears that the results obtained in the first experiment were not an artifact of using a group reasoning situation. In group and individual situations, subjects are able to generate analogies based on deep structural features. The retrieval of sources was not highly constrained by superficial similarity. We suggest that this is due to our use of a production paradigm. But this difference in method is not the only difference between ours and that of previous studies. Another important difference may be found in the materials used, which, in our studies, differ sharply from anything else used in the literature to date. Real-world political problems are complex and ill defined. Most studies of analogical retrieval have used short stories or even a set of short sentences (Gentner et al., 1993; Wharton et al., 1994) in which the characters and story line are well defined. Another way in which this material is different is that the subjects were familiar with the topic. The deficit problem was one of the most important political problems when the experiment was run. Although the subjects' level of political expertise varied, all had at least some understanding of this problem prior to the experiment and had heard about it in the media. It is well known that the content of a problem can affect performance (Ericsson, 1994). Furthermore, the ability to use structure in problem solving can be related to expertise in a specific domain (Ericsson, 1994). It is possible that the subjects' ability to use structure in the first two ex-

Table 5
List of Materials: Experiment 3

Target No.	Target Text	Sources	
		Superficial Match	Structural Match
1	<p>The Canadian government has an important debt.</p> <p>The Canadian population also needs a good health care system.</p> <p>When the government only focuses on reducing the debt and does not attend to the important need for health services, the population suffers.</p>	<p>Buying medication for a child who is sick can be very expensive for parents who are not covered by an insurance plan.</p> <p>This is one of the reasons why the government created the Quebec drug insurance plan under which everybody is insured.</p>	<p>A plant is in much need of some water.</p> <p>Plants in general also need sunshine.</p> <p>If the owner of the plant only gives it adequate amounts of water and forgets to put it where it can get enough sunshine, it will die.</p>
2	<p>The national debt grows exponentially.</p> <p>If the government reduces expenses now, interest on this debt will stop accumulating.</p> <p>If he waits until later to proceed with the cuts, interest will continue to accumulate and the debt will be harder to pay off.</p>	<p>A young couple is unsure about whether they should buy a house or not.</p> <p>Buying a house implies having a big debt and paying lots of interest for many years to come.</p> <p>On the other hand, this is an investment unlike rent.</p>	<p>Cancer is an illness that progresses rapidly.</p> <p>If a patient undergoes chemotherapy immediately, the tumors will stop growing.</p> <p>If this person waits to proceed with therapy later, tumors will continue to multiply and the cancer will be a lot more difficult to eradicate.</p>

periments was not due to the method (a production paradigm), but rather to the content of the problem, a real-world relevant and known political problem.

EXPERIMENT 3

To test this possible alternative explanation, we ran a third experiment with the same analogies but in a reception paradigm. The goal was to keep constant the content of the analogies while varying the task, to see whether this would influence the subjects' ability to use structure. In this experiment, we used the same type of method employed in most studies of analogical retrieval: a reception paradigm. The subjects were first asked to read a set of sources. Later, they were presented with targets and were asked to retrieve one of the initial sources. The sources and targets used were taken from the subjects' productions in the first two experiments. The interest here lay in seeing whether retrieval would be based on structural or superficial similarity. If the content of a problem determines whether people can use structure in the retrieval of analogical sources, retrieval in this case should be based on structural similarity. If the increased ability to use structure, seen in the first two experiments, was in fact a function of using a production paradigm, retrieval here should be based on superficial similarity.

We selected two target arguments, one for the anti-zero deficit side, and one for the pro-zero deficit side. For each

target, we selected one superficial match and one structural match. These four matches, in addition to eight fillers, constituted the sources that were read by the subjects in the first part of the experiment. Later, they were presented with the two target arguments and two target fillers. For each target, they were asked to retrieve a similar source.

METHOD

Subjects

Twenty-two undergraduate students participated. They were recruited through the subject pool of the Department of Psychology, and they obtained course credit for their participation.

