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

COGNITIVE SCIENCE 15, 89-132 (1991)

Systematicity as a Selection Constraint in Analogical Mapping

CATHERINE A. CLEMENT AND DEDRE GENTNER

University of Illinois, Champaign, IL

Analogy is often viewed as a partial similarity match between domains. But not all partial similarities qualify as analogy: There must be some selection of which commonalities count. Three experiments tested a particular selection constraint in analogical mapping, namely, systematicity. That is, we tested whether a given predicate is more likely to figure in the interpretation of and prediction from an analogy if the predicate participates in a common system of relations. In Experiment 1, subjects judged two matches to be included in an analogy: an isolated match, and a match embedded in a larger matching system. Subjects preferred the embedded match. In Experiments 2 and 3, subjects made analogical predictions about a target domain. Subjects predicted information that followed from a causal system that matched the base domain, rather than information that was equally plausible, but that created an isolated match with the base. Results support Gentner's (1983, 1989) structure-mapping theory in that analogical mapping concerns systems and not individual predicates, and that attention to shared systematic structure constrains the selection of information to include in an analogy.

In an analogy, a familiar domain is used to understand a novel domain in order to highlight important similarities between the domains, or to predict new features of the novel domain. For example, we use our knowledge about water flow to elucidate properties of electric circuitry. Such an analogy can lead to useful inferences, and reveal deep structural features about a domain. In this research we ask how an analogical mapping is constructed. In particular, we ask whether systematic relational structure acts as a psychological selection constraint in interpreting an analogy.

An analogy can be seen as a partial similarity match between situations. But not all partial similarity matches qualify as analogy: There must be some selection of which commonalities count. The selection problem exists

This research was supported by a Child Health and Human Development Grant No. HD07205-08 and by the Office of Naval Research under Contract No. N00014-85-K-0559.

We thank Lise Dobrin, Robert Knippen, Alan Jeung, Maureen Clifford, Walter Kintsch, Laura Kotovsky, Ronald Mawby, Robert Mitchell, Betsy Perry, Mary Jo Rattermann, Brian Ross, Robert Schumacher, and an anonymous reviewer.

Correspondence and requests for reprints should be sent to Catherine A. Clement, Psychology Department, Eastern Kentucky University, Richmond, KY 40475-3108.

on at least two levels. First, as most researchers now agree, in an analogy common relations are more important than common object descriptions (Collins & Burstein, 1989; Gentner, 1983). For example, the analogy between a plumbing system and an electric circuit should not tell us that electric wires are hollow and large in diameter like water pipes. The second selection problem is more subtle: Given that two analogs may have a large number of common relations, how does a person decide *which* set of common relations should belong to the interpretation? In the *plumbing/electricity* example, we could preserve various different relations: *both are distributed to persons in a city, degree of pressure determines flow rate,* or even *both can be used in cooking*. Indeed, the number of common relations is potentially unmanageable (Rips, 1989; see also Goodman, 1970). Thus, to account for people's observed fluency in interpreting analogy, we must postulate constraints on the space of possible "commonalities that matter" between the base and target. The problem is even more acute when it comes to prediction. For example, suppose we know three facts: *It rains in San Francisco; There is a mime troop in San Francisco;* and *It rains in Urbana.* Clearly we do not want our theory of analogy to tell us that *There is a mime troop in Urbana.* We must specify constraints that tell us, given a set of matching facts, which other facts in the base can be mapped to the target as candidate inferences. We first review the kinds of selection constraints that have been proposed, and then present research testing one particular constraint, *systematicity.*

Selection Constraints

Three broad classes of selection constraints on mapping have been proposed in computational models of analogy and metaphor (see Collins & Burstein, 1989; Hall, 1989): constraints stemming from the current *goal state,* constraints stemming from the *relative importance* of the information in the analogous domains, and constraints stemming from a tacit preference for common *systems of relations.*[1]

Goal Relevance as a Selection Constraint. The selection of information to map between domains may be constrained by the goal for which the analogy is used. For example, Holyoak (1985) discusses analogy as similarity defined with respect to a goal, and on this account predicates to map are selected on the basis of their relevance to the reasoner's current goal (e.g., the reasoner's goal to solve a problem in the target domain, or to make a point about the target.[2] Carbonell (1981) offers a related approach to meta-

[1] We treat analogy and metaphor together as nonliteral similarity matches.

[2] Since his article in 1985, Holyoak has modified his account of analogy. Holyoak and Thagard (1989) propose an account of analogy that combines structural and goal-based constraints on mapping.

phor. He proposes an invariance hierarchy in which the selection of information to map between domains follows a prespecified order through 10 possible classes of information. In deciding what to map, people first seek common goals, then plans, then causal information and so on, until as a last resort, if none of the other types of commonalities are found, they seek common object descriptions. These goal-based models appeal to our intuitions that analogies are often used to map contextually important information, such as causal relations relevant to a goal. However, such models are limited in scope. They can only be applied when an analogy is used in the context of a preexisting goal that provides the analogizer with knowledge of what inferences and matches to seek. In the absence of such a prior goal, these models either make no predictions or predict failure to achieve analogical mapping.

Relative Importance as a Selection Constraint. A second class of selection constraints focuses on the relative importance of information in the analogous domains. Ortony's (1979) salience imbalance view of metaphor exemplifies this approach. On this account, the interpretation of metaphor maps high-salient features of the base domain onto low-salient features of the target domain. Ortony, Vondruska, Foss, and Jones (1985) provided evidence that subjects tend to prefer metaphors that obey salience imbalance. However, other research suggests that salience imbalance by itself does not specify a metaphor interpretation (Gentner & Clement, 1988). That is, the interpretation of a metaphor cannot be predicted from the relative salience of features in subjects' prior representations of its constituent terms. Thus, salience imbalance may be part of the story, but it does not appear to provide an adequate selection rule for metaphor.

An insightful aspect of Ortony's salience imbalance model is its modification of Tversky's (1977) contrast model of similarity to include the notion that the salience of a feature can be context dependent. This article will pursue a different sort of departure from the contrast model, namely, the possibility that interconnections among features explicitly enter into the processing of similarity and analogy.

Relational Structure as a Selection Constraint. Finally, selection constraints may arise from structural relationships among the matching features. In Gentner's (1980, 1983, 1989) structure-mapping account of analogy, the predicates included in an analogy are those that are embedded in a larger shared structure (e.g., a structure linked by causal relations). That is, the value of a given matching component depends not only on the match itself, but also on other matches to which it is connected (Forbus & Gentner, 1989). The structure-mapping account stems in part from the observation that useful analogies, such as those used in science or education, involve rich, interconstraining systems of mappings between two domains, rather than a set of independent correspondences (e.g., Black, 1962; Gentner, 1980, 1983; Hesse, 1966; Holyoak, 1984; Kittay, 1987). Analogies that underlie our understanding of many everyday concepts also appear to have this coherent structure (Carbonell, 1983; Lakoff & Johnson, 1980). The proposal, then, is that analogy is essentially a cognitive device for mapping systematic relational structures from a base to a target domain, and that people's attention to connected relational structure guides the analogical mapping process.

Specifically, on the structure-mapping account, the interpretation of an analogy is guided by two kinds of implicit constraints: (1) two structural consistency constraints—specifically, a preference for consistent, one-to-one object correspondences, and a preference for consistently mapped dependency structure—and (2) two selection constraints. The first selection constraint is that only *relational* commonalities count; similarities in nonrelational object attributes are irrelevant unless they are governed by a common relation. This follows from the position that analogy is a mapping of relations between objects rather than of independent object descriptions. This first selection principle is fairly widely accepted. For example, Anderson (1989) has suggested a similar idea in his "no function in identity" principle. However, as discussed above, this constraint is not sufficient, for there is, potentially, an indefinite number of common relations between objects. The second selection constraint is the *systematicity principle*, which holds that among the set of common relations, the relations that are mapped are just those that participate in common systems of relations: sets of common relations connected by higher-order relations that can themselves be mapped.[3] Thus, common isolated lower-order relations—relations not linked to a larger matching system—are typically disregarded in analogy interpretation.

Psychological Evidence for the Importance of Systematicity in Analogy. Although we know of no prior research directly testing systematicity as a selection constraint, several lines of research provide evidence for the general role of systematic relational structure in analogical processing. First, Gentner and Landers (1985) and Rattermann and Gentner (1987) showed that judgments of the soundness of a comparison between two situations (that is, whether the analogy would yield justifiable inferences) are positively related to the systematicity of the mapping. Second, Gentner and Schumacher (1986) and Schumacher and Gentner (1988) found that systematicity can facilitate accurate analogical mapping. In their research, subjects were taught a device model and then asked to transfer their knowledge to an analogous device. Subjects were able to achieve accurate transfer

[3] First-order relations are relations between objects. Higher-order relations (e.g., causal relations) are relations among relations.

in substantially fewer trials when they were taught a domain model with systematic structure (even though the device procedures were held constant). Third, in a study of analogical problem solving, Holyoak and Koh (1987) varied the degree of structural correspondence between analogous problems. Consistent with the systematicity principle, they found that when the causal system describing the initial problem state differed between the base and target problems, subjects were less likely to transfer the solution from the base to target, even though the solution would have been adequate in the target problem. Finally, Gentner and Toupin (1986) found that 9-year-old children could more accurately map a story plot between two sets of characters when the story included explicit systematic structure. In sum, prior evidence indicates that systematicity enters into the evaluation of analogies, and that it aids in accurate transfer. However, this research has not tested the claim that systematicity acts as a selection filter during analogy interpretation.

