On Adaptation in Analogy:

Tests of Pragmatic-importance and Adaptability in Analogical Problem Solving

By

Mark T. Keane

Department of Computer Science,
University of Dublin, Trinity College,
Dublin 2, Ireland
FAX: +353-1-6772204
E-mail: mark.keane@cs.tcd.ie
WWW: http://www.cs.tcd.ie/www/mkeane/mkeane.html

Running head: ADAPTATION IN ANALOGY

ABSTRACT

When people use analogies to solve problems they form an *analogical mapping* between two domains of knowledge. This mapping may support inferences by analogy that suggest a novel solution to a problem. Several factors have been proposed to be important in selecting this mapping, from among several alternative mappings: structural factors (systematicity and structural consistency) and pragmatic factors (the exploitation of higher-order planning categories). We suggest another set of factors play a role in selecting mappings: adaptability factors. Specifically, if a mapped solution can be adapted easily to a problem, then it will be preferred over an alternative mapping that is less adaptable.

Two experiments are reported which test the effects of pragmatic and adaptation factors; using a novel technique in which the story analogue used has two alternative plans either of which can be used to solve an insight problem. In Experiment 1, these plans were varied in terms of their pragmatic-importance (success or failure) and their adaptability. In Experiment 2, the relative adaptability of plans was manipulated. The results suggest that there is little evidence for these specific pragmatic factors, but that adaptability plays a definite role in selecting an analogous plan. The findings suggest that most models need to be extended to include adaptation constraints.

ACKNOWLEDGEMENTS

This research was supported by The Open University, The University of Wales College of Cardiff and Trinity College Dublin. The author would like to thank Stuart Duff for his help in running some of the experiments and Ruth Byrne for reading and criticising several drafts of the paper. The paper has also benefited from discussions with Mike Brayshaw, Peter Cheng, Marc Eisenstadt, Dedre Gentner, Art Markman and two anonymous reviewers. Some of the results of Experiment 2 were presented at the *Sixteenth Annual Conference of the Cognitive Science Society*, Altanta, Georgia, USA.

Human thought excels in its capacity to deal with novelty and to overcome the unexpected and unfamiliar. In recent years, it has been demonstrated that the use of analogy is one strategy for dealing with novelty; novel problems can be solved by analogy to past experiences (see e.g., Burstein, 1986; Gentner, 1980, 1983; Gick & Holyoak, 1980; Keane, 1985, 1988). Gick & Holyoak (1980, 1983) demonstrated the phenomenon of analogical problem solving experimentally using Duncker's (1945) "radiation problem". The radiation problem involves a doctor's attempt to destroy a malignant tumour using rays. The doctor needs to use high-intensity rays to destroy the tumour but these high-intensity rays will destroy the healthy tissue surrounding the tumour. If the doctor uses low-intensity rays then the healthy tissue will be saved but the tumour will remain unaffected too. This dilemma can be solved by a "convergence solution" which proposes that the doctor send low-intensity rays from a number of different directions so that they converge on the tumour, summing to a high intensity to destroy it. However, only about 10% of subjects produce this solution if they are given the problem on its own.

Gick & Holyoak (1980) gave subjects an analogous story about a general attacking a fortress. The general was prevented from using his whole army to take the fortress because the roads leading to it were mined to explode if large groups of men passed over them. Therefore, he divided his army up into small groups of men and marched them along different roads to the fortress so that they converged on it. When subjects were given this analogous story to memorise and later asked if they could use it to solve the radiation problem, the rates of convergence solutions rose to about 80%. Hence, this experiment showed that people could solve problems by analogy.

Five main processes have been demonstrated to be important in analogical problem solving: representation, retrieval, mapping, adaptation and induction. Clearly, the way an analogue and a problem are mentally *represented* has an important impact on subsequent problem solving attempts by analogy (see e.g., Novick, 1988). If the

representation of a domain is partial or its causal structure has not been encoded adequately then analogies between it and another domain may be missed. The *retrieval* of an analogous case to solve a problem is harder if the domains are semantically-distant to one another (such as the General story and the radiation problem; see e.g., Gentner & Landers, 1985; Gentner, Rattermann & Forbus, 1993; Holyoak & Koh, 1987; Keane, 1987; Wharton et al., 1994). In fact, Gick & Holyoak (1980) found that subjects were unlikely to notice a semantically-distant story unless they were explicitly directed to do so. In *analogical mapping*, the core process in analogical thinking, the analogy is drawn between the two domains (e.g., between the story and the problem). After an analogy is mapped it may have to be *adapted* to make the analogous solution fit the target problem (see Keane, 1990; Novick & Holyoak, 1991). With respect to *induction*, Gick & Holyoak (1983) have shown that subjects can induce a generalisation from the correspondences between two analogous stories that facilitates subsequent problem solving.

In this study, we will be concerned with factors that impact the mapping and adaptation stages of analogising. Specifically, I address the issue that mapping and adaptation are intimately related, that the adaptability of a solution plays a role in the mapping process. In the remainder of this section, the current state of theory on mapping and adaptation is reviewed, extensions to current theoretical views are proposed and the novel paradigm used in the present experiments is introduced.

Factors Affecting Analogical Mapping

In retrospect, the analogy between the story and radiation problem seems obvious. When you know the analogy, it appears to be the only possible interpretation of the comparison between the two domains. However, many other interpretations are possible. The General story could be taken to suggest that "the doctor deploy the rays in a planned way, just like the general" or that "the doctor view the tumour as the enemy" or that "the doctor should use teams of junior doctors (troops) in the

operation". All of these interpretations of the analogy are valid given the information in the story but they do not appear to be as informative as the interpretation involving the convergence plan. Furthermore, when people use the story they tend to avoid these other interpretations. The central challenge for analogy theory has been to explain how analogical mapping is achieved and how people select such optimal interpretations.

This challenge has been answered by theoretical proposals that specify the processes in analogical mapping and the factors that affect the selection of one interpretation over other competing interpretations. In analogical mapping, people are said to map the conceptual structure of one domain (e.g., the story, termed the *base* or *source domain*) to another domain (e.g., the radiation problem, termed the *target domain*). During mapping, subjects *match* corresponding parts of the two domains. For example, subjects might match "the general's desire to attack the fortress" with "the doctor's desire to destroy the tumour". Typically, the mapping will also involve the *carry-over* of knowledge from the base domain to the target; knowledge which constitutes a set of *candidate inferences* by analogy. The candidate inferences are constructed from those parts of the base domain that are not present in the target domain. For example, in the story analogy, the information about *dividing the troops* and *the troops converging on the fortress* is carried over into the medical domain to generate candidate inferences about *dividing the rays* and *making the rays converge on the tumour*. It has been proposed that this interpretation of the analogy is selected during the mapping process because of certain structural and pragmatic factors.

Structural factors

The structural account maintains that the convergence interpretation is selected because it is part of a coherent set of matches between the two domains, it is the most *systematic* interpretation of the analogy between the two domains (see Gentner, 1980, 1983, 1989). That is, the convergence plan is causally-connected in a rich way to other parts

of the general story (e.g., to the general's intention to attack the fortress and the prevention of his whole army being used) and these parts of the story correspond to parts of the radiation problem (i.e., the doctor's intention to destroy the tumour and the prevention of high-intensity rays being used). This mapping is *structurally consistent* in that there are many one-to-one correspondences between the entities (objects and relations) in the story and in the problem (e.g., general->doctor, attack->destroy, troops->rays, fortress->tumour). In short, the convergence interpretation comes to be selected because it involves a systematic and coherent mapping between the two domains.