Materials

As mentioned previously, the stimuli for Experiment 3 were taken from the subjects' productions in the previous two experiments. These analogies were separated into sources and targets (see Table 5 for a list of the materials). Of the selected targets, one was against zero deficit (Target 1) and one was for zero deficit (Target 2). For each target, there was one structural match and one superficial match. The structural matches were identical in structure to the target but had very low superficial similarity. They were taken from the domains of medicine (cancer) and gardening (plant), domains very different from the target domain of politics and economics. The superficial matches contained elements that were highly similar to the ones in the targets, such as money, government, debts, interest, and so on, but they had no structural overlap. The sources had to be slightly modified from their original formats (as produced by the subjects in the first studies) to maximize or minimize superficial and structural similarities.

Table 6
Example of Filler Paragraphs: Experiment 3

Filler	Paragraph
Source	A woman is planning a 1500 mile solo-trek across Antarctica. She has been training for this expedition for the past two years. Her training has included 10-mile runs and hours of weight training each week. During her expedition, she'll take in about 5,000 calories a day and expend about 6,000.
Target	A school teacher recently started a Girls' Club. In this club, the girls get to talk about anything they want. The goal is to empower girls to retain strong identities and avoid different problems such as depression and eating disorders.

Because the two targets were from the same domain—both addressed the problem of public finances—the superficial match for one target was also a good superficial match for the other target and vice versa. Since all subjects were presented with the two targets and their corresponding matches, there were two superficial matches for each target.

In addition to these target and source matches, eight fillers were designed to be presented with the sources and two with the targets (see Table 6). All the filler paragraphs were taken from current events reported in the media. Only events that did not mention money, finances, government, or anything that could be superficially or structurally similar to the targets were selected. All paragraphs were carefully equated for the number of words (between 40 and 50), the number of propositions (5 or 6), and the number of times a given element was repeated across the different sentences (2).

Procedure

The subjects were told that the Experiment was about creative thinking and would take place in two parts. In the first part, they were told that they would have to read short paragraphs and evaluate them for pleasantness. The four sources and eight fillers were presented in a random order on a computer screen. After the subjects read each paragraph, a seven-point scale appeared on the screen. Their task was to enter a number corresponding to their evaluation of the pleasantness of the text. During the second part, which was in fact a distractor task, the subjects were asked to generate ideas for completing a cartoon strip that also appeared on the computer screen. They were given 2 min to generate as many ideas as possible for each of two successively presented cartoon strips.

The subjects were then asked to read four more paragraphs. They were instructed that after reading each one, they should tell the experimenter which of the paragraphs that they had read in the first part of the experiment the current text reminded them of. The four paragraphs consisted of the two targets and two fillers. The subjects' answers were tape recorded.

There were two different conditions. In one condition, the subjects were asked to say "Which of the paragraphs you read in the first part is this one similar to." In another condition, they were asked "Which of the paragraphs you read in the first part would make a good analogy to this one?" We wanted to examine whether asking specifically for an analogical match would increase the number of retrievals based on structural similarity. Traditionally in this type of experiment, people have asked the first type of question (Gentner et al., 1993), probing for "similar," and not specifically "analogically similar," sources. We know that when no retrieval of sources from memory is involved, people interpret analogical similarity as structural correspondence (Gentner et al., 1993). It might be the case that asking specifically for analogical similarity would provide different results than simply asking for similarity.

Materials Check

Before drawing any conclusions, we needed to verify whether the structural matches selected for this experiment were actually per-

ceived as good analogies by independent raters. We asked another group of subjects ($n = 8$) to read the targets, sources, and fillers, and to choose, for each target, which other paragraph would make the best analogy. For each of the two targets, all subjects chose the structural match as the best analogy, confirming that these matches were indeed perceived as better analogies than the superficial matches when no retrieval from memory was needed.

RESULTS

For each target, retrievals could be one of three types: a superficial match, a structural match, or a filler. Because all sources were presented to all subjects, the structural match for Target 1 was coded as a filler for Target 2, and vice versa. Also, if a subject answered any of the two superficial matches for any target, it was taken as a retrieval of a superficial match, since the two targets were from the same domain and were highly similar in surface to each other.