The Question of Selection Constraints in Empirical Research

In general, psychological research has bypassed the selection problem in describing analogical processing, with a few exceptions (Gentner & Clement, 1988; Ortony et al., 1985). Research on the role of complex analogy in learning and problem solving has typically focused either on demonstrating the conceptual power of analogies when learning a new domain. (e.g., Burstein, 1983; Gentner & Gentner, 1983; Reed, 1987; Rumelhart & Norman, 1981), or on discovering factors that determine access of analogs in memory (e.g., Catrambone & Holyoak, 1989; Gentner & Landers, 1985; Gick & Holyoak, 1980, 1983; Novick, 1988). There has been relatively little concern for explaining how subjects construct an analogical mapping, or for isolating the basis of their mapping choices.

In some lines of investigation, the selection issue has not been addressed because the research has utilized very simple analogies. For example, much research has examined proportional analogies, or four-term analogy problems: for example, APPLE:EAT::MILK:______, in which subjects must find the missing term(s) from a set of alternatives (e.g., WHITE, DRINK, COW, SWEET). Sternberg (1977) suggested a componential analysis identifying the sequence of steps used in solving such problems. However, this account does not include the selection of which relation to map, since the selection issue does not arise in these kinds of analogies; they are carefully designed so that only one clear relation is meant to be mapped. Thus, accounts of the processing required to solve these four-term analogy tasks have little to say about how people interpret complex, explanatory analogies that contain many possible mappings, and that may be a continuing source of hypotheses about the target domain. In a similar vein, Anderson (1989), whose PUPS system models the use of analogy in learning programming, suggests that constraints on the selection of relations to be mapped are not needed in his system. One could argue this point, but the general principle, again, is that the more limited the space of possible mappings (and programming may be a domain in which the possibilities are relatively few), the less pressing is the constraints problem. However, as discussed above, in richer domains, such as many that occur in physical science and natural discourse, the number of possible commonalities is large enough that models that ignore the selection problem are untenable.

Experiments to Address Systematicity as a Selection Constraint

As previously discussed, despite the theoretical plausibility of the existence of selection constraints in analogy, very little psychological research has attempted to identify these constraints. In this research we tested whether systematicity provides a constraint on the selection of information to include in an analogy. We describe three experiments that looked separately at two distinct components of analogical mapping: (a) *matching* existing information in the base and target, and (b) *inferring* new information about the target that follows from the analogy with the base domain.

Our first experiment examined whether systematicity constrains the matching process. We constructed a set of novel analogies that were specifically designed to test whether mapping choices are guided by systematicity. In these analogies, subjects could choose which of *two* facts (lower-order relations) to map from a base to a target domain. In all cases, both facts were equally acceptable when considered as independent matches. But the two facts differed in whether they were part of a shared systematic relational structure (specifically, a shared causal structure). Under the systematicity hypothesis, subjects should prefer those matches that are embedded in a larger system of matches, and they should reject those matches that are relatively unconnected. Thus, mapping decisions should not be based exclusively on a matching fact itself. Decisions should be determined by a match's interdependence with other information shared by the two domains. With our materials, we could distinguish the effect of the lower-order facts themselves on mapping choices from the effect of the higher-order embedding of these facts.

Our second and third experiments examined whether systematicity constrains inferences carried over from the base domain to the target. Subjects were asked to make an analogical prediction about a target domain. Two candidate facts were present in the base domain that were equally plausible as inferences about the target. However, only one fact was linked to a causal system shared by the base and target. If a systematicity principle governs analogical inference, subjects should identify matching systems in the two domains, and map those relations that are present in the base system but missing in the target system. Subjects should not select just any base fact on the selection of relations to be mapped are not needed in his system. One could argue this point, but the general principle, again, is that the more limited the space of possible mappings (and programming may be a domain in which the possibilities are relatively few), the less pressing is the constraints problem. However, as discussed above, in richer domains, such as many that occur in physical science and natural discourse, the number of possible commonalities is large enough that models that ignore the selection problem are untenable.

Experiments to Address Systematicity as a Selection Constraint

As previously discussed, despite the theoretical plausibility of the existence of selection constraints in analogy, very little psychological research has attempted to identify these constraints. In this research we tested whether systematicity provides a constraint on the selection of information to include in an analogy. We describe three experiments that looked separately at two distinct components of analogical mapping: (a) *matching* existing information in the base and target, and (b) *inferring* new information about the target that follows from the analogy with the base domain.

Our first experiment examined whether systematicity constrains the matching process. We constructed a set of novel analogies that were specifically designed to test whether mapping choices are guided by systematicity. In these analogies, subjects could choose which of *two* facts (lower-order relations) to map from a base to a target domain. In all cases, both facts were equally acceptable when considered as independent matches. But the two facts differed in whether they were part of a shared systematic relational structure (specifically, a shared causal structure). Under the systematicity hypothesis, subjects should prefer those matches that are embedded in a larger system of matches, and they should reject those matches that are relatively unconnected. Thus, mapping decisions should not be based exclusively on a matching fact itself. Decisions should be determined by a match's interdependence with other information shared by the two domains. With our materials, we could distinguish the effect of the lower-order facts themselves on mapping choices from the effect of the higher-order embedding of these facts.

Our second and third experiments examined whether systematicity constrains inferences carried over from the base domain to the target. Subjects were asked to make an analogical prediction about a target domain. Two candidate facts were present in the base domain that were equally plausible as inferences about the target. However, only one fact was linked to a causal system shared by the base and target. If a systematicity principle governs analogical inference, subjects should identify matching systems in the two domains, and map those relations that are present in the base system but missing in the target system. Subjects should not select just any base fact that could be plausibly inferred in the target, but should infer a fact that follows from a shared interdependent set of relations.

Before describing the structure of our materials in more detail, two further criteria in developing our tasks should be mentioned. First, to preserve a realistic degree of complexity, the situations described were fairly rich in information. Second, in the interests of generality, we did not wish to limit ourselves to cases in which goal relevance was a possible selection constraint on mapping. Thus, subjects did not have to map information in order to solve problems or prove points in the target domain. Finally, it was important that subjects' answers be governed by use of the analogy rather than simply by prior beliefs about the target. Therefore, we developed analogies between novel, fictional domains rather than using real-world analogies.

Design of the Materials. Each of the analogies developed consisted of a base and target passage describing novel objects or organisms on fictional planets. Each passage included two chief paragraphs. One paragraph described a causal structure that matched between the base and target, and the other described a causal structure that did not match. Subjects had to make mapping choices between key *facts* that were embedded either in the matching or in the nonmatching causal structures. The key facts themselves always matched between the base and the target.

To help understand the materials, we first describe in abbreviated prose the base passage for one analogy. This passage described creatures called *Tams* who live on a distant planet. (The actual passages were about one page long and were written in the style of an encyclopedia article.)

> Paragraph 1: The Tams live on rock and can grind and consume minerals from the rock through the constant action of their underbelly. However, periodically they run out of minerals in one spot on the rock and must relocate. At this time they stop using their underbellies.
>
> Paragraph 2: Although at birth the Tams have rather inefficient underbellies, eventually the underbellies adapt and develop a texture that is specially suited to the rock the Tam lives on. As a consequence, a grown Tam's underbelly cannot function on a new rock.