Structural factors such as these have been shown to be critical to analogical mapping (see e.g., Clement & Gentner, 1991; Gentner & Toupin, 1986; Gick & Holyoak, 1980; Holyoak & Koh, 1987; Keane, 1988). For example, Clement & Gentner (1991) have shown that structural systematicity acts as a selection constraint on mapping. When subjects were given a judgement task involving matches between two analogous stories, they preferred the match embedded in a system of relations shared by both stories to a match embedded in a system of relations which differed. Clement & Gentner have also shown that these effects hold for the selection of candidate inferences as well as matches. These proposals have also been substantiated by computational models which demonstrate that structural constraints are the key to the simulation of analogical mapping; see the Structure-Mapping Engine (SME; Falkenhainer, Forbus & Gentner, 1986, 1989; Forbus, Ferguson & Gentner, 1994), the Analogical Constraint Mapping Engine (ACME; Holyoak & Thagard, 1989, 1995) and the Incremental Analogy Machine (IAM; Keane, 1995; Keane & Brayshaw, 1988; Keane, Ledgeway & Duff, 1994).

In the experiments reported here, we hold structural factors constant and look for the effects of other factors (such as pragmatic factors). The effects of structural factors are well established, but less is known about other possible factors.

Pragmatic factors

Pragmatic factors have also been proposed to be important to selecting optimal interpretations of an analogy (Holyoak, 1985; Keane, 1985). Holyoak (1985) originally suggested that pragmatic factors alone were sufficient to account for analogy, in arguing that the convergence solution was selected because it contains goal-related information (see also Holland, Holyoak, Nisbett & Thagard, 1986). The reasoning behind this prediction was that goal-relevant information in the story should also be goal-relevant information in the problem and hence help to solve it. However, this strong position on pragmatic factors is untenable and must be diluted with structural factors to work. For example, pragmatic factors alone would not favour the selection of the convergence interpretation over an interpretation proposing that "that the doctor should use teams of junior doctors", because both of them are goal-relevant. Some other factor is required to select the former over the latter.

Holyoak has subsequently abandoned this strong position and come to view pragmatic factors as working in tandem with structural factors (see Holyoak & Thagard's, 1989, ACME model). The conception of pragmatic factors has also been re-defined. In ACME, one brand of pragmatic factors work off exploiting higher-order planning structures within a domain. The story and problem domains can be viewed as having higher-order divisions between its setting-part and a goal-relevant, solution-part. In ACME, pragmatic constraints act to map similar parts of the domains separately (e.g., the propositions in the story-setting are only matched with the propositions in the problem-setting). So pragmatic factors are seen as having a role in simplifying the mapping process.

Keane & Brayshaw (1988) had a similar conception of pragmatic factors in their IAM model (which also implemented structural factors). They proposed that higher-order categories in the story could be identified as being pragmatically-important. For example, if a story had planning categories corresponding to setting information, a

successful plan and a failed plan then pragmatically the former would be viewed as being more important than the latter. In Keane & Brayshaw's IAM model, the success or failure of a plan was used to identify the relevant part of the story to be mapped. So, the first interpretation selected would involve the successful plan. However, the unique effects of these pragmatic factors have never been substantiated empirically.

Holyoak & Thagard (1989) have another conception of pragmatic factors in which some aspect of the current context suggests that a particular mapping should be *presumed*. In ACME, this is modelled by providing extra activation to certain matches (see also Forbus & Oblinger's, 1990, implementation of pragmatic factors in SME). Evidence for the influence of presumed mappings has been found by Spellman & Holyoak (1992) using different analogies about the Gulf War and World War II or Vietnam. For example, if one of the mappings is presumed, by asking subjects the question "If Saddam is Hitler..." then various coherent mappings follow: Bush may be Churchill, in which case the US of 1991 is Great Britain.

In this paper, I test the predictions for pragmatic factors proposed by Keane & Brayshaw (1988). Holyoak & Thagard's ACME model makes the same predictions, although they never considered the successful-plan/failed-plan scenario. Previous studies have tended to use stories in which there is only one solution plan (e.g., see the General story). In this paper, we depart from this standard paradigm by using a story analogue that has two possible solution plans; either of which can be mapped to solve the target problem. One of these plans will succeed in the story and the other fail, hence according to Keane & Brayshaw the interpretation of the analogy (and solution to the problem) should always be based on the successful plan, because it is pragmatically more important.

Adaptability As a New Factor in Analogical Mapping

The second main focus of this paper is on the adaptation stage or, more precisely, the *adaptability of solutions*. Informally, adaptation is concerned with "how well an analogous solution can be fitted to the target problem". Even though most cognitive theories of analogy include an adaptation stage, little is known about adaptation and how it influences the course of analogical problem solving [In Artificial Intelligence, adaptation is a key stage of case-based reasoning (see e.g., Kolodner, 1993; Riesbeck & Schank, 1989; Smyth & Keane, 1994, 1995)]. In this paper, I test the novel hypothesis that adaptability factors can also affect the interpretation selected for an analogy.

Novick & Holyoak (1991) have demonstrated the importance of adaptation in solving mathematics problems by analogy to stories. They showed that even when subjects know the correspondences between two domains, they had difficulty applying the solution-plan in the base story (the mathematical procedure of finding the lowest common multiple) to a target problem. However, Novick & Holyoak's goal was to demonstrate the separation between the mapping and adaptation stages. They did not specify a theory of adaptation and neither did they view adaptability as playing a role in analogical mapping. Furthermore, it is not clear how their operational definition of adaptability in the mathematics domain, applies to the non-mathematical domains used in standard, analogy paradigms.

Table 2 About Here

Some insight into the processes involved in the adaptation stage can be gleaned from considering solution plans that differ in their adaptability. Table 1 shows three possible solution-plans to the General story that might be used to solve the radiation problem. Assume that the rest of the story and problem have been matched already and that these solutions are being used to create an analogous solution to the radiation problem (i.e., a

set of candidate inferences by analogy). Intuitively, these solutions differ in their adaptability, in how well they "fit" the radiation problem. The first solution is easy to fit and requires little adaptation; if we assume *death-laser* has already been matched to *ray*, then the laser solution-plan can be used directly in the problem domain. However, the second solution is harder to fit and less adaptable; the troops solution-plan mentions objects that are not present in the target problem (like *roads*), and the actions proposed (like *marching*) cannot be used directly in the problem. The analogiser has to adapt the "marching of the small groups of troops" to be "the sending of low-intensity rays", and adapt the "marching of troops along roads to the fortress" to be the "sending of rays along paths to the tumour". This adaptation involves creating new objects (like *paths*) and actions (like *sending*) and assessing whether these actions solve the problem. Finally, the third solution is the least adaptable. It requires all the adaptation required in the second solution, plus an action that does not seem to be at all useful in the radiation problem. A possible adaptation for "troops inspect their rifles" might be "rays inspect their intensities", but such an action is useless to the solution of the radiation problem.

These examples indicate that the adaptability of a solution is likely to depend on the entity-parallelism and entity-utility of solution parts. *Entity-parallelism* refers to whether the entities (i.e., objects and relations) in the solution have known corresponding objects in the target problem. If few of the base entities (especially objects) have parallels in the target (i.e., known matches) then the solution will be harder to adapt. *Entity utility* refers to how useful a base object or relation is to the solution of the target problem (e.g., the *march* relation is more useful than the *inspect* relation). If the utility of a relation or object is poor then the solution will be less adaptable [1].