Differences Between Conditions

Whether we asked for similarity or analogical similarity did not result in any difference in the proportion of superficial and structural matches retrieved. A chi-square performed on type of match retrieved revealed that the two conditions did not differ in the proportion of structural matches, surface matches, and fillers retrieved [$\chi^2(2, N = 44) = 0.838, p > .05$]. This was true for each of the two targets separately [Target 1, $\chi^2(2, N = 22) = 1.077, p > .05$; Target 2, $\chi^2(2, N = 22) = 2.207, p > .05$]. We therefore combined the data from the two conditions for all subsequent analyses.

Retrieval of Structural and Superficial Matches

The subjects retrieved matches mainly on the basis of surface similarity. A chi-square was performed to compare the number of times that each type of match was retrieved and expected frequencies. This analysis indicated that retrieval was based mainly on superficial similarity [$\chi^2(2, N = 44) = 97.228, p < .001$]. As can be seen in Figure 3, the proportion of superficial matches retrieved far exceeds the proportion of structural matches and of fillers. For the two targets combined, 65% of retrievals were based on surface similarity and 16% were based on structural similarity. Specifically, for Target 1, 16 subjects retrieved a superficial match, whereas only 1 retrieved the structural match [$\chi^2(2, N = 22) = 53.73, p < .001$]. For

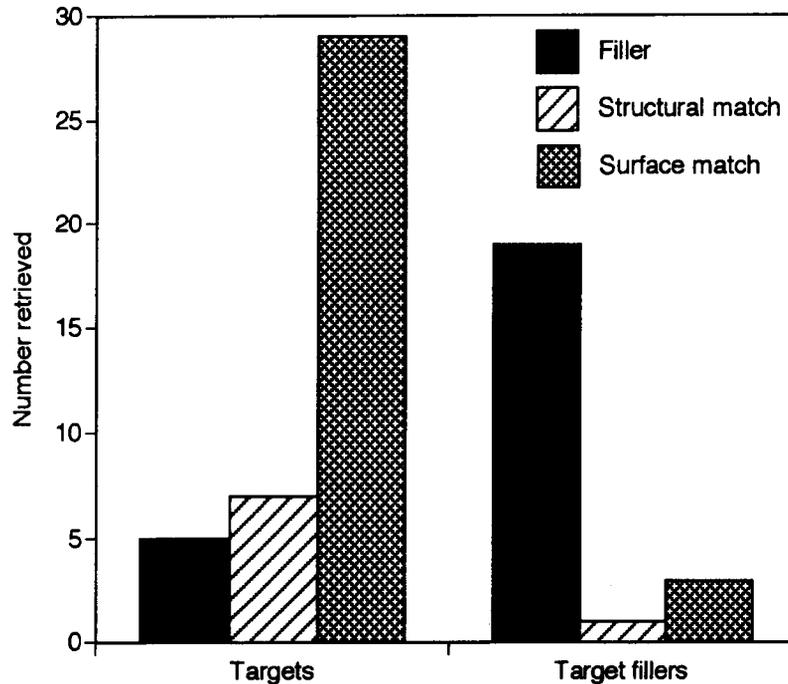


Figure 3. Retrieval of superficial and structural matches for targets and fillers.

Target 2, 13 subjects retrieved a superficial match and 6 retrieved the structural match [$\chi^2(2, N = 22) = 52.367, p < .001$].

Retrieval based on the two target fillers did not show such a pattern. Retrieval in this case was not significantly different from expected frequencies [$\chi^2(2, N = 44) = 0.492, p > .05$].

GENERAL DISCUSSION

The main finding to emerge from this study is that the use of a production paradigm allows subjects to retrieve analogies based on deep structural features. The subjects were not highly constrained by superficial similarity. These findings are surprising, given the results of previous laboratory tasks; but they closely mirror what we have found in real-world investigations of analogical reasoning (Blanchette & Dunbar, 1997; Dunbar, 1995, 1997). Previous work on analogical reasoning has mainly, although not exclusively (Wharton et al., 1994), focused on the influence of superficial similarity in the retrieval of source analogs (Gentner et al., 1993; Keane, 1987). We have found in studies of real-world scientific and political reasoning that a large proportion of the generated analogies were based on deep structural features rather than superficial features (Blanchette & Dunbar, 1997; Dunbar, 1995, 1997). The results presented here also show that most of the sources generated by the subjects had little superficial similarity to the target problem. This was especially true when the subjects worked individually. The groups' production was somewhat more influenced by superficial