Each paragraph describes a causal structure consisting of a *key fact* and a causal antecedent governing that key fact. The two key facts in this base domain are

1. *The Tams sometimes stop using their underbellies.*
2. *The Tams' underbellies cannot function on new rocks.*

This base domain, and the analogous target domain called "The Robots," which describes robots who use probes to gather data from planets, are outlined in Table 1. The left column of the table shows the two causal structures

TABLE 1
Relational Structure of the Base Domain, *The Toms*, and the Target Domain, *The Robots*

Base: *The Toms*	Target: *The Robots*	
	Version 1	Version 2
Consume minerals with underbellies	Gather data with probes	Gather data with probes
Exhaust minerals in one spot and must relocate on the rock	**Exhaust data in one place and must relocate on the planet**	Internal computers over-heat when gather a lot of data
So stops using underbelly	*So stops using probes*	*So stops using probes*
Born with inefficient underbelly	Designed with delicate probes	Designed with inefficient probes
Underbelly adapts and becomes specialized for one rock	Robots cannot pack probes to survive flight to a new planet	**Probes adapt and become specialized for one planet**
So underbelly can't function on new rock	*So probes can't function on new planet*	*So probes can't function on new planet*

Note. Key facts are shown in italics. Matching causal information is shown in boldface. In Experiment 2, italicized facts were removed from the target.

of the base, with each key fact shown in italics, and the causal antecedents shown in boldface. The middle column of the table shows Version 1 of The Robots which, like the base domain, contains two causal structures. Importantly, the key facts in *each* structure of the target *match* the key facts in the base:

1. *The robots sometimes stop using their probes.*
2. *The probes cannot function on new planets.*

However, although both key facts in the target match the base domain, only the first key fact is linked to a causal system that also matches the base domain. We will call this the *shared-system* key fact. We predicted that subjects should prefer this shared-system fact in mapping. Although the other key fact matches the base domain, it is linked to a causal antecedent that does not match the base. (It should be noted that for ease of reporting, the matching causal structures in Table 1 are described in language that is similar at the surface level. In the actual passages the matching key paragraphs were written in more domain-specific language, and we tried to avoid extensive similarity in sentence structure within key paragraphs.)

To avoid confoundings with particular content there were two versions of a target domain, as shown in Table 1. In Version 1 of the target, the cause for the first key fact matches the base domain, but the cause for the second key fact does not. Version 2 of the target contains the *same* key facts but reverses which key fact is linked to a shared causal system: In this version, the second key fact is the shared-system fact. This ensures that mapping preferences for shared-system key facts cannot be attributed to preferences for a particular key fact in itself. A further control task was included to ensure that subjects' responses were not due to a preference for a particular fact in the context of a particular version of the target. In this task control subjects read only the target stories and judged which key fact was more important for each story. If these "target-only" subjects show no bias toward one key fact or the other, then any preference among experimental subjects can be attributed to the specific effects of the match with the base system.

Experiment 1 examined the matches people include in an analogy. Subjects had to judge how well each key fact in the target contributed to the analogy with the base domain. We predicted that shared-system key facts would be preferred over different-system key facts, even though both key facts in themselves match the base equally well. In Experiments 2 and 3, the materials were altered slightly, and we examined subjects' inferences about the target domain, given its analogy with the base domain. If subjects prefer the shared-system key facts as analogical matches and inferences, this would indicate that subjects employ a systematicity principle in selecting information to map between analogous domains.

EXPERIMENT 1

Method

Overview

Four novel analogies were created, each consisting of a base and target passage designed according to the structure described above.[4] Subjects in the analogy group read both the base and target passage for each analogy, and subjects in a control group (target-only group) read only the target passage. For each analogy, after learning the passage(s), the two groups were given parallel tasks. These tasks required subjects to rate or choose between the two key facts in the target passage. As discussed above, each key fact matched the base passage, and each was embedded in a causal system. The manipulation was that this causal system matched between the base and target for the *shared-system fact* but did not match for the *different-system fact*. The analogy subjects evaluated how well the two key facts contributed to the analogy. The target-only subjects evaluated the importance of the facts to the target passage. We predicted that, in their ratings and choices, analogy subjects would prefer the shared-system fact, and target-only subjects would show no preference for either fact.

The experimental variables were fact-type (shared-system vs. different-system), a within-subjects variable, and condition (analogy vs. target-only), a between-groups variable. Additional variables were passage (the four target

[4] Materials used in all the experiments may be requested from the first author.

passages), a within-subjects variable, and the counterbalancing factor version-set (two versions), a between-groups variable (see Table 1). Version-set refers to which version of the target passage was given. Within each condition there were two subgroups of subjects, each received a different version of each of the four target passages.

Subjects

The subjects were 48 paid undergraduate students at the University of Illinois. Half were assigned to the analogy condition and half to the target-only control condition.

Procedure

Subjects were run in groups of three to six. The experimenter read aloud the instructions for each task. (The instructions were also given in writing.) Subjects had as much time as they needed. Sessions for target-only subjects lasted approximately 1 hour. Sessions for analogy subjects, who had more material to read, lasted approximately 2 hours.

Task for Analogy Subjects. Our interest was in subjects' judgments of how well each key fact in the target contributed to the analogy with the base domain. We used a rating task and a choice task to assess these judgments. Since the materials are complex, we first gave subjects learning tasks to be sure that they understood the materials.

Learning Tasks. The subjects' first task was to read the base passage carefully. Then, they were told that the next passage they were to read—the target passage—was analogous to the base passage. For the first analogy in the session, before reading the target, subjects were given an example of what we meant by "analogous." They were told,

> We can say that plant stems and drinking straws are analogous. Even though they are two different things, they are both used as channels to bring liquid nutrients from below up to a living thing....

After reading the target, subjects were asked to match the objects that corresponded between the base and target domains. They were given a list of the approximately four or five central objects in the base passage. For example, in the analogy described in Table 1, the list included *Tams, rocks, underbellies, minerals, internal organs*. Subjects were to identify which objects in the target (if any) corresponded to each base object. Thus, the answers here were *Robots, planets, probes, data, and microcomputers*. The first correspondence, that between the subjects of the stories (e.g., the *Tams* and the *Robots*) was always filled in to illustrate the task. In order to avoid biasing subjects, the stories were designed so that there was no inconsistency in the object correspondences relevant to the shared- and different-system facts.

Subjects then wrote out the ways in which the two domains were and were not analogous. The instructions were, for example:

> Describe what is and is not analogous about the Tams and the Robots. When two things are analogous some things fit better than other things. Thus, some facts about the Robots contribute to making the analogy with the Tams a good analogy and some facts do not contribute to the analogy. Describe those facts that support the analogy and describe the facts that do not support the analogy.

Subjects were encouraged to refer back to the passages if necessary. Subjects were given no feedback. The purpose of the preliminary learning tasks was simply to lead subjects to process the analogies thoroughly.

Rating Task. After the learning tasks, subjects were given two experimental tasks in which they evaluated how well the two key facts from the target contributed to the analogy. First, in a rating task, the two key facts were presented on a separate page and subjects rated the degree to which each fact supported the analogy on a scale of 1 to 7. For example, the italicized facts from the target in Table 1 were presented, stated as in the passage:

1. *Sometimes a robot must shut down its probes,* and
2. *Robots cannot go from one planet to another to gather data.*

The rating instructions, which were given in writing and read aloud, were, for the Robots/Tams example:

> When two things are analogous some things fit better than other things. Some facts support the analogy better than other facts. Assume you are in a debating match and you have to defend the claim that the Robots are analogous to the Tams. Below are two facts about the Robots. Please rate them on a scale of one to seven according to how well they support the claim.

Subjects were told to give a 7 to a statement that contributed well to making the analogy a good analogy, and a 1 to a statement that did not contribute at all to the analogy.

Choice Task. Next, subjects were given the two key facts again on a new page. Subjects were asked to choose between them and to explain their choice:

> Still assuming that you are in a debating match, reread the two facts, they are printed again below. Which statement best supports the claim that the Robots and Tams are analogous?

Subjects were required to give a brief written explanation for their choice of one fact and their rejection of the other. During both judgment tasks, subjects were encouraged to reexamine the passages as needed. Thus, there were no memory requirements.

TABLE 2
Experiment 1. Results of Rating Task: Mean Ratings[a] of Shared-System (Correct)[b] and Different-System Key Facts for Each Passage

	[Analogy Group (n=24)]				[Target-Only Group (n=24)]			
	Shared System		Different System		Shared System		Different System	
Passage	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
1	6.12	.95	4.96	1.88	4.79	1.72	4.96	1.68
2	6.04	1.46	5.37	1.79	5.42	1.50	5.96	.99
3	6.42	1.10	4.33	1.86	5.62	1.28	5.42	1.06
4	6.54	.93	4.87	1.62	5.04	1.68	4.92	1.64
Overall	6.28	.96	4.88	1.32	5.22	.80	5.31	.57

[a] Facts were rated on a 1-7 scale.
[b] In this and the following tables we will use the term "correct" to remind the reader that it is the shared-system key fact that would be preferred by analogy subjects if they were reasoning according to the systematicity hypothesis.

Tasks for Target-Only Subjects. The target-only subjects read only the target passage. To ensure that these subjects thoroughly processed the passage, they were asked to summarize it. Subjects were told that their summaries should include discussion of the key objects from the target passage (these objects were listed for them). Following this learning task, target-only subjects were given the same two key facts given to analogy subjects and asked to perform judgment tasks in the same order as the analogy subjects. First they *rated,* on a scale of 1 to 7, how important each fact was to the target passage. Following this, they were presented with the facts again and were asked to *choose* which fact was most important to the target passage. Subjects were asked to explain their choices. (We were not interested in these justifications, but included this question in order to maintain similarity to the task given to analogy subjects.)