The second main hypothesis tested in this paper is that adaptability plays a role in analogical mapping. That is, that adaptability influences the selection of an

interpretation for an analogy between two domains. Using our new paradigm, it is possible to construct a two-plan story in which the adaptability of the plans is different. If adaptability influences the selection of the interpretation, then the more adaptable plan should be chosen and appear as the solution to the target problem. In the next section, I introduce the specific story and problem materials used in this paradigm.

When an Analogue has Two Solution Plans

Consider Maier's (1931) "two-string" or "pendulum" problem and an analogous story about a fire (see Gick & Holyoak, 1983; Keane, 1988). The two-string problem involves a situation in which two strings, hanging from the ceiling of a room, have to be tied together. However, the two strings are too far apart for a person to reach one string while holding the other. Traditionally, subjects propose a variety of solutions to this problem including a *stick* solution (use a stick to bring the string that is furthest away, within reach) and a *swing* solution (swing one of the strings so that it comes within reach while holding the other string).

Consider a putative analogue to this problem, consisting of a story about a towering inferno. In this story, several people are trapped on the upper stories of a burning skyscraper and can only be rescued using a helicopter. So, the helicopter flies as close to the building as possible and lets down a rope. Initially, the people grab hold of the rope using a stick to bring it within reach and are carried to the ground by the helicopter. But, this plan fails because it does not evacuate the people fast enough and it will not rescue all the people in time. So, another plan is attempted in which people swing on the rope from the burning building to a nearby building. This plan is executed much faster than the other one and all the people are rescued.

If we wish to solve the two-string problem by analogy to this story, we are faced with two possible solution plans to map to the target problem. One mapping suggests the analogous solution of using a stick to reach the string (using the first plan) and the other

mapping suggests the analogous solution of swinging one of the strings to bring it within reach (using the second plan). In this case, either of these plans may be mapped to solve the problem. This competition between solution plans in this situation allows us to look at the factors affecting the selection of one interpretation over another, by varying aspects of the solution plans.

First, predictions based on pragmatic-importance can be tested by varying the success or failure of the plan used in the problem (see Keane & Brayshaw, 1988). Second, the effects of adaptability can be examined by using solutions that differ in their adaptability. In the Fire story, the stick plan is less adaptable than the swing plan, because the former has less entity-parallelism. The stick plan proposes that a stick be used to reach one of the strings, but a stick which is not present in the statement of the two-string problem. The swing plan proposes various actions, like swinging, without the need to introduce a new object. As such, the swing plan should be easier to adapt and be chosen as the interpretation for the analogy.

EXPERIMENT 1: PRAGMATIC-IMPORTANCE & ADAPTABILITY

In the first experiment, subjects were given a story to memorise and were then asked to use this story to solve a problem by analogy. We used the two-string problem and the Fire story. The Fire story always had the same general structure consisting of setting information, an account of a failed rescue plan followed by a successful rescue plan. In all versions of the story the setting information was the same, but the specific plan used for the successful and failed plan was varied (see Appendix A). In roughly half the stories, the successful plan was the swing plan and the failed plan was the stick plan (see Story 1, Appendix A), whereas in the other half of the stories, the stick plan was the successful plan and the swing plan was the failed plan (see Story 2, Appendix A).

In drawing the analogy, the setting information in the Fire story and the two-string problem will be matched: for instance, both domains involve propositions about getting

from one location to another location to achieve some goal. Both plans constitute collections of candidate inferences which follow from matched propositions in the setting. So both plans form alternative interpretations of the analogy, one of which will be selected as the analogous solution to the problem. The critical measure was the solution subjects develop for the problem by analogy to the story (classified as stick, swing or neither solutions).

If pragmatic-importance is a factor in selecting one interpretation over another, then we should find differences in the frequencies of stick and swing solutions being produced as a function of the changes in the pragmatic-importance of the stick and swing plans. On the whole, if adaptability plays a role in selecting interpretations then we should see more solutions based on the highly-adaptable swing solution than on the less adaptable stick solution. Similarly, overall, if pragmatic factors play a role then we should see more solutions based on the successful plan than on the failed plan.

Method

Subjects

Eighty-five subjects from the subject panels of the Open University and the University of Wales College of Cardiff were randomly assigned to the two conditions in the experiment. Eight of these subjects were excluded from the data analysis because they had either encountered the problem before (3 subjects) or they misunderstood the task (5 subjects; i.e., they proposed further solutions to the Fire story).

Materials

The materials consisted of two versions of the Fire story (see Appendix A) and the following version of the two-string problem:

> Imagine you are in a room in which two strings are hanging from the ceiling. Your task is to tie the ends of the two strings together. However, when you

take hold of one string and try to get the other string you find that it is too far away to touch.

This problem statement was followed by a diagram depicting the problem situation.

Design

The design was a between-subjects design involving two conditions -- the stick-pragmatically-important and swing-pragmatically-important conditions -- where the independent variable was the pragmatic-importance of a plan [2]. In the stick-pragmatically-important condition, the stick-plan was the successful plan while the swing plan was the failed plan. In the swing-pragmatically-important condition the swing-plan was the successful plan while the stick-plan was the failed plan. The dependent variable was the type of solution produced: a stick solution, a swing solution or neither (see scoring).

Procedure

Subjects received written instructions, in booklet form, telling them that the experiment consisted of two separate tasks: a memory/comprehension task and a problem solving task. In the memory/comprehension task they had three minutes to read and memorise the Fire story. They were told that after the three minutes they would have to recall it in writing. These recollections were analysed later using a propositional-analysis to determine how well the story had been encoded. Each subject in the experiment received one of the two versions of the Fire story.

After subjects had recalled the story in writing and it was taken away, they were given the two-string problem to read. They were asked "If you were to apply the story about the fire to this problem what answer would it suggest to you (irrespective of whether you think the answer is a good or a bad one) " [3].

Subjects were tested in small groups of varying sizes and each experimental session took between 20 and 30 minutes.

Scoring

In analysing subjects' responses the interest was in the solution that was most immediately suggested to them by the story. Therefore, when subjects produced more than one solution, only the first solution produced was considered. In fact, 90% of subjects produced only one solution (10% produced two solutions). When subjects' solutions made use of one or other of the provided plans, they were categorised as being: (i) *stick solutions* (which suggested using a stick to bring one string closer) (ii) *swing solutions* (which proposed swinging one string and catching it on the up-swing while holding the other) or *neither* (some other type of solution).

Table 2 About Here

Results & Discussion

The pattern of subjects' solutions to the problem suggest that both pragmatic-importance and adaptability have an effect on solving problems by analogy. Of the 77 subjects in the experiment, 87% (67 subjects) generated a solution based on one or other of the story's plans, whereas 13% (10 subjects) generated a solution based on neither plan (see Table 2).

In the swing-pragmatically-important condition more swing solutions (72%) were produced than in the stick-pragmatically-important condition (42%), and in the stick-pragmatically-important condition more stick solutions were produced (43%) than in the swing-pragmatically-important condition (15%) ($chi^2(2) = 8.52$, $p < .01$). This suggests that changing the pragmatic-importance of a plan increases or decreases the likelihood that it will be selected as the interpretation for the analogy and used to solve

the problem. However, as we shall see later, there are some grounds for caution regarding this result.