similarity. The groups produced a greater proportion of analogies with high superficial similarity to the target problem, and their evaluation of analogies was also influenced by superficial similarity, which was not the case for the individuals. Still, the subjects in both conditions were able to generate analogies exhibiting deep structural similarity between source and target. The third experiment showed that this ability to use structure in the generation of analogies was not due to the type of materials used. If the ability to use structure had been due to the subjects' familiarity with the content of the problem, we would expect that in a different paradigm, but with the same analogies, the subjects would have been able to retrieve on the basis of structure. This was not the case. When using the same analogies in a reception paradigm, the subjects were not able to retrieve sources based on structure, but they retrieved overwhelmingly on the basis of surface similarity.

Why does the generation of analogies lead to a focus on structural features, whereas being asked to remember which source is most similar to the target leads to a focus on superficial features? One hypothesis is that the types of encoding required for production and reception tasks are different. In the reception task, the subjects were asked to rate each passage (sources and fillers) for pleasantness. Previous studies that have concluded that people retrieve on the basis of superficial similarity have used similar methods and instructions. It may be the case that the standard paradigm implicitly cues the subjects to encode superficial features and hence these features are used in retrieving analogical sources. In the production task used

in the present experiments, it is highly likely that the subjects encoded the target problem in terms of structural features and therefore were able to retrieve analogical sources on the basis structural similarity. Thus, the type of encoding of both the source and the target will determine whether superficial or structural matches will be used.

Another possible explanation for the finding that subjects generated structural analogies is that they had considerable knowledge of the target and source domains before entering the experiment. However, in the third experiment we used the same target problem and picked familiar sources that the subjects had generated in the first two experiments: cancer and gardening. Thus, it is not the amount of knowledge of the source domains but rather the type of encoding and retrieval that differed between the production and reception tasks. The results suggest that the type of cognitive processing required will determine the type of similarity on which the retrieval of sources is likely to be based.

Another alternative explanation for the findings of these experiments could be that people retrieve, but spontaneously reject, all superficial matches, especially because they are asked to generate persuasive analogies. If this were the case, most of the retrieval would be based on surface similarity, but the analogies that the subjects reported would have been based on structural similarity because those would have been the only analogies that survived the evaluation stage. However, there are two reasons why this is an unlikely explanation for the results. First, it is possible for analogical sources to be both structurally and superficially similar to the target. If the subjects' retrieval was based on surface similarity, but they then rejected all those that did not have structural similarity, most of the analogies would still be superficially similar—that is, both superficially and structurally similar. Second, although it is true that the subjects in the generation task were asked to come up with persuasive analogies (which could promote the rejection of exclusively superficial matches), the subjects in the reception task were asked to retrieve “what would make a good analogy.” We used these instructions knowing that people generally interpret “a good analogy” to mean “a good structural match” (Gentner et al., 1993). If evaluation were responsible for the differences observed between the two tasks, people should also have rejected the exclusively superficial match in the reception task.

One final possible explanation for our results is that the subjects in the generation task were not really retrieving sources but rather retrieving analogies already stored in memory, ones they had heard in the media or elsewhere. It is possible that the subjects had heard such analogies, given that the deficit was one of the most important political issues when we ran the study. However, it is unlikely that this was responsible for the results obtained here. First, if the subjects in the generation task really were familiar with a set of analogies on the deficit issue, this should also have been true of the subjects in the reception task. In this case, the subjects in the reception task would

have been even more likely to retrieve these already stored analogies, since they had read the sources just a few minutes before they read the target. This should have facilitated retrieval of the structural match. The second reason why it is unlikely that the subjects were retrieving already stored analogies is the diversity present in the subjects' production. The subjects generated sources from nine different source categories, and substantial amounts were not even classifiable into any category. We saw in the subjects' production a diversity that we, as careful observers of the debate over the deficit, have never seen in the media. If the subjects really were retrieving prestored analogies, there would have been much less variation in the types of sources used.