Results

Rating Task

Table 2 shows the mean ratings for shared-system and different-system key facts for each target passage. As predicted, when asked to rate how well each fact contributed to the analogy, analogy subjects rated shared-system key facts higher than different-system key facts (overall $M = 6.28$ and $M = 4.89$, respectively). Thus systematicity governed the analogy subjects' preference. In contrast, the target-only subjects gave equivalent importance ratings to the two types of key facts, ($M = 5.22$ and $M = 5.31$), indicating that the analogy group ratings are not due to differences in the *importance* of the key facts in the two versions of the target. To compare the ratings of each type of key fact in each condition a $2 \times 2 \times 4 \times 2$ repeated-measures ANOVA was conducted. The experimental variables were fact-type (shared-system vs. different-system) and condition (analogy group vs. target-only group).

Additional variables were passage (the four target passages) and the counterbalancing variable version-set (1 vs. 2). Both subjects and passage were treated as random variables in the analysis (see Clark, 1973).

There was a significant main effect for passage, indicating that overall, ratings varied across passages, $F(3, 132) = 3.09, p < .05$. There were no other significant main effects. Significant two-way interactions were found between condition and passage, $F(3, 132) = 3.57, p < .05$, and fact-type and passage, $F(3, 132) = 3.00, p < .05$. These interactions simply indicate that averaged over fact-type, condition had an effect that varied across passage, and, averaged over condition, fact-type had an effect that varied across passage.

The key prediction is an interaction between condition and fact-type. This interaction was significant, $F'(1, 44) = 13.99, p < .001$, indicating that the difference in ratings of shared-system and different-system key facts was different for the analogy and target-only groups. When this interaction was analyzed, the simple effect of fact-type within the analogy group confirmed that the mean rating for shared-system facts by analogy subjets is significantly higher than the mean rating for different-system facts, $F'(1, 11) = 10.44, p < .01$. This difference does not hold for the control group. No other two-way interactions were significant.

There is no evidence that the key interaction between condition and fact-type was specific to particular passages or to a particular version-set. That is, there was no triple interaction among condition and fact-type and either of these other factors. Finally, a significant triple interaction was found among version-set, fact-type and analogy, $F(3, 132) = 19.77, p < .001$. This simply indicates that, overall, a difference in ratings to the two fact types was found that depended on the version of particular analogies. No other effects were significant. In sum, the rating task responses support the prediction that systematicity would govern analogy subjects' preferences for particular matches between analogous domains.

Choice Task

The results for the choice task are consistent with those of the rating task. As predicted, analogy subjects most often chose the shared-system fact as best contributing to the analogy. In contrast, target-only subjects showed no preference for this fact over the different-system fact in their importance judgments. To test the difference between groups, subjects were assigned a score for the number of choices of shared-system key facts across the four analogies (giving a possible score of 0–4). A 2 × 2 ANOVA (Condition × Version-Set) revealed a significant main effect of condition, indicating that the mean for analogy subjects (3.17) is significantly greater than the mean for target-only subjects (2.0), $F(1, 44) = 18.43, p < .001$. There was also a main effect of version-set, $F(1, 44) = 6.02, p < .05$, indicating that, overall, the number of choices of shared-system facts varied across version-set. As expected, the interaction between condition and version-set was not significant.

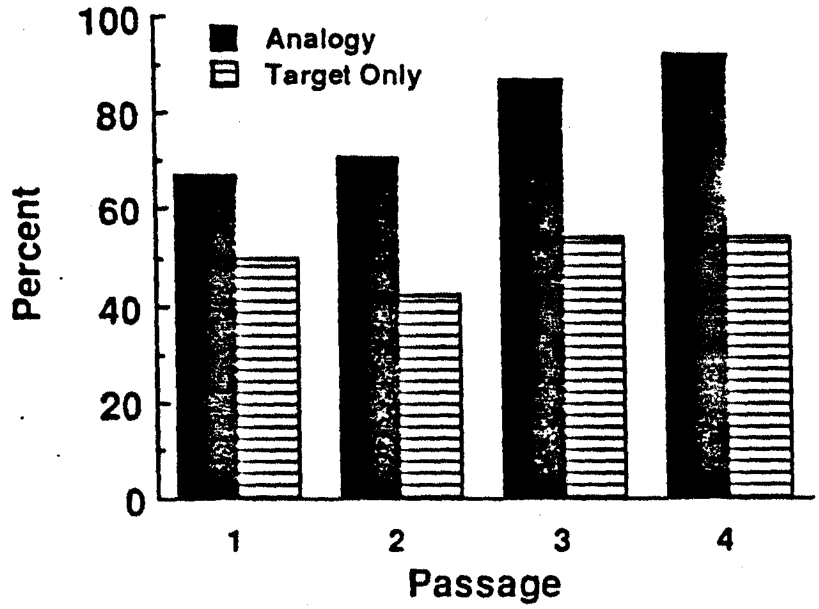

Figure 1. Experiment 1 Choice Task: Percent of subjects in each condition choosing the shared-system (correct) fact for each passage.

Figure 1 shows that the pattern held for each of the four passages. The analogy subjects chose the shared-system fact 67% to 92% of the time, while target-only subjects chose it 42% to 54% of the time. Fisher exact tests reveal that the difference between groups is significant for three of the four passages: ($p < .05$, one-tailed). In sum, these results show that the analogy subjects viewed the matching fact embedded in a shared causal system as better support for the analogy than an equally good match that was not so embedded.

The rating and choice task results support the position that analogical processing is not a matter of concatenating independent matches, but of finding a connected system of matches. The goodness of a particular match is influenced by its neighboring matches that form a mutually constraining system. We next asked whether subjects were explicitly *aware* of the higher-order constraints that appear to govern their choices.

Justifications of Choices

In the choice task, analogy subjects were asked to write a brief justification for their responses. They were to explain (a) why they thought their *chosen* fact supported the analogy best, and (b) why the *rejected* fact did not support the analogy as well. The justifications for choices and rejections were separately scored. Our main interest was in whether subjects would show an explicit concern for the shared higher-order embedding of key facts. Thus, we scored justifications as to whether subjects referred to the causal systems supporting or not supporting a key match, or referred only to the lower-order match of the key fact itself. Responses of all subjects were coded, regardless of whether they chose the shared-system or the different-system key fact. Two judges coded subjects' justifications for choosing a fact into four categories which can be ordered according to the level of focus:

1. *Focus on similarity in causal structure* (*higher-order similarity*): Subjects stated that the cause of the chosen fact was similar in the base and target, or described the similar cause for the chosen fact. For example, if subjects chose the key fact *Robots cannot go from one planet to another to gather data,* they might give a justification such as, "Robots cannot go to another planet because they become adapted to one planet, and Tams cannot go to another rock because they are adapted to one rock"; or "The Robots cannot go to another planet for the same reasons that Tams cannot go to another rock."
2. *Focus on the similarity of chosen fact* (*lower-order similarity*): Subjects simply noted the direct similarity between the base and target key facts, for example, "Robots cannot gather data on another planet and Tams cannot function on another rock."
3. *Focus on the chosen fact alone* (*importance of the target fact*): Subjects asserted that the chosen fact is the most important fact.
4. *Other:* Gives any other response.

Subjects' justifications for rejecting a fact were classified into the same four categories, except that now justifications focused on: (1) *dissimilarities* in the cause for the rejected fact to the cause for the corresponding fact in the base, (2) the *dissimilarity* of the rejected fact to the corresopnding fact in the base, or (3) the *lesser* importance of the rejected fact.

The important contrast is between the first and second categories. The first category reflects subjects' concern for the higher-order embedding of the matching fact, whereas the second category reflects only a concern for the lower-order match itself. The third and fourth types of responses could not be clearly categorized in this respect.

Table 3 shows the justification data. The left side of the table shows the distribution of responses when subjects made the predicted choice, that is, when the shared-system fact was chosen (and thus the different-system fact was rejected). The right side shows responses when the different-system fact was chosen. The top half of the table shows justifications for choosing a fact, and the bottom half shows justifications for rejecting the other fact. Interrater agreement in classifying choice justifications was .86 before discussion, and agreement in classifying rejection justifications was .82 before discussion. (Data shown throughout this article reflect final scoring decisions.)