To achieve a more fine-grained analysis, the solutions of the 67 subjects who produced a stick or swing solution were analysed, collapsing across the two conditions in the experiment. In this analysis, subjects' solutions were categorised according to theoretically-derived categories based on pragmatic-importance and adaptability. Four categories were used: (i) low-pragmatic/low-adaptable (a stick solution produced from a failed stick-plan), (ii) low-pragmatic/high-adaptable (a swing solution produced from a failed swing-plan), (iii) high-pragmatic/low-adaptable (a stick solution produced from a successful stick-plan), (iv) high-pragmatic/high adaptable (a swing solution produced from a successful swing-plan).

Table 3 About Here

Table 3 shows the distribution of subjects' solutions in these theoretically-defined solution categories. The results are revealing about the predictive nature of the pragmatic-importance and adaptability. According to the pragmatic predictions, more subjects should produce solutions based on the successful plan (high in pragmatic-importance) than on the failed plan (low in pragmatic-importance). This prediction is supported in that almost twice as many subjects select the successful plan over the failed plan; 66% (44) as opposed to only 34% (23), respectively ($chi^2(1) = 6.58$, $p < .05$). This result indicates that there is a pronounced tendency to make use of the pragmatically-important solution. If the adaptability of a solution plan plays a key role in selecting between alternative plans then we should find that more subjects produce solutions based on the highly adaptable, swing plan than on the less adaptable stick plan. This prediction was confirmed; 67% of subjects (45 out of 67) produced solutions based on the high-adaptable plan as opposed to 33% (22 out of 67) producing solutions based on the low-adaptable plan ($chi^2(1) = 7.89$, $p < .01$). This demonstrates an effect of adaptability that has not been recognised in previous research.

There was no evidence of an interaction between pragmatic-importance and adaptability.

Problems With the Results

At first glance, these results suggest that both pragmatic-importance and adaptability have an effect on the production of a solution by analogy. However, there are reasons to be cautious about each of these conclusions. First, the pragmatic finding may have been due to recency effects. Second, the adaptability results hinge on the production of a particular solution, the swing solution.

The results show that subjects use the successful plan more often than the failed plan. The pragmatic account says that this selection occurs because the successful plan is more pragmatically-important than the failed plan. However, all the versions of the story in the experiment had the same "setting--> failed plan -->successful plan" structure. Hence, subjects may have used the successful plan because it was always the most-recently memorised plan (a recency effect). In Experiment 2, we controlled for recency by ensuring that half the story materials had a "setting --> failed plan --> successful plan" structure and the other half had a "setting --> successful plan --> failed plan" structure.

The results also show that subjects are more likely to produce swing solutions than stick solutions. I have argued that this occurs because the swing solution is more adaptable than the stick solution, because it does not involve any new objects whereas the stick solution requires the introduction of a stick into the problem domain. However, it could well be the case that subjects were spontaneously generating swing solutions, without using the story at all. To test this we had a separate control group of 21 subjects solve the two-string problem without a story analogy, after an irrelevant task of an equivalent length of time (see Table 2). If spontaneous-generation is responsible for the results found then we should find that this control condition is

predictive of the pattern of results found in the experimental conditions. However, Table 2 shows that this is clearly not the case as neither experimental condition parallels the control condition: the stick-pragmatically-important condition is reliably different to the control ($chi^2(2) = 18.48$, $p < .001$), as is the swing-pragmatically-important condition ($chi^2(2) = 14.55$, $p < .001$).

Accepting this conclusion, there is still a niggling worry about the high rate of swing solutions in the control (43%), relative to the experimental condtions (42% and 73% for the stick-pragmatically-important and swing-pragmatically-important conditions, respectively). One counter to this worry is that the control is a very conservative one. A more accurate control would be to give subjects a story that did not contain a swing plan, ask them to use it by analogy to solve the problem and then look at the rates of swing solutions produced [4]. This would parallel the experimental conditions more closely and give us a purer measure of the spontaneous-production of swing solutions in this paradigm (i.e., subjects producing swing solutions in spite of the plans given in the story). Just such a condition was introduced into Experiment 2. Furthermore, to allay any remaining worries about adaptability effects being based on the swing solution, in Experiment 2, the stick solution was used as the most adaptable solution. In Experiment 2, the perceived adaptability of the stick solution was manipulated by pairing it with solution plans that were more or less adaptable to it. As we shall see, this provides us with a much tighter test of adaptability.

EXPERIMENT 2: TESTING RELATIVE ADAPTABILITY

Experiment 2 was concerned with testing four main issues. First, the materials were counterbalanced to control for recency effects. Second, a tighter test of adaptability was performed. Third, the issue of spontaneous production of swing solutions was examined. Finally, I gathered ratings to substantiate my intuitions on the adaptability of different plans.

First, as we saw earlier, one possible explanation for effects of pragmatic-importance in Experiment 1 was that subjects selected the successful plan over the failed plan because it was always the most recently-memorised plan in the story. In this experiment, to control for such recency effects between the story plans the order of the plans in different versions of the story were reversed: in half of the stories used the story structure was "setting --> successful plan --> failed plan" while in the other half it was as in Experiment 1, "setting --> failed plan --> successful plan" (see e.g., Stories 1 and 2 in Appendix B). If the pragmatic predictions are correct, overall, we should see more subjects using the successful plan than the failed plan.

In Experiment 1, the adaptability effects found rest on the use of the swing plan as opposed to the stick plan; these effects are questionable because some other extraneous factor connected to the swing plan, may be the source of these results. We need a stronger test of adaptability, which shows that the perceived adaptability of the *same* solution-plan can be systematically altered depending on the plan with which it is paired. Clearly, it would be a good idea to demonstrate this effect of "relative adaptability" for a plan other than the swing plan.

To meet these constraints we used the stick plan as the focus for this experiment and paired it with either the swing plan or a new, lasso plan. The *lasso plan* is similar to the stick plan, with the exception that the rope is reached by lassoing it with another piece of rope. Even though this plan is very similar to the stick plan, it is less adaptable for several reasons. The lasso plan has poor entity-parallelism (it requires that an object be found to correspond to the *lasso*) and poor entity-utility (a non-rigid object like a lasso is not a very effective means for reaching in this context). Thus, in the stick-high-adaptable condition where the stick and lasso plans are paired, the stick plan should appear to be highly-adaptable relative to the lasso plan. Whereas in the stick-low-adaptable condition where the stick and swing plans are paired, the stick plan should appear less adaptable relative to the swing plan. The adaptability predictions

are that subjects should prefer the stick plan when it is high-adaptable relative to its paired-plan, than when it is low-adaptable relative to the paired-plan.

Third, the use of the "lasso-stick" stories also has the added benefit of allowing us to assess the spontaneous production of the swing solution. This condition provides us with a better control than that used in Experiment 1. As such, it should allow us to assess whether the high rates of swing solutions in Experiment 1 reflected a baseline of spontaneous production of the swing solution or the selection of a highly-adaptable plan.

Finally, to substantiate our intuitions about the relative adaptability of plans, we asked subjects to rate how good the swing, stick and lasso solutions were as solutions to the two-string problem. If our intuitions are correct, then the swing solution should be rated as being better as a solution than the stick solution and the stick solution should be rated above the lasso solution.

Method

Subjects

Sixty-three undergraduates in Computer Science at Trinity College Dublin were randomly assigned to the conditions of the experiment. Four of these subjects were excluded before data analysis because they had insufficient memory of the plans in the story. Twenty more undergraduates were used in a separate control condition for the rating-task.