The different types of structure underlying the analogies produced by the subjects in the first two experiments were very complex. A set of first-order relational statements was used in each of the anti- and pro-zero deficit conditions. These first-order relational structures were combined in different ways and were accompanied by additional premises or consequences to produce the different structure categories. Each of these structure categories basically corresponds to the underlying structure of an argument that can be used to argue for or against zero deficit. For example, Structure Category 1 (X is needed; If X is eliminated; Then negative consequence; Therefore don't eliminate X) corresponds to a basic anti-zero deficit argument to the effect that social costs such as insecurity, poverty, and violence will result from the elimination of social support systems. In the pro-zero deficit condition, Structure Category 10 (Y makes you subservient to other; Therefore eliminate Y; Therefore can be free) also corresponds to a basic argument of the pro-zero deficit side—namely, that substantial debts cause governments to be vulnerable and dependent on the lenders.

The subjects were able to abstract out the underlying structures of political arguments and in turn to use those in the generation of analogies. This is in line with other findings showing that, through reasoning, people are able to extract structure out of complex political problems and in turn to use this structure for analogical reasoning (Faries & Reiser, 1990). Some of the structures used by the subjects in the present experiments integrated parts of arguments presented in the introduction (Structure Categories 1, 2, 3, 6, 8, and 10). Other structure categories correspond to arguments that were not presented in the introduction (Structure Categories 4, 5, 7, and 9). In any case, the surprising finding is that the subjects were able to use the underlying structure of these arguments to produce analogies and were not constrained to generate analogies based on superficial similarity to the target problem, as could have been expected given the findings of previous experiments on the retrieval of source analogs (Gentner et al., 1993; Keane, 1987).

The use of a production paradigm allowed us to broaden our understanding of analogical reasoning. Similarly, other researchers have recently resorted to “generative” paradigms in different areas of reasoning. Ward (1994)

asked subjects to imagine animals living on another planet. Bassok, Chase, and Martin (1998) asked subjects to construct arithmetic word problems. In both cases, the results showed that people relied on some type of structural knowledge, derived from semantic interpretation or general knowledge, to generate new instances. The results obtained here also exemplify how people can use underlying structure in generating new exemplars. There are, however, also some differences. The study on the generation of new animals showed that the same features people attend to when categorizing, including superficial and structural features, influence the generation of new instances. In our case, we have shown that the constraints governing subjects' behavior in reception tasks do not necessarily apply in the same way in production tasks. Inevitably, there will be differences between production tasks, depending on the domain of reasoning investigated. However, it is important to use production or generative tasks to identify whether or not and how the constraints that are identified with the use of reception paradigms apply to other reasoning situations.

The analogies produced by the subjects in this study resemble those studied in real-world reasoning environments in many ways (Blanchette & Dunbar, 1997; Dunbar, 1995, 1997). First, they were based on deep structural features and few were taken from domains similar to the target domain. Second, elements of the source were not all explicitly related to their counterparts in the target. In fact, many analogies were mostly implicit. The sources were described, but explicit connections were not drawn to the elements in the target. This was left up to the hypothetical audience. In real-world reasoning contexts too, analogies presented are not totally mapped out. Inferences about the mapping are left up to the audience. It is possible that sources are chosen so that these mappings are obvious.

Another interesting set of findings is related to the evaluation of analogies. We have shown that it is possible for superficial features, and for the degree of elaboration of the source, to play a role in the evaluation of analogies. In general, analogies selected as best contained a greater number of elements in the source and target. It is possible, in fact, that the analogies showing a greater degree of elaboration are perceived by subjects as being better, more convincing, or more persuasive. It is also possible that when an analogy is being proposed, the analogies that are perceived to be good are given more attention, and that the group or individual spends more efforts on the elaboration of those sources. This finding adds to other findings relating to the evaluation of analogies (Clement & Gentner, 1991; Holyoak & Thagard, 1995). The importance of structural soundness in the evaluation of analogies, as identified by Clement and Gentner (1991), is probably a general criterion used for analogies of all types. However, it is possible that more specific criteria related to the goal of the "analogizer" (Dunbar, 1995;

Holyoak & Thagard, 1989) are added for the evaluation of certain types of analogies—in this case, persuasive political analogies. In fact, it is also possible that the degree of elaboration in the source is closely linked to the structural soundness of an analogy. The greater the structural similarity, the greater the number of elements mentioned in the source and target.