As shown in Table 3, the results confirm that subjects who chose the shared-system fact did so because shared systematicity was important. Looking at the left side of the table, we see that the majority of these sub-

TABLE 3
Experiment 1. Justifications by Analogy Subjects for Facts Selected in Choice Task: Proportion and Frequency of Shared-System and Different-System Choices (and Rejections)

	Chosen Fact			
	Shared-System Fact (Correct)		Different-System Fact (Incorrect)	
Justification Type	Prop.	Freq.	Prop.	Freq.
Similar cause for chosen fact	.73	(56)	.35	(7)
Similarity of chosen fact	.16	(12)	.20	(4)
Importance of chosen fact	.04	(3)	.30	(6)
Other	.07	(5)	.15	(3)
Total	100	(76)	100	(20)
	Rejected Fact			
	Different-System Fact (Correct)		Shared-System Fact (Incorrect)	
Justification Type	Prop.	Freq.	Prop.	Freq.
Dissimilar cause for rejected fact	.56	(43)	.05	(1)
Dissimilarity of rejected fact	.20	(15)	.30	(6)
Importance of rejected fact	.04	(3)	.20	(4)
Other	.20	(15)	.45	(9)
Total	100	(76)	100	(20)

Note. Out of 96 opportunities, subjects chose the shared-system fact (the predicted response) 76 times and the different-system fact 20 times. For each choice, justifications are shown for both the chosen fact and the rejected fact.

jects' justifications both for their choice of shared-system fact, or their rejection of the different-system fact were concerned with shared causal information. Relatively few justifications were concerned only with the lower-order match itself. Some examples of these justifications in which subjects explicitly mention the common causal constraints are shown in Table 4. (Note that these examples are all from subjects who gave the predicted response. The examples of choice and rejection justifications are not from the same subjects.)

Subjects did not always give higher-order justifications when choosing the shared-system fact. This suggests that people may sometimes operate on the basis of shared relational structure without necessarily being able to articulate their reasoning. In this connection, it is interesting that people seemed more often to give higher-order justifications for *acceptance* of shared-system facts than for *rejection* of different-system facts, even though these represent the same set of ("correct") choice responses. This pattern is intelligible if we assume that there is some processing cost for negativity (e.g., Clark & Chase, 1972). Conceivably, it is easier in the positive case for

TABLE 4

Experiment 1. Examples of Justifications for Shared-System (Correct) Choices[a]

Justifications for choosing the shared-system fact:

Robots cannot go from one planet to another to gather data:

1. . . .there is a direct comparison. Robots cannot go and be effective from one planet to another because they have adapted to one particular planet. Toms cannot move from one type of rock to another because they also have adapted.
2. . . .the robots cannot move planets after it has developed probes and the Tom cannot move to a new rock once it has developed its underbelly.

Justifications for rejecting the different-system fact:

Robots sometimes shut down their probes.

1. . . .the reason the robot shuts down is because it has received *too much* information, whereas the reason the tom must stop sucking and move on is because it is no longer *receiving* enough minerals.
2. Although a robot must shut down its probes sometimes and a Tom's underbelly sometimes no longer functions, they are not as similar. The robot shuts down when it has too much of that which it is seeking (data) and the Tom shuts down when it does not have enough of what it is seeking (minerals).

[a] Responses are for the *Toms/Robot* analogy, given Target Version 2.

subjects to introspect and articulate the higher-order causal connections that govern their choice.

When subjects ("incorrectly") chose the different-system fact (right side of Table 3), their justifications were evenly distributed among the categories, focusing either on causal information, the similarity of the fact itself, or the importance of the chosen fact. Interestingly, when subjects mentioned causal information in justifying the different-system choice, they often invented or imputed a cause in the target that was similar to that in the base, even though this interpretation went beyond and/or was inconsistent with the information provided. For example, in Version 1 of the target domain *The Robots,* the robots are described as *not* specialized for a particular environment (see Table 1). However, some subjects read the passage as saying that the robots *were* specialized, thus creating a match in causal information with the base domain. In sum, subjects' choice-task justifications indicate that links to a shared causal system guided which lower-order matches were judged to belong best to an analogy.

Discussion

As predicted, systematicity appears to constrain which similarities are selected for an analogy. In both their ratings and choices, given two equally good lower-order matches between the base and target, analogy subjects preferred whichever matching fact was embedded in a matching causal system. Let us review this finding in more specific terms. Analogy subjects were given two lower-order relations, T1 and T2, which both matched key relations in the base domain. Subjects had to choose which match, B1—T1 or B2—T2, best belonged to the analogy. Note that each key fact was embedded in some causal system in both the base and target. Thus, the immediately dominating relation was the cause in all cases. The manipulation was in whether the causal antecedents matched between domains, for example,

Base: CAUSE (X, B1) and CAUSE (Y, B2)

Target: CAUSE (X', T1) and CAUSE (Z, T2)

Analogy subjects selected the fact with like embedding: In this case, T1, whose causal antecedent, X' matches the antecedent (X) of B1 in the base.

Subjects' response justifications further confirmed their concern for the higher-order embedding of matching facts. Finally, the fact that control subjects, who saw only the target domain, showed no preference for the shared-system facts indicates that the analogy responses are not due to the relative importance of the key facts in the target domain.

These results indicate that analogical matching is not merely a feature-by-feature decision: Analogical matching concerns systems of predicates, not individual predicates.

EXPERIMENT 2

Experiment 1 focused on *matching* known facts between two analogous domains. The results indicated that systematicity constrains which matching facts are selected as belonging to an analogy. Experiment 2 focuses on the use of analogies to predict new facts about a target domain. We asked whether systematicity would guide the *inference* process in which relations in the base domain are carried over to the target domain as candidate inferences about the target. In this experiment, instead of only asking subjects to choose among specified mapping possibilities, we included a task that allowed subjects themselves to find information to map from the base to the target. Specifically, after reading the base and target, subjects were asked to predict a fact about the target domain that followed from the analogy with the base. If systematicity guides analogical inference, then subjects' new predictions should center on facts that (if true in the target) would link into a causal structure shared by the two domains.

Method

Overview

Subjects read base and target passages that were similar to those of Experiment 1 and that again followed the design described in Table 1. However, for this experiment the two key facts were removed from a target passage; thus, the italicized facts shown in the table were present only in the base passage. Of course, although these facts were removed from the target, they were plausible in that domain. The target domain still included the antecedent information that could potentially cause each key fact. For example, in the target shown in Table 1, in Version 2, the reader learns that the robots' computers overheat, and learns that the robots become adapted for one planet. As in Experiment 1, this antecedent information does not match the base domain in one case, but does match the base domain in the other case.[5]

In this experiment subjects were asked to make a prediction (and later to judge given predictions) about the target domain. Our question was whether they would predict one of the two key facts omitted from the target, and more importantly, if so, which one. Thus, given the analogy in Table 1, we were interested in whether subjects would predict either that *the robots sometimes stop using their probes* or that *the robots fail to function on a new planet.* If subjects are simply trying to predict facts about the target that correspond to individual facts in the base, or that are plausible in the target, they should be equally likely to predict either of the omitted key facts. If, however, subjects are guided by systematicity in making predictions, they should predict the one key fact for which there was a matching antecedent in the base and target, that is, the shared-system fact. Thus, given Version 2, subjects should predict that *robots' probes cannot function on a new planet.*

As before, in order to avoid confoundings with particular content, two versions of each target were used; which key prediction would follow from a matching antecedent was counterbalanced across the two versions (see Table 1).

It was important that both the shared-system and different-system predictions be equally easy to construct in the target domain. We took several steps to ensure this. First, given all the information in the base and target, subjects could easily identify the appropriate object correspondences and many relational correspondences between the two domains. Therefore, new analogous relations in the target corresponding to either key fact in the base could be easily created. Second, as already mentioned, the target included antecedent information making both predictions plausible. Finally, to control for any differences in the ease or plausibility of constructing the two key facts in the target, the design was counterbalanced as discussed above,

[5] Aside from removing the key facts, some further modification of the materials used in Experiment 1 was sometimes necessary. The goal was to have the key facts in the target be plausible in the target and predictable given the analogy with the base domain. However, we did not want these facts to be obvious or necessarily true given the target alone. Thus, to ensure that the key facts were not obvious consequences of the antecedent information in the target passages, we sometimes rewrote or removed some of the original antecedent information. However, enough antecedent remained to make the key facts plausible, and where desired, to allow a match with antecedent information in the base passage. The specific content of some of the analogies was also modified for clarity.

and a target-only group created and judged predictions based only on reading the target passages.

In sum, in this experiment we asked analogy subjects to make a prediction about the target domain based on the analogy with the base domain. Because much of the content of the base and target passages already matched, subjects were constrained in their possible new predictions. However, they had two particularly plausible choices: There were always two facts presented in the base that were not present in the target for which parallel target facts could be straightforwardly constructed. Subjects could potentially carry over either of these key facts from the base domain. However, only one of them would follow from a causal system shared by the base and target.

Subjects

Subjects were 32 paid undergraduate students at the University of Illinois. Half were assigned to the analogy condition and half to the target-only condition.

Procedure

Task for Analogy Subjects. For each analogy, subjects first had learning tasks identical to those in Experiment 1. They then were given three experimental tasks requiring them first to make a prediction about the target domain and then to judge some possible predictions.