Materials

The materials consisted of four versions of the Fire story and the two-string problem. Half of the stories used in the experiment had a "setting --> successful plan --> failed plan" structure while the other half had a "setting --> failed plan --> successful plan"

structure. Within each of these sets, half the stories were swing-stick versions of the story and half were lasso-stick versions of the story (see Appendix B).

The materials for the rating task consisted of a booklet containing descriptions of the swing, stick and lasso solutions to the two-string problem. Under each solution was a seven-point scale, ranging from 1 (very bad) to 7 (very good). Written instructions informed subjects that they were to rate how good they thought the solution was, as a solution to the problem. Each solution was printed on a separate sheet. The order of presentation of solutions in the rating task was randomised.

Design

The design was a between-subjects design involving two conditions -- the stick-low-adaptable and stick-high-adaptable conditions -- where the independent variable was the relative adaptability of the stick plan. In the stick-low-adaptable condition, the stick-plan was paired with the swing-plan. In the stick-high-adaptable condition the stick-plan was paired with the lasso-plan. Each of these conditions were counterbalanced for recency effects. The dependent variable was the type of solution produced: a stick solution, a swing solution, a lasso solution or neither (see scoring).

Procedure

The experiment had three distinct stages: the memory/recall task, the problem solving task and the rating task. All the booklets handed out at each stage were collected before those of the next stage were handed out. The first two tasks were as in Experiment 1.

In the rating task, subjects were asked to rate the goodness of a solution plan as a solution to the problem. The experimental groups received the rating booklet after the memory/recall and problem solving tasks. The control group subjects received an irrelevant task of equivalent length to the memory/recall and problem solving tasks and then were given a description of the two-string problem and asked to complete the

rating booklet. This control group was included to rule out any influences on the rating task that may have emerged from seeing the story and using it to solve the problem.

Subjects were tested in two groups of roughly equal numbers and each experimental session took between 20 and 30 minutes.

Scoring

Subjects' responses were classified as being either swing, stick or lasso (which suggested using a piece of rope/string to lasso the string that was furthest away) or neither of these solutions. As before, the majority of subjects produced a single solution to the problem (20% of subjects produced two solutions). The analysis was based on the first solution produced.

Table 4 About Here

Results & Discussion

The pattern of subjects' solutions revealed the effects of relative adaptability but showed little evidence of pragmatic effects. The pattern of goodness ratings confirmed our intuitive ranking of the adaptability of different plans. Of the 59 subjects in the experiment, 81% (48 subjects) generated a solution based on one or other of a story's plans, whereas 19% (11 subjects) generated a solution based on none of the presented plans in a story.

The pattern of solution-types produced in the stick-high-adaptable condition was reliably different to that found in the stick-low-adaptable condition ($chi^2(3) = 16.32$, $p = .001$). Notably, there were reliably more stick solutions produced in the stick-high-adaptable condition (67%) than in the stick-low-adaptable condition (28%; $chi^2(1) = 9.03$, $p = .002$).

Table 5 About Here

As before, to achieve a more fine-grained view I analysed the solutions of the 48 subjects who produced a solution based on the plans presented in the story, collapsing across the two conditions in the experiment (see Table 5). Solutions were categorised according to the four theoretically-derived categories: low-pragmatic/low-adaptable (a lasso solution produced from the failed, low-adaptable lasso plan), low-pragmatic/high-adaptable (a swing solution produced from the failed, high-adaptable swing plan), high-pragmatic/low-adaptable (a stick solution produced from the successful, stick plan paired with the swing plan), high-pragmatic/high adaptable (a stick solution produced from the successful stick plan paired with the lasso plan).

First, with respect to pragmatic-importance, marginally more subjects used the successful plan (58%; 28) over the failed plan (42%, 20) but this is not a reliable difference ($chi^2(1) = 1.3$, $p > .05$). So, when we control for recency the pragmatic effects disappear. In contrast, solutions based on the high-adaptable plan were produced more often (77%) than solutions based on the low-adaptable plan (23%; $chi^2(1) = 14.88$; $p < .0001$). There was no significant interaction between pragmatic-importance and relative-adaptability. These adabtability effects, which replicate the findings of Experiment 1, indicate that the adaptability of a plan influences the selection of one interpretation over another in analogical problem solving.

If we compare the pattern of solutions in the control condition, run as part of Experiment 1 (see Table 4), with the pattern of solutions found here, it is clear that neither experimental condition parallels the control condition. In this experiment, we also have a more accurate control for assessing the baseline spontaneous-generation of swing solutions, since in the lasso-stick versions of the story no swing plan was given. When subjects are given these stories, the baseline found for the spontaneous production of swing solutions was 13% (4 subjects). This is considerably lower than the rate of 58% (17 subjects) which we see in conditions receiving the swing plan as part of the story ($chi^2(1) = 13.19$, $p < .001$). Hence, in this experiment and

Experiment 1, we can be confident that subjects are using the swing plan to solve the problem, rather than simply spontaneously producing it without reference to the Fire story. Therefore, this finding should clear any remaining worries about the adaptability effects in Experiment 1 being based on the spontaneous production of the swing solution.

Figure 1 About Here

Ratings Results

Subjects were also asked to rate the relative goodness of the three solutions, as solutions to the two-string problem, to provide independent support for the proposals about the relative adaptability of these solutions. Overall the goodness ratings confirmed these intuitions (see Figure 1). In a 2 (experiment v control) x 3 (solution type) analysis of variance, with repeated measures on the last factor, of the ratings there were no reliable main effects of condition and no reliable interactions. However, as expected, there was a reliable main effect of solution type ($F(2, 154)= 38.58$, $p < .0001$). Planned comparisons, using Newman-Keuls tests, between solution types revealed that the goodness of the swing solution ($m = 5.09$) is rated as being significantly different to the goodness of the stick solution ($m = 4.17$, $p < .01$) which in turn is rated more highly than the goodness of the lasso solution ($m = 2.81$, $p < .01$).

Summary

The results of this experiment also confirm the generality of the findings in Experiment 1 for adaptability but overturn the findings on pragmatic effects. In this experiment, the same plan (i.e., the stick plan) is conceived of as being more or less adaptable depending on the plan with which it is paired (either the swing or lasso plan). When the stick plan is high-adaptable relative its competing plan (i.e., the lasso plan), subjects tend to select it 77% of the time, but when it is low-adaptable relative to its

competing plan (i.e., the swing plan) it is selected only 23% of the time. Thus, this experiment provides strong evidence that subjects assess the relative adaptability of one mapping interpretation relative to another and then select that mapping that has been evaluated as the most adaptable.

GENERAL DISCUSSION

The present study reveals a number of interesting and novel findings which have significant implications for current theories and models of analogy. These findings have emerged from a novel experimental manipulation that enables us to assess the factors affecting the selection of one interpretation (solution plan) of an analogy over another in a problem solving context. In particular, the study has two main findings. First, that there is little evidence for effects based on exploiting the pragmatically-important, higher-order planning structure in a domain. Second, that the adaptability of a solution plays an important role in selecting one analogous interpretation over another.

These results raise a number of issues for analogy research that are discussed in the following sections. First, I consider the implications of these findings for the role of pragmatic factors in analogical mapping and for the computational modelling of pragmatic constraints. Second, the explanation of adaptability effects by structural factors, like structural consistency or Forbus & Oblinger's goal-marking mechanism, is considered. Third, I relate the implications of the adaptability findings for current theory and computational modelling of analogy.