Overall, this study has provided new insights into the ways in which people use analogical reasoning. The use of a production paradigm was motivated mainly by the discrepancies seen between the findings of standard psychology experiments and real-world investigations of analogical reasoning. This supports the necessity for cross-fertilization between real-world (in vivo) and laboratory investigations (in vitro) of cognitive phenomenon (Dunbar, 1993, 1995, 1997). Whereas laboratory experiments allow for more detailed and controlled investigations of the processes to be studied, real-world data can, and do, provide new knowledge and avenues to be investigated. These new insights, obtained from looking at real-world examples of the cognitive processes under study, can be brought back into the laboratory and scrutinized in a more controlled fashion, as was done in the present experiments. The interplay between these two types of investigation is crucial to further our understanding of human cognitive abilities.

REFERENCES

- BASSOK, M., CHASE, V. M., & MARTIN, S. A. (1998). Adding apples and oranges: Alignment of semantic and formal knowledge. *Cognitive Psychology*, *35*, 99-134.
- BLANCHETTE, I., & DUNBAR, K. (1997). Constraints underlying analogy use in a real-world context: Politics. In M. G. Shafto & P. Langley (Eds.), *Proceedings of the Nineteenth Annual Conference of the Cognitive Science Society* (p. 867). Mahwah, NJ: Erlbaum.
- BRUNER, J. S., GOODNOW, J. J., & AUSTIN, G. A. (1956). *A study of thinking*. New York: Wiley.
- CATRAMBONE, R. (1997). Reinvestigating the effects of surface and structural features on analogical access. In M. G. Shafto & P. Langley (Eds.), *Proceedings of the Nineteenth Annual Conference of the Cognitive Science Society* (pp. 90-95). Mahwah, NJ: Erlbaum.
- CLEMENT, C. A., & GENTNER, D. (1991). Systematicity as a selection constraint in analogical mapping. *Cognitive Science*, *15*, 89-132.
- DUNBAR, K. (1993, March). *In vivo cognition: Knowledge representation and change in real-world scientific laboratories*. Paper presented at the Society for Research in Child Development, New Orleans.
- DUNBAR, K. (1995). How scientists really reason: Scientific reasoning in real-world laboratories. In R. J. Sternberg & J. E. Davidson (Eds.), *The nature of insight* (pp. 365-395). Cambridge, MA: MIT Press.
- DUNBAR, K. (1997). How scientists think: On-line creativity and conceptual change in science. In T. B. Ward, S. M. Smith, & S. Vaid (Eds.), *Conceptual structures and processes: Emergence, discovery and change* (pp. 461-493). Washington, DC: American Psychological Association.
- ERICSSON, A. K. (1994). Expert performance: Its structure and acquisition. *American Psychologist*, *49*, 725-747.
- FARIES, J. M., & REISER, B. J. (1990, March). Terrorists and spoiled children: Retrieval of analogies for political arguments. In *Proceedings of the 1990 AAAI Spring Symposium on Case-Based Reasoning*. Menlo Park, CA: AAAI Press.