Prediction Task. After performing the learning tasks, subjects were asked to make one prediction about the target domain. They were told to predict new information about the target that was suggested by the analogy with the base passage. The specific instructions (which were given in writing and read aloud) were, for example:

> Because the Robots are analogous to the Tams, we might add some information to the story about the Robots. Aside from the information already stated about the Robots, the Tam story can suggest some predictions. Look again at the two stories; then in the space below predict one thing that might be true of the Robots that is suggested by analogy with the Tams.

It was further emphasized that this should be a prediction about what *might be* true about the Robots and not something already written explicitly in the passage, and that this should be a prediction based on the *analogy* and not on the target passage alone.

Prediction Rating Task. After making their own predictions, subjects rated possible predictions about the target according to how well they followed from the analogy. Subjects were told:

> Professor Zee answered the same question you just did. She said that the analogy with the Tams suggested two things about the Robots. She made the two predictions stated below. Neither of these were explicitly written about the Robots though both of these predictions are equally plausible. But, are these predictions suggested by the analogy with the Tams? Rate these predictions on a scale of 1 to 7 according to how well they follow from the analogy. (Note: one of her predictions may be the same as the one you just made).

The two key facts that appeared in the base but not in the target were then given to subjects as predictions about the target. For example, subjects were given:

1. *A Tam sometimes stops using its underbelly. Similarly, sometimes a robot shuts down its probes.*
2. *Tams cannot go from one rock to another. Similarly, Robots cannot go from their current planet to another planet to gather data.*

Prediction Choice Task. Finally, subjects were again presented with the two predictions on a new page and asked to "Choose the prediction that is best suggested by the analogy." Subjects were also asked to write explanations for choosing one prediction and rejecting the other. For all tasks subjects were told to refer back to the base and target passages while deciding on a response.

Task for Target-Only Subjects. Subjects who only read the target domain performed *prediction, rating* and *choice* tasks which paralleled the tasks given to analogy subjects. Before completing these tasks, subjects were asked to summarize the target by answering general questions that directed them to the essential parts of the passage, for example, "Explain what happens when the data gathered by the robots are no longer new." These questions were intended to ensure that subjects thoroughly attended to the parts of the passages describing the causal structures and key facts.

Prediction Task. Following the learning task, subjects were asked to make a prediction about the target passage. Instructions were similar to those given to analogy subjects, except that subjects were told to make a prediction about the target "given the information in the target passage."

Prediction Rating Task. After making their own predictions, target-only subjects, like analogy subjects, were told that a fictional person had made two predictions about the target passage. Subjects were given the same two predictions given to analogy subjects, except that there was no mention of the analogous facts from the base domain:

1. *Sometimes a robot shuts down its probes.*
2. *Robots cannot go from their current planet to another planet to gather data.*

TABLE 5
Experiment 2. Results of Prediction Task: Proportion of Subjects in Each Group Predicting Shared-System (Correct) and Different-System Key Facts for Each Passage

	[Analogy Group (n=16)]			[Target-Only Group (n=16)]		
Passage	Shared System	Different System	Other	Shared System	Different System	Other
1	.44	.06	.44	.06	.25	.69
2	.62	.06	.25	.06	0	.94
3	.44	.19	.25	.06	.12	.81
4	.62	.19	.12	.31	.12	.56
Overall	.53	.12	.27	.12	.12	.74

Note. A few analogy subjects predicted both the shared-system and different-system fact (.08 of all responses). These responses are not shown here.

They were asked to rate these according to "how well they followed from, or were predicted by, information in the target passage."

Prediction Choice Task. Finally, subjects were given the predictions again and asked to "choose which claim best follows from the information in the passage," and to explain their choices. (As in Experiment 1, these explanations were not coded and will not be mentioned further.)

Results

Prediction Task

Two judges, who were blind to the condition of subjects being scored, grouped subjects' predictions into three categories[6]: (a) predictions of key facts that would follow from the shared causal system in the base and target (b) predictions of key facts that would follow from the different-system (c) all other predictions about the target. Interrater agreement before discussion was .88. Disagreements were readily resolved. A few subjects (.08 of the responses) predicted both the shared-system and different-system fact. These responses were discarded and omitted from further analysis.

As predicted, analogy subjects most frequently predicted the shared-system rather than the different-system key fact. Table 5 shows that, across the four analogies, .53 of the analogy subjects' responses were predictions of the shared-system fact, and only .12 were predictions of the different-system fact. As shown, the same pattern of results holds for each of the four passages used. Thus, although both predictions were clearly possible, analogy subjects made the inference that was connected to a larger matching structure.

Two examples of shared-system predictions by analogy subjects for the *Tams/Robots* analogy follow. These subjects were given Version 2 of the

[6] In some cases the content of subjects' responses would indicate which condition the subject was in. However, one judge was blind to the experimental manipulation.

Robots passage, in which the shared-system prediction was that robots cannot go from one planet to another to gather data.

1. I would predict that once the robots were specialized they would be unable to probe for data on other planets than what they were used to just as the Tams would not have the right textured underbelly for a new kind of rock.
2. The robot may eventually be strictly unable to switch planets as the Tams cannot switch rock types.

The frequency of shared-system predictions among analogy subjects cannot be attributed to a bias in the materials. The responses of the target-only subjects indicate that the shared-system predictions were not highly salient or obvious predictions in the target domains. Not surprisingly, the most frequent response for the target-only subjects was to predict information other than the key facts. Some of these were rather creative. For example, two target-only subjects given Version 2 of the target predicted:

1. Since the robots are so sensitive to the different planets and they need to develop their own probes which related to that particular planet it may be predicted that the robots have trouble analyzing data and make incorrect assumptions and conclusions.
2. Robots are able to control the spaceship which takes them from planet to planet.

When target-only subjects did predict one of the key facts, their responses were evenly distributed between the two types of key facts.

To assess whether the preference for shared-system predictions among analogy subjects was reliable, subjects were scored for the number of shared-system predictions minus the number of different-system predictions across the four analogies. (Note that shared system and different system predictions do not exhaust the possible responses.) A 2×2 ANOVA (the variables were condition and the counterbalancing variable version-set) showed that the mean difference for analogy subjects was significantly higher than the mean for target-only subjects ($M = 1.63$, and $M = 0$ respectively, $F(1, 28) = 20.05$, $p < .001$. The only other significant result was the main effect for version-set, $F(1, 28) = 9.61$, $p < .01$, simply indicating that, overall, the number of shared-system predictions varied across version-set. No interactions were significant. Thus, a preference for shared-system predictions was found among analogy subjects but not target-only subjects and this was true for each version-set.

Prediction Rating Task

The mean ratings of each fact-type for each target are shown in Table 6. As predicted, when asked to rate key facts according to how well they were predicted by the analogy, analogy subjects gave higher ratings to shared-system than to different-system facts (overall $M = 6.03$ and $M = 4.41$, respectively).

TABLE 6
Experiment 2. Prediction Rating Task: Mean Ratings[a] of Shared-System (Correct) and Different-System Predictions for Each Passage

	[Analogy Group (n=16)]				[Target-Only Group (n=16)]			
	Shared System		Different System		Shared System		Different System	
Passage	M	SD	M	SD	M	SD	M	SD
1	6.43	1.50	3.69	2.57	4.00	2.31	4.50	2.66
2	5.50	1.93	4.44	2.13	5.31	2.02	5.56	1.46
3	6.12	2.03	4.75	2.14	3.50	2.03	5.87	1.36
4	6.06	1.44	4.75	1.88	4.50	1.97	4.94	1.77
Overall	6.03	1.16	4.41	1.51	4.33	1.03	5.22	.76

[a] Predictions were rated on a 1–7 scale.

The ratings of target-only subjects, who rated key facts according to how well they were predicted by the target passage, showed no such difference (and, if any, showed the reverse preference), indicating that the materials were not biased in favor of the shared-system prediction (overall $M=4.32$ and $M=5.22$, respectively, for shared- and different-system facts).

As in Experiment 1, a $2 \times 2 \times 4 \times 2$ repeated-measure ANOVA was conducted to test the difference in ratings of shared-system and different-system facts in each condition. In addition to the experimental variables, condition and fact-type, the analysis also included passage, and the counterbalancing variable, version-set. Both subjects and passage were treated as random variables in this analysis.

As expected, no main effects were significant. The predicted interaction between condition and fact-type was significant $F'(1, 9) = 10.99$, $p < .01$. The analysis of the simple effect of fact-type within the analogy condition confirmed that analogy subjects gave significantly higher ratings to shared-system facts than to different-system facts $F'(1, 12) = 7.86$, $p < .05$. This difference was not found for target-only subjects. No other effects tested by the overall ANOVA were significant.