Implications of Findings for Pragmatic Factors

Previous evidence has supported one particular set of pragmatic factors (e.g., Spellman & Holyoak's, 1991, results support the effects of presumed mappings). Here we have examined a different subset of pragmatic factors; namely, the use of higher-order planning structures in a domain. The fact that predictions for these pragmatic

influences have failed, does not contradict findings based on other pragmatic factors. However, these results suggest that people do not use these specific planning structures to direct the course of analogical problem solving. Hence, the functionality built into some models to handle these structures is *not* required. Consider what this means for the two models that have suggested that such structures be used (i.e., ACME and early versions of IAM).

ACME and Keane & Brayshaw's IAM have the functionality to deal with higher-order structurings of a domain and to exploit them in drawing analogies. Holyoak & Thagard (1989) designed ACME with a facility for partitioning domains into different fields, and then just mapping the contents of like fields with like fields. ACME has used such higher-order structurings in the mapping of other examples in the literature (e.g., the radiation problem). In ACME, the Fire story could be encoded as having the following fields: SETTING, SUCCESSFUL-PLAN, FAILED-PLAN. However, there is no specific and unique empirical evidence to support the proposal that people exploit such plan categories in their analogising. Indeed, the present experiments provide negative evidence for the use of such categories [5]. Similar criticisms can be made of Keane & Brayshaw's IAM which had the functionality to deal with such higher-order categories like SUCCESSFUL-PLAN and FAILED-PLAN (later versions of IAM do not have this feature, see Keane et al., 1994). Even if people form such categories, the evidence suggests that they do not exploit them when solving problems by analogy. [These criticisms do not apply to Forbus & Oblinger's (1990) pragmatic-marking mechanism for matches in SME, because it does not rest on the use of higher-order categories in a domain (see next section for a discussion of this mechanism).]

Alternative Explanations of Adaptability Effects

A critical question that remains is whether the adaptability effects found can be explained by other mechanisms posited already in the literature. There are two such candidate explanations, one that relies on the structural factor of structural consistency

and another that uses the pragmatic factor of pragmatic-marking (Forbus & Oblinger, 1990).

Structural consistency refers to the whether a match involves a one-to-one correspondence and preserves relational structure. This is very like the idea of the entity-parallelism identified here to underlie adaptability. In Experiment 1, the adaptability effects found hinge on there being an extra, unmatched object in the stick plan than in the swing plan. So one interpretation of these results, might be said that the stick plan is less structurally-consistent than the swing plan. However, this account does not work for the results of Experiment 2. In it we find that the stick plan is preferred over the lasso plan even though both have the same lack of structural consistency, that both involve an extra, unmatched object (stick or lasso).

This failure of structural-consistency to explain the results of both experiments is enough to rule it out as an alternative account, but a further problem must be mentioned. That is, structural consistency does not even predict the differences found in Experiment 1 (that is, a difference between the swing and stick plans as sets of candidate inferences). In SME, the structural goodness of one interpretation is computed before candidate inferences are even constructed (the structural-goodness score is based on structural consistency and systematicity). Hence, any entities added to the base domain that only contribute to the candidate inference set (e.g., the stick in the stick solution), do not change the structural goodness score for a mapping interpretation. In SME structural-consistency only applies to matches not to candidate inferences.

Structural-consistency and entity-parallelism are related in that they are both concerned with one-to-one correspondences between entities in two domains; they differ in that the former deals with correspondences established in matching the main body of the analogy, whereas the latter concerns correspondences that have to be established for the unmatched, candidate inferences of the analogy. They are different things, a fact that

is underlined by the asymmetric relationship between them. Structural consistency establishes the degree of entity-parallelism, but when parallel entities are found in the target domain to complete entity parallelism, these new correspondences may violate previously-established structural consistencies.

The second candidate explanation involves the use of pragmatic-marking. Forbus & Oblinger (1990) proposed a pragmatic mechanism in SME that involved marking matches (or match hypotheses) that are pragmatically-important or specified in the target goal. The target goal in the two-string problem is to "bring the string that is furthest away within reach" (see Keane, 1989). For the materials in Experiment 1, this mechanism would favour any interpretation that involved a match to "bring-within-reach" or "string". Both the stick and swing plans involve a *bring-within-reach* relation so there is a *bring-within-reach ->bring-within-reach* match. This relational match would establish object-matches between *stick->string* (in the stick plan) and *rope->stick* (in the swing plan). So, interpretations based on either plan would both be pragmatically-marked. Hence, to the best of my understanding, pragmatic-marking would not predict the result found in Experiment 1 (a similar argument holds for Experiment 2).

In short, no current mechanism in literature accounts the adaptability effects found here. It could be argued that some variant of structural-consistency for candidate inferences (what we have called entity-parallelism) could be proposed but this misses the point of the work. These experiments suggest a new set of constraints on analogical mapping; the effects should be explained, not explained away. In the next section, I consider the character of such an explanation.

Putting Adaptability into Theory and Model ?

The present results suggest that a computational level account of analogical thinking should include *adaptation constraints*; that the adaptability of alternative mappings

needs to be computed. Current theory admits that various informational constraints (i.e., structural, similarity and pragmatic constraints) and behavioural constraints (e.g., working memory limitations) influence analogical mapping (see Keane et al., 1994). The informational constraints on analogical mapping should now be expanded to include constraints that evaluate the adaptability of alternative mappings.

This evaluation could rely purely on local tests of the candidate inferences being made: for instance, by noting whether the inferences suggest new objects that are absent from the original specification of the target problem (i.e., assessing the lack of entity-parallelism). However, such local evaluation does not seem sufficient. It would result in a ranking of the stick solution as being less adaptable than the swing solution (because the former requires an extra object) but would characterise the lasso and stick solutions as being equally adaptable (because both require an extra object). A more plausible assumption is that the evaluation is performed by simulating the proposed solution (i.e., assessing entity-utility as well); the action of swinging a rope-like object or reaching for it with a stick is simulated in the target domain. Only a detailed simulation of this type predicts the adaptability ordering supported by the results of these experiments; an ordering that is reflected in subjects' ratings of the solutions (i.e., swing better than stick, and in turn stick better than lasso).

Implications for Computational Models

At present, none of the models in the literature can deal with the adaptability effects found here, because they do not have adaptation components. In Artificial Intelligence, Falkenhainer's (1987) PHINEAS system carries out "verification-based analogical learning" using Qualitative Process Theory descriptions of physical situations (see Forbus, 1984). PHINEAS uses the Structure-Mapping Engine as a sub-component and has a method for testing candidate inferences where no specific knowledge of the target domain is available. However, people know quite a bit about the domain of the two-string problem, so the verification carried out in PHINEAS is

likely to be different to the adaptation found here. Furthermore, PHINEAS has no mechanism for evaluating the relative adaptability of mappings. A better candidate model might be found in Falkenhainer's (1990) *contextual structure-mapping* idea, as it provides an architecture for integrating background knowledge about a target domain into the analogy process.

Clearly, adaptation components need to be developed for cognitive models of analogy. The results of this modelling venture should be interesting because the different architectures of these models are likely to require the integration of an adaptation component in different ways. Indeed, these differences that may give rise to new empirical predictions that distinguish between the models.

APPENDIX A

Examples of Fire Story Used in Experiment 1

1) Stick-Swing Story (with Stick as Failed Plan and Swing as the Successful Plan)

A large number of people were trapped on the upper stories of a burning skyscraper. The authorities decided that the only way of saving the people would be to use a helicopter. So, the air force were called and soon a helicopter arrived on the scene.