- FORBUS, K. D., GENTNER, D., & LAW, K. (1994). MAC/FAC: A model of similarity-based retrieval. *Cognitive Science*, *19*, 141-205.
- GENTNER, D., & MARKMAN, A. B. (1997). Structure mapping in analogy and similarity. *American Psychologist*, *52*, 45-56.
- GENTNER, D., RATTERMAN, M. J., & FORBUS, K. (1993). The roles of similarity in transfer: Separating retrievability from inferential soundness. *Cognitive Psychology*, *25*, 524-575.
- GICK, M. L., & HOLYOAK, K. J. (1980). Analogical problem solving. *Cognitive Psychology*, *12*, 306-355.
- HOLYOAK, K. J., & THAGARD, P. (1989). Analogical mapping by constraint satisfaction. *Cognitive Science*, *13*, 295-355.
- HOLYOAK, K. J., & THAGARD, P. (1995). *Mental leaps: Analogy in creative thought*. Cambridge, MA: MIT Press.
- HOLYOAK, K. J., & THAGARD, P. (1997). The analogical mind. *American Psychologist*, *52*, 35-44.
- KEANE, M. [T.] (1987). On retrieving analogues when solving problems. *Quarterly Journal of Experimental Psychology*, *39A*, 29-41.
- KEANE, M. T., LEDGEWAY, T., & DUFF, S. (1994). Constraints on analogical mapping: A comparison of three models. *Cognitive Science*, *18*, 387-438.
- SPELLMAN, B. A., & HOLYOAK, K. J. (1992). If Saddam is Hitler then who is George Bush? Analogical mapping between systems of social roles. *Journal of Personality & Social Psychology*, *62*, 913-933.
- SPELLMAN, B. A., & HOLYOAK, K. J. (1996). Pragmatics in analogical mapping. *Cognitive Psychology*, *31*, 307-346.
- WARD, T. B. (1994). Structured imagination: The role of category structure in exemplar generation. *Cognitive Psychology*, *27*, 1-40.
- WHARTON, C. M., HOLYOAK, K. J., DOWNING, P. E., LANGE, T. E., WICKENS, T. D., & MELZ, E. R. (1994). Below the surface: Analogical similarity and retrieval competition in reminding. *Cognitive Psychology*, *26*, 64-101.

NOTES

1. This problem involves two analogous stories. One involves a fortress that a general wants to attack. To achieve this goal, the general splits his army into small groups that converge onto the fortress. The other story is about a doctor with a patient that has a tumor. The tumor must be destroyed, but rays of the sufficient intensity would destroy not only the tumor but also the healthy tissues. Subjects are asked to find a solution to this second problem.

2. This term was first used by Bruner, Goodnow, and Austin (1956) in concept attainment experiments, to denote situations in which subjects are given instances of a concept and were then asked to determine what the concept was.

APPENDIX A

Introduction to the Issue of Zero Deficit

Important problems faced by most governments and societies are national debts and budgetary deficits. Most people agree that the debts governments have contracted over the years have become a very important problem which needs to be addressed. The budgetary crisis and its related problems occupy an important place in the media.

The basic problem is that governments spend more than they earn, which produces a deficit (the difference between earnings and expenses for a given year). The income of governments mainly comes from taxes and government-owned industries, and the spendings are devoted to services to the population and support for those who need it. When spending is greater than incomes, the governments have to borrow money either from banks, citizens (savings bonds, for example), or lenders outside the country. The total amount that is borrowed is the debt. Interest runs on this debt, and the government has to pay this interest or else it increases the amount of the debt. Annual deficits accumulate and produce an enormous debt which means more interest to pay, which means more expenses, therefore more borrowing, and so on. An important problem with substantial government debts is that, like for individuals, the borrower is vulnerable and dependent upon the lender. This problem is especially acute when money is borrowed from lenders outside the country.

Our governments in Canada face this problem. The Canadian federal debt amounts to approximately 620 billion dollars with an annual deficit of 24 billion dollars. The Ontario deficit for 1996-1997 is of approximately 8.1 billion dollars. The Québec deficit for the same period is 3.2 billion dollars with a total debt of 75 billion dollars. In Québec, for each dollar paid in taxes, 16 cents go to the management of the debt.

This state of affairs has led to a wide consensus that governments have to reduce deficits. Most people consider that public finances represent a very important problem that must be addressed. The goal, ideally, is to eliminate the deficit so the debt can at least stop growing.

This is where agreement stops and disagreement begins. Although a multitude of different positions exist, two basic views oppose one another in this war on the deficit. One view is that reducing the deficit should be the number one and most urgent priority. The opposing view considers that it is important to reduce the deficit, but not at any cost.