Prediction Choice Task

The results for subjects' choice of which key fact was the best prediction also support the systematicity hypothesis. Analogy subjects, but not target-only subjects, preferred the shared-system predictions. When subjects are scored for the number of choices of shared-system key facts across the four analogies (possible score is 0–4), a 2×2 (Condition $\times$ Version-Set) ANOVA revealed that the mean for analogy subjects ($M=2.9$) is significantly greater than the mean for target-only subjects ($M=1.8$), $F(1, 28) = 16.43$, $p < .001$. As expected, no other effects were significant.

In contrast to the previous two tasks, here the results across the individual passages are somewhat varied. The pattern of results holds for three of the four passages, though the difference between the analogy and target-only groups reaches significance only for two passages. For these two passages, analogy subjects chose the shared-system fact 87% and 69% of the time, whereas target-only subjects chose this fact only 44% and 25% of the time, ($p < .05$, Fisher exact, one-tailed tests.) For the remaining two passages, analogy subjects chose the shared-system fact 62% and 75% of the time, and target-only subjects chose it 69% and 44% of the time; these differences between groups are not significant.[7]

Justifications of Choices

As in Experiment 1, analogy subjects were asked to write a brief explanation for their choice task responses. They were to (a) explain why they thought their *chosen* prediction followed best from the analogy, and (b) explain why the *rejected* prediction did not follow from the analogy as well. The justifications for choices and rejections were separately scored. Our main interest was in whether subjects would refer to the shared or differing causal systems supporting a prediction in justifying their responses.

Two judges coded subjects' justifications for choosing a prediction into four categories similar to those used in Experiment 1. Responses in the first category show a concern for the higher-order embedding of the key prediction, and responses in the second simply show a concern for the similarity of the prediction to the base key fact. For responses in the final two categories it was either not clear whether subjects were responding on the basis of the analogy, or it was not clear *how* the analogy guided their responses. Thus, responses were categorized according to the level of focus:

1. *Focus on the similar causal structure* (*higher-order similarity*). The subject described the cause for the chosen prediction that is similar between the base and target, or asserted *that* there is a similarity in cause.
2. *Focus on the similarity of the prediction* (*lower-order similarity*). The subject simply noted the similarity or correspondence between the chosen prediction and the corresponding fact in the base.
3. *Focus on the prediction alone—on its plausibility* (*properties of the target*). Subjects in this category did not refer to the base domain. They

[7] Note that for one passage the percentage of target-only subjects choosing the shared-system prediction is 69%. This gives a slight suggestion that this passage may be biased in favor of the systematicity predictions (though the percentage of analogy subjects making the shared-system choice is smallest, 69% for this passage). The possibility that this passage was biased in favor of the systematicity predictions suggested that the analyses of the *prediction task* data should be reconsidered in the absence of this passage. The original findings still remain. That is, the ANOVA comparing the number of shared-system minus difference-system key facts predicted in each condition still shows a significant effect for condition even when predictions for this possibly biased passage are not included, $F(1, 28) = 13.03$, $p < .001$.

either stated that the prediction was likely because of the causal information in the target, or simply asserted that the fact was plausible in the target.

4. *Other.* Gave any other response.

Subjects' justifications for rejecting a fact were classified into the same four categories, except that now justifications focused on: *dissimilarities* in the cause of the rejected prediction to the cause for the corresponding fact in the base (some subjects who focused on the dissimilar cause also stated that the prediction was unlikely in the target), the *dissimilarity* of the rejected prediction to the corresponding fact in the base, or the *implausibility* of the rejected prediction in the target.

Table 7 shows the justification data. As before, the left side of the table shows the distribution of responses when the shared-system prediction was chosen. The right side shows responses when the different-system prediction was chosen. The top half of the table shows justifications for choosing a prediction; the bottom half, for rejecting the other prediction. Interrater agreement for scoring of choice justifications according to the four categories was .84 before discussion. Interrater agreement for scoring rejection justifications was .84 before discussion.

As shown in the left side of Table 7, when subjects chose the shared-system prediction, the majority of the justifications focused on the similar causal information in the base and target. Correspondingly, their justifications for rejecting the different-system prediction focused on the dissimilar causal information.

Examples of such causal justifications follow. Again examples are taken from subjects given Version 2 of the *Robots* passage in which the shared-system prediction is that robots cannot go from one planet to another to gather data, and the different-system prediction would be that the robots shut down their probes.

1. The second prediction [robots cannot go from planet to planet] follows the analogy best. Robots will not be able to go from one planet to the next. We can assume this by looking at the story from the Tams. Since their underbellies (Tams) and filters and sensitivities (robots) are specified they can't go from one place to the other. The first prediction [robots shut down probes] doesn't follow the analogy because the Tams stop using its underbelly because there are no nutrients left. While the robot doesn't stop because there is no data, rather it will overheat if it doesn't.
2. Two [robots cannot go from planet to planet] is better because it relies on the analogy that both the Tams and the robots specialize to the extent

TABLE 7

Experiment 2. Justifications by Analogy Subjects for Predictions Selected in Choice Task: Proportion and Frequency of Shared-System and Different-System Choices (and Rejections)

	Chosen Prediction			
	Shared-System (Correct)		Different-System (Incorrect)	
Justification Type	Prop.	Freq.	Prop.	Freq.
Similar cause for chosen fact	.77	(36)	.29	(5)
Similarity of chosen fact itself	.04	(2)	.18	(3)
Likelihood of chosen fact in target	.06	(3)	.29	(5)
Other	.13	(6)	.23	(4)
Total	100	(47)	100	(17)

	Rejected Prediction			
	Different-System (Correct)		Shared-System (Incorrect)	
Justification Type	Prop.	Freq.	Prop.	Freq.
Dissimilar cause for rejected fact	.55	(26)	.29	(5)
Dissimilarity of rejected fact itself	.13	(6)	.23	(4)
Likelihood of rejected fact in target	.19	(9)	.29	(5)
Other	.13	(6)	.18	(3)
Total	100	(47)	100	(17)

Note. Out of 64 opportunities, subjects chose the shared-system prediction (the predicted response) 47 times and the different-system prediction 17 times. For each choice, justifications are shown for both the chosen prediction and the rejected prediction.

> that they are not transferable from rock to rock or planet to planet. One [Shut down probes] isn't good because when a Tam stops using its underbelly, it has exhausted the supply of minerals: whereas when a robot stops probing, it has been getting too much data and must shut down to avoid overheating.

These comments reveal subjects' belief that predictions must be based on a shared higher-order structure. Few justifications were concerned only with the match between the target prediction itself and the corresponding base fact. Thus, the justifications for selecting predictions are consistent with the justifications for selecting matches found in Experiment 1. When subjects can make explicit their reasons for selecting information to map from a base to a target domain, their inferences are guided by systematicity and shared structure.

Discussion

The results of Experiment 2 indicate that systematicity constrains analogical inference processes. That is, systematicity determines which predicates in the base domain will be imported as predictions into the target domain. In the prediction task, either the shared-system or different-system prediction was possible: Each was a fact given in the base but not in the target, each could be easily constructed in the target, and each had antecedent conditions in the target. Yet subjects showed a strong preference for making the prediction that was supported by antecedent conditions that matched the base domain. The results of the rating and choice tasks provide converging evidence that subjects prefer predictions sanctioned by systematicity. Overall, the predictions linked to a matching causal system were rated most highly and chosen most often as the predictions that follow well from the analogy. Furthermore, subjects explicitly focused on the matching causal structure in their choice justifications. Finally, the difference in performance between the target-only and analogy subjects for each of the three tasks indicates that the results cannot be attributed to a bias in the materials in which the shared-system facts were inherently more salient or plausible in the target.

Convergent with the results of Experiment 1, the results of Experiment 2 indicate that analogical mapping concerns corresponding systems of predicates and not merely independent correspondences among individual predicates. Subjects ignored or rejected possible predictions that represented an isolated correspondence between the base and target, even though in themselves these predictions created a good match with the base. Rather, subjects generated analogical predictions that were supported by a larger systematic matching structure. It appears that in generating candidate inferences, just as in selecting predicates that belong to a match, people tacitly and sometimes explicitly seek connectivity and interdependency.

EXPERIMENT 3

In Experiments 1 and 2, subjects were always allowed to examine the base and target passages as they made mapping judgments and predictions. In the next experiment we asked whether the systematicity constraint would operate when subjects had to rely on their memory representations of a base domain. In ordinary life, people frequently reason by analogy without the benefit of written material. Thus, our tasks so far could be allowing subjects to attain unrealistic levels of rigor in processing, or worse, could be somehow suggesting an unnatural strategy. In the more natural case, when people reason from a stored representation of a base domain, perhaps they are content with whatever correspondence they can most readily create with the target. Their selection of information to map may be less constrained by a concern for shared structure.[8] We were somewhat reassured by evidence of various kinds that systematicity plays a role in natural analogizing, for example, in scientific reasoning (Burstein, 1983; Clement, 1983; Gentner & Gentner, 1983; Gentner & Jeziorski, 1898). However, we wanted to investigate the generality of the phenomenon. The tasks here check whether systematicity operates as a selection constraint even when subjects had to rely on a base domain represented in memory.