The helicopter positioned itself near the burning building and a rope was lowered to where the people were. One of the people then grabbed hold of the rope using a stick to bring it within reach. Then, holding onto the rope he was carried to the ground by the helicopter. However, it soon became clear that this method was too slow and would not rescue all the people before the fire reached them. So, another method was tried.

The helicopter positioned itself between the burning building and another nearby building. The rope was thrown to the people in the burning building and one of them grabbed hold of it. Then, holding onto the rope this person swung from the burning building to the nearby building. Using this swing method all the people were able to evacuate the burning building before the fire reached them.

2) Swing-Stick Story (with Swing as Failed Plan andStick as the Successful Plan)

A large number of people were trapped on the upper stories of a burning skyscraper. The authorities decided that the only way of saving the people would be to use a helicopter. So, the air force were called and soon a helicopter arrived on the scene.

At first the helicopter positioned itself between the burning building and another nearby building. A rope was thrown to the people in the burning building and one of them grabbed hold of it. Then, holding onto the rope this person swung from the burning building to the nearby building. However, it soon became clear that this method was too slow and would not rescue all the people before the fire reached them. So, another method was tried.

The helicopter positioned itself near the burning building and a rope was lowered to where the people were. One of the people then grabbed hold of the rope using a stick to bring it within reach. Then holding onto the rope he was carried to the ground by the helicopter. Using this carrying method all the people were able to evacuate the burning building before the fire reached them.

APPENDIX B
Examples of Fire Stories Used in Experiment 2

1) Swing-Stick Story (Setting--> Failed Plan --> Successful Plan)

A large number of people were trapped on the upper stories of a burning skyscraper. The authorities decided that the only way of saving the people would be to use a helicopter. So, the air force were called and soon a helicopter arrived on the scene.

At first, the helicopter positioned itself between the burning building and a nearby building. A rope was thrown to the people in the burning building and one of them grabbed hold of it. This person tied himself onto the rope. Then, he jumped clear of the building and swung on the rope from the burning building to the nearby building. However, it soon became clear that this method was too slow and would not rescue all the people before the fire reached them. So, another method was tried.

The helicopter positioned itself near the burning building and the rope was lowered down to where the people were. One of the people then quickly grabbed hold of the rope by hooking it with a stick. Then this person tied himself onto the rope and was carried to the ground. Using this method all the people were able to evacuate the burning building before the fire reached them.

2) Swing-Stick Story (Setting--> Successful Plan --> Failed Plan)

A large number of people were trapped on the upper stories of a burning skyscraper. The authorities decided that the only way of saving the people would be to use a helicopter. So, the air force were called and soon a helicopter arrived on the scene.

The helicopter positioned itself near the burning building and a rope was lowered down to where the people were. One of the people then quickly grabbed hold of the rope by hooking it with a stick. Then this person tied himself onto the rope and was carried to

the ground. Using this method all the people were able to evacuate the burning building before the fire reached them. However, before this method was used, another method had been tried.

At first, the helicopter had positioned itself between the burning building and a nearby building. The rope was thrown to the people in the burning building and one of them grabbed hold of it. This person tied himself onto the rope. Then, he jumped clear of the building and swung on the rope from the burning building to the nearby building. However, it soon became clear that this method was too slow and would not rescue all the people before the fire reached them. So, the other method was tried.

3) Lasso-Stick Story (Setting--> Failed Plan --> Successful Plan)

A large number of people were trapped on the upper stories of a burning skyscraper. The authorities decided that the only way of saving the people would be to use a helicopter. So, the air force were called and soon a helicopter arrived on the scene.

At first, the helicopter positioned itself near the burning building and a rope was let down. One of the people then grabbed hold of the rope by lassoing it with another piece of rope. This person then tied himself onto the rope and was carried to the ground by the helicopter. However, it soon became clear that this method was too slow and would not rescue all the people before the fire reached them. So, another method was tried.

The helicopter positioned itself near the burning building and the rope was lowered down to where the people were. One of the people then quickly grabbed hold of the rope by hooking it with a stick. Then this person tied himself onto the rope and was carried to the ground. Using this method all the people were able to evacuate the burning building before the fire reached them.

4) Lasso-Stick Story (Setting--> Successful Plan --> Failed Plan)

A large number of people were trapped on the upper stories of a burning skyscraper. The authorities decided that the only way of saving the people would be to use a helicopter. So, the air force were called and soon a helicopter arrived on the scene.

The helicopter positioned itself near the burning building and a rope was lowered down to where the people were. One of the people then quickly grabbed hold of the rope by hooking it with a stick. Then this person tied himself onto the rope and was carried to the ground. Using this method all the people were able to evacuate the burning building before the fire reached them. However, before this method was used, another method had been tried.

At first, the helicopter had positioned itself near the burning building and the rope was let down. One of the people then grabbed hold of the rope by lassoing it with another piece of rope. This person then tied himself onto the rope and was carried to the ground by the helicopter. However, it soon became clear that this method was too slow and would not rescue all the people before the fire reached them. So, the other method was tried.

REFERENCES

Burstein, M.H. (1986). Analogical learning with multiple models. In T.M. Mitchell, J.G. Carbonell & R.S. Michalski (Eds.), *Machine learning: A guide to current research*. Lancaster: Kluwer Academic Publishers.

Clement, C.A., & Gentner, D. (1991). Systematicity as a selection constraint in analogical mapping. *Cognitive Science, 15*, 89-132.

Duncker, K. (1945). On problem solving. *Psychological Monographs, 58* (Whole No. 270).

Falkenhainer, B. (1987). An examination of the third stage in the analogy process: Verification-based analogical learning. *Proceedings of the Tenth International Joint Conference on Artificial Intelligence*. Los Altos, CA: Morgan Kaufmann.

Falkenhainer, B. (1990). Analogical interpretation in context. *Proceedings of the Twelfth Annual Conference of the Cognitive Science Society.* Hillsdale, NJ: Erlbaum.

Falkenhainer, B., Forbus, K.D., & Gentner, D. (1986). Structure-mapping engine. *Proceedings of the Annual Conference of the American Association for Artificial Intelligence* Los Altos, CA: Morgan Kaufmann.

Falkenhainer, B., Forbus, K.D., & Gentner, D. (1989). Structure-mapping engine. *Artificial Intelligence, 41*, 1-63.

Forbus, K.D. (1984). Qualitative process theory. *Artificial Intelligence, 24*, 85-168.

Forbus, K.D., Ferguson, R.W., & Gentner, D. (1994). Incremental structure mapping. *Proceedings of the Sixteenth Annual Conference of the Cognitive Science Society.* Hillsdale, NJ: Erlbaum.

Gentner, D. (1980). The structure of analogical models in science. *BBN Technical Report No. 4454*.

Gentner, D. (1983). Structure-mapping: A theoretical framework for analogy. *Cognitive Science, 7*, 155-170.

Gentner, D. (1989). Mechanisms of analogical learning. In S. Vosniadou & A. Ortony (Eds.), *Similarity and analogical reasoning*. Cambridge: CUP.

Gentner, D., & Landers, R. (1985). Analogical reminding: A good match is hard to find. *Proceedings of the International Conference on Systems, Man and Cybernetics*. Tucson, Arizona, November.

Gentner, D., Rattermann, M.J., & Forbus, K.D. (1992). The roles of similarity in transfer. *Cognitive Psychology*, *25*, 431-467.