The governments of Ontario and Alberta have adopted the first view and made the reduction of the deficit their number one and most urgent priority. To attain this goal, they have had to announce or proceed to a drastic reduction in government

APPENDIX A (Continued)

spending. This has amounted to massive cuts in social assistance, university funding, health services, arts and culture, and school financing, to name a few. It has also meant the loss of many jobs in the public services. In Ontario, up to 1/3 of government jobs could be cut in order to reduce the government's expenses. This policy has resulted in a balanced budget for Alberta, and possibly a budgetary surplus for the year 1996-97 that could be allocated to the payment of the debt. In Ontario, the deficit will be eliminated by the year 2001 and taxes have been and will again be lowered to reach a total reduction of 30%.

A growing number of voices are starting to speak out against what they consider to be ruthless cuts to the social programs. Specifically, substantial portions of the union movement, of community groups, and of the student movement are saying that they agree with the reduction of the deficit, but not at any cost. "The elimination of the deficit cannot result in a social deficit." They advocate that cuts to social welfare and health services in particular will translate into social chaos. The human costs will be substantial. Some people argue that insecurity, poverty and violence will result from the elimination of social support systems. For these reasons, this movement argues that it is preferable to take more time to attain deficit zero but not to sacrifice the social security net along the way. An important movement has developed in Ontario, protesting against the cuts made by Premier Harris. In Québec, community groups are asking that the government stop targeting the most vulnerable groups (e.g., welfare recipients, single-parent families, etc.) when making budgetary cuts. Also, a strike movement by the government workers has been mentioned as a possibility in the last few days.

APPENDIX B Instructions

Anti-Zero Deficit Condition

You are a group of consultants and you have been hired by the National Association of Community Groups (NACG). The NACG is currently organizing an important nationwide campaign to sensitize both citizens and governments (federal and provincial).

The NACG is aware that analogies, comparing one thing to another, are a powerful tool for persuasion. Some analogies have already been used in a number of occasions in politics. For example, in 1992 President Bush wanted to convince the American population of the importance of the military intervention in the Persian Gulf. In his efforts to do so, he made extensive use of the analogy to World War II, comparing Saddam Hussein to Hitler.

The NACG has hired your group of consultants to come up with analogies that they could use to persuade the public and the government leaders that the reduction of the deficit is an important goal but that it should not be pursued at any cost. They do not agree with reducing the deficit if it means less support for those who need it and massive cuts to health, education and social assistance, less jobs, and more poverty. They want you to provide them with a list of analogies that support and illustrate this position.

You have twenty minutes to complete this task. You will be videotaped. You should brainstorm and come up with as many analogies as you regardless of how good you think they are. You write down briefly each analogy you come up with on a piece of paper, again, regardless of how good you think they are. Do not worry also about what is and what is not an analogy. Include metaphors, comparisons, everything that you think might be an analogy or something resembling. I will be in a separate room. I will come back when the time allowed for the task is up.

Pro-Zero Deficit Condition

You are a group of consultants and you have been hired by the National Association of Responsible Citizens (NARC). The NARC is currently organizing an important nationwide campaign to sensitize both citizens and governments (federal and provincial).

APPENDIX B (Continued)

The NARC is aware that analogies, comparing one thing to another, are a powerful tool for persuasion. Some analogies have already been used in a number of occasions in politics. For example, in 1992 President Bush wanted to convince the American population of the importance of the military intervention in the Persian Gulf. In his efforts to do so, he made extensive use of the analogy to World War II, comparing Saddam Hussein to Hitler.

The NARC has hired your group of consultants to come up with analogies that they could use to persuade the public and the government leaders that the reduction of the deficit should be our number one priority. They think that governments do not have a choice and that citizens have to make some sacrifices. They want you to provide them with a list of analogies that support and illustrate this position.

You have twenty minutes to complete this task. You will be videotaped. You should brainstorm and come up with as many analogies as you can, regardless of how good you think they are. You should write down briefly each analogy you come up with, again, regardless of how good you think they are. Do not worry about what is and what is not an analogy. Include metaphors, comparisons, everything that you think might be an analogy or something resembling. I will be in a separate room. I will come back when the time allowed for the task is up.

(Manuscript received May 29, 1998;
revision accepted for publication January 28, 1999.)