Gentner, D., & Toupin, C. (1986). Systematicity and surface similarity in the development of analogy. *Cognitive Science*, *10*, 227-300.

Gick, M.L., & Holyoak, K.J. (1980). Analogical problem solving. *Cognitive Psychology*, *12*, 306-355.

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

Holland, J.H., Holyoak, K.J., Nisbett, R.E., & Thagard, P. (1986). *Induction: Processes in Inference, Learning and Discovery*. Cambridge, MASS: MIT Press.

Holyoak, K.J. (1985). The pragmatics of analogical transfer. *The Psychology of Learning and Motivation*, *19*, 59-87.

Holyoak, K.J., & Koh, K. (1987). Surface and structural similarity in analogical transfer. *Memory & Cognition*, 15, 332-340.

Holyoak, K.J., & Thagard, P. (1989). Analogical mapping by constraint satisfaction. *Cognitive Science, 13,* 295-355.

Holyoak, K.J., & Thagard, P. (1995). *Mental leaps*. Cambridge, MASS: MIT Press.

Keane, M. (1985). On drawing analogies when solving problems: A theory and test of solution generation in an analogical problem solving task. *British Journal of Psychology*, *76*, 449-458.

Keane, M.T. (1987). On retrieving analogues when solving problems. *Quarterly Journal of Experimental Psychology, 39A*, 29-41.

Keane, M.T. (1988). *Analogical problem solving*. Chichester: Ellis Horwood (New York: Simon & Schuster).

Keane, M.T. (1989). Modelling "insight" in practical construction problems. *Irish Journal of Psychology*, *10*, 201-215.

Keane, M.T. (1990). Incremental analogising: Theory and model. In K.Gilhooly, et al. (Eds.). *Lines of thinking*. Chichester: John Wiley.

Keane, M.T. (1994). Adaptation as a selection constraint on analogical mapping. *Proceedings of the Sixteenth Annual Conference of the Cognitive Science Society*. Hillsdale, NJ: Erlbaum.

Keane, M.T. (1995). On order effects in analogical mapping: Predicting human error using IAM. *Proceedings of the Seventeenth Annual Conference of the Cognitive Science Society*. Hillsdale, NJ: Erlbaum.

Keane, M.T., & Brayshaw, M. (1988). The Incremental Analogical Machine: A computational model of analogy. In D. Sleeman (Ed.), *European working wession on learning*. London: Pitman.

Keane, M.T., Ledgeway, T, & Duff, S. (1994). Constraints on analogical mapping: A comparison of three models. *Cognitive Science*, *18*, 287 - 334.

Kolodner, J. (1993). *Case-based reasoning*. San Mateo, CA: Morgan Kaufmann.

Maier, N.R.F. (1931). Reasoning in humans II: The solution of a problem and its appearance in consciousness. *Journal of Comparative Psychology*, *12*, 181-194.

Novick, L.R. (1988). Analogical transfer, problem similarity, and expertise. *Journal of Experimental Psychology: Learning, Memory & Cognition, 14*, 510-520.

Novick, L.R., & Holyoak, K.J. (1991). Mathematical problem solving by analogy. *Journal of Experimental Psychology: Learning, Memory and Cognition, 17,* 398-415.

Riesbeck, C. K., & Schank, R. C. (1989). *Inside case-based reasoning*. Hillsdal, NJ: Lawrence Erlbaum Associates.

Smyth, B., & Keane, M.T. (1994). Retrieving adaptable cases. In Wess, S., Althoff, K-D., & Richter, M.M. (Eds.), *Topics in case-based reasoning*. Amsterdam: Springer-Verlag.

Smyth, B., & Keane, M.T. (1995). Remembering to forget: A competence-preserving deletion policy for CBR systems. *Proceedings of the Fourteenth International Joint Conference on Artificial Intelligence*. Los Altos: Morgan Kaufmann.

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.

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.

FOOTNOTES

1 There is probably a distinction to be made here between *easily adapted* and *correct* inferences. Gentner (1989) has argued that incorrect analogical inferences (inferences that are demonstrably false in the target domain) can lead to an analogy being rejected. Adaptability acts on inferences that appear to be correct or allowable in the target domain.

2 Originally, the amount of detail in the story plan was manipulated in this experiment. However, this manipulation produced no reliable differences. This manipulation is not reported here, but treated as a variable controlled for by counterbalancing of the materials presented to subjects in both conditions. Details of the results for this manipulation are available from the author.

3 I wanted a pure measure of the solution-plan used, so the caveat in parentheses was added to make sure that subjects did not use a solution by analogy and then reject it because they felt it was a poor solution.

4 The *neither* solution category probably gives us a good indication of the number of spontaneous solutions produced. In both of the experimental conditions this stands at 13%. If we assume that 13% of swing solutions were spontaneously produced then clearly the observed rates in the experimental conditions are considerably higher.

5 It should also be noted that ACME is really a matcher and cannot form the candidate inferences required by this task; that is, it cannot form extended sets of candidate inferences without some additional mechanism or the use of implausible, dummy predicates for each inference in the target. In short, it cannot really do the candidate inferencing required to solve these problems.

Table 1 Three Possible Solutions to the General Story that Differ in Their Adaptability to the Radiation Problem

Solution	*Adaptability*
(1) The General directed low-intensity death-lasers at the fortress, so that they converged on the fortress to destroyed it.	easy
(2) The General ordered small groups of troops to march down each of the roads to the fortress, so that they converged on the fortress to destroy it.	moderate
(3) The General ordered the small groups of troops to inspect their rifles and then to march down each of the roads to the fortress, so that they converged on the fortress to destroy it.	hard

Table 2 The Percentage of Subjects Producing Different Solutions in Experiment 1

	Solution Type			
Condition	*Swing*	*Stick*	*Neither*	*N*
Stick-pragmatically-important	42% (16)	45% (17)	13% (5)	38
Swing-pragmatically-important	72% (28)	15% (6)	13% (5)	39
Control*	43% (9)	0% (0)	57% (12)	21

* Independently run.

Table 3 The Percentage of Subjects in Experiment 1 Producing Solutions Classified According to Their Pragmatic-Importance and Adaptability

Pragmatic-Importance	Adaptability: Low	Adaptability: High	
Low	9% (6)	25%(17)	34% (23)
High	23% (16)	42% (28)	66% (44)
	33%(22)	67% (45)	67

Table 4 The Percentage of Subjects Producing Different Solution Types in Experiment 2

	Solution Type				
Condition	*Swing*	*Stick*	*Lasso*	*Neither*	*N*
Stick-low-adaptable	58% (17)	28% (8)	0% (0)	14% (4)	29
Stick-high-adaptable	13% (4)*	67% (20)	10% (3)	10% (3)	30
Control**	43% (9)	0% (0)	0% (0)	57% (12)	21

* This the baseline for the spontaneous-production of swing solutions because no swing plan was present in the story analogue given to these subjects (the lasso-stick story).
** For comparison purposes, run as part of Experiment 1.

Table 5 The Percentage of Subjects in Experiment 2 Producing Solutions Classified According to Their Pragmatic-Importance and Adaptability

Pragmatic-Importance	Adaptability: Low	Adaptability: High	
Low	6% (3)	36%(17)	42% (20)
High	17% (8)	41% (20)	58% (28)
	23%(11)	77% (37)	48

FIGURE CAPTIONS

Figure 1 Subjects' Mean Ratings of the Goodness of the Solutions to the Two-String Problem in the Conditions of Experiment 2