A second reason for conducting this experiment was to confirm that subjects' responses in the previous two experiments were guided by similarities in the underlying structures of the base and target passages, and not by uninteresting superficial features of the passages. In writing the analogies for all the experiments we attempted to avoid similarities in the surface form of sentences used to describe matching causal systems. However, replicating the previous experiments with memorized base domains would confirm that conceptual similarities in the absence of surface commonalities can support the systematicity constraint. (Note that we assume here that when material is committed to memory, the semantic content is better represented than the surface form of information.)

Thus, in Experiment 3, we again examined the effects of systematicity on the analogical predictions subjects make about a target domain. The basic method of this experiment was identical to that of Experiment 2, except that subjects first committed the base domains to memory.

Method

Subjects

Subjects were 48 paid undergraduate students at the University of Illinois. Half were assigned to the analogy condition and half to the target-only condition.

Materials

Materials were designed identically to those used in Experiment 2. Three rather than four analogies were used because of the added length of the learning tasks. Two of the analogies were the same as those used in Experiment 2 (though modified slightly to facilitate comprehension) and one was a new analogy (derived from a previous analogy).

Procedure

Task for Analogy Subjects. Subjects first memorized the base domain. Then, given the target domain, subjects performed learning, prediction,

[8] Note that we are not talking here about how people *access* analogs in memory. This is generally agreed to be strongly influenced by surface similarities (e.g., Gentner & Landers, 1985; Holyoak & Koh, 1987; Ross, 1984, 1989). Rather we are concerned with how the analogical match is constructed given that a particular analog has been accessed.

and judgment tasks identical to those in Experiment 2, except that subjects now could not refer back to the base passage.

Memorizing the Base Domain. Subjects studied a base domain for five minutes and then summarized it from memory. Next they reviewed the base passage and corrected or elaborated their summary by comparing it to the passage; no feedback was given by the experimenter. Finally, the base passage was removed and a multiple-choice test was given to assess subjects' understanding of the central events in the base passage. The test included questions about each key fact and causal system. Note that we tested knowledge of the causes for both key facts, and all subjects received the same test. Thus, the test did not bias subjects' attention toward a particular causal system. Subjects were given no feedback after their test.

To promote energetic performance on the memorization task, subjects were given a monetary motivation. That is, they were told that if they did well on their summaries and multiple-choice test, they would be paid an additional two dollars at the end of the session.

Learning Tasks. After learning the base domain, subjects were given the target domain and, as in the previous experiments, worked out the object correspondences and described the ways in which the base and target were and were not analogous.

Prediction and Judgment Tasks. A prediction task, in which subjects made their own prediction about the target domain, as well as rating and choice tasks were given which were identical to those in Experiment 2. The instructions were altered slightly for clarification. Subjects were allowed to refer back to the target passage but not to the base passage.

Task for Target-Only Subjects. As in Experiment 2, in order to determine that the shared-system predictions were not simply the most plausible predictions in the target, a group of subjects made predictions and judgments based only on the information in the target passage.

Results

Prediction Task

Two judges who were blind to the condition of subjects being scored[9] grouped predictions into the same categories as in the previous experiment: (a) predictions of key facts that would follow from the shared causal system in the base and target, (b) predictions of key facts that would follow from the

[9] Again, in some cases the content of subjects' responses would indicate which condition the subject was in.

TABLE 8
Experiment 3. Prediction Task: Proportion of Subjects in Each Group Predicting Shared-System and Different-System Key Facts for Each Passage

Passage	[Analogy Group (n=24)] Shared System	Different System	Other	[Target-Only Group (n=24)] Shared System	Different System	Other
1	.46	.21	.29	.04	.08	.88
2	.375	.375	.17	.04	0	.96
3	.75	.08	.08	.25	.125	.625
Overall	.53	.22	.18	.11	.07	.82

Note. A few analogy subjects predicted both the shared-system and different-system fact (.07 of all responses). These responses are not shown here.

different system, and (c) all other predictions about the target. (A few subjects, .07 of the responses, predicted both the shared-system and different-system fact. Again these responses were discarded and omitted from further analysis.) Interrater agreement before discussion was .96. Disagreements were readily resolved.

Table 8 shows the proportion of responses in each category. As predicted, analogy subjects most frequently predicted the shared-system key fact. Furthermore, responses of the target-only subjects indicate that this prediction was not especially salient in the target passages. As before, in order to assess whether the preference for shared-system predictions among analogy subjects was reliable, subjects were scored for the number of shared-system minus different-system facts predicted across the three passages. The mean for the analogy group is .92 and the mean for the target-only group is .13. A 2×2 ANOVA (the variables were condition and the counterbalancing variable version-set) showed that the predicted main effect for condition was significant $F(1, 44) = 5.15$, $p < .05$. Analogy, but not target-only subjects, predicted more shared-system than different-system facts. The ANOVA also showed a significant main effect for version-set, simply indicating that, overall, the number of shared-system predictions varied across version-set, $F(1, 44) = 4.12$, $p < .05$. As expected, there was no interaction between condition and version-set.

Table 8 shows that for one passage the data are not consistent with the systematicity predictions. For Passage 2, analogy subjects predicted each fact-type equally often. These inconsistent results simply appear to be due to some subjects' inaccurate representations of the causal information in the base passage for this analogy. As will be discussed later, several subjects incorrectly answered the multiple-choice test questions about the base domain causal structure, and these errors were apparently related to responses on the prediction, as well as the rating and choice tasks.

In sum, consistent with the result of Experiment 2, analogy subjects tended to make the prediction sanctioned by systematicity. Because subjects

TABLE 9
Experiment 3. Prediction Rating Task: Mean Ratings[a] of Shared-System (Correct) and Different-System Predictions for Each Passage

	[Analogy Group ($n=24$)]				[Target-Only Group ($n=24$)]			
	Shared System		Different System		Shared System		Different System	
Passage	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
1	5.96	1.92	4.46	2.30	3.33	1.95	5.83	2.16
2	5.75	1.75	5.50	2.02	5.29	2.05	5.21	2.06
3	5.79	1.72	4.71	2.37	4.29	1.83	5.12	1.94
Overall	5.83	1.08	4.89	1.70	4.31	1.34	5.39	1.18

[a] Predictions were rated on a 1–7 scale.

had to memorize the base domains in this experiment, the previous findings do not appear to be specific to the artificial situation of having a written base domain available for inspection. This experiment shows that analogical inferences are based on a shared higher-order structure even when subjects must rely on their memory representations of the base domain.

Prediction Rating Task

Table 9 shows the mean ratings for shared-system and different-system key predictions. Again, as predicted, analogy subjects rated shared-system facts more highly than different-system facts (overall $M=5.83$, and $M=4.89$, respectively) whereas target-only subjects showed the reverse pattern of ratings ($M=4.31$ and $M=5.39$, respectively).

A $2\times2\times3\times2$ (Condition × Fact-Type × Passage × Version-Set) repeated-measures ANOVA was conducted. (Passage was treated as a fixed-effects variable because the number of passages was small.) There was a significant main effect for condition, $F(1, 44)=4.37$, $p<.05$, simply indicating that overall, ratings varied across the two groups. No other main effects were significant.

As before, the key prediction is an interaction between condition and fact-type. This interaction was significant $F(1, 44)=10.82$, $p<.01$. Also, the simple effect of fact-type within the analogy group showed that the difference in ratings to shared-system and different-system facts by these subjects is significant, $F(1, 22)=5.25$, $p<.05$. Interestingly, the difference in ratings to the two fact types is in the opposite direction for target-only subjects (these subjects gave higher ratings to different-system facts) and this effect is also significant, $F(1, 22)=5.58$, $p<.05$. This finding suggests that the materials were biased *against* the systematic choice.

Finally, the triple interaction, Condition × Fact-Type × Passage, was significant $F(2, 88)=5.41$, $p<.01$. As previously noted, responses to Passage 2 do not follow the predicted pattern (again, apparently due to poor memory for the base passage), but the predicted interaction between condition and fact-type holds for the other two passages. No other effects were significant.

Overall, the rating task results again confirm findings of the previous experiment. With the exception of one analogy, shared-system facts were rated higher than different-system facts. It is interesting that these results were found even when materials were apparently biased against the shared-system fact. The analogy subjects valued a prediction which followed from the systematicity of the analogy even when an alternative prediction may have been more salient or plausible in the target domain.

Prediction Choice Task

Results for the choice task also support the systematicity hypothesis. Analogy subjects, but not target-only subjects, preferred the shared-system prediction. Subjects were scored for the number of choices of shared-system predictions across the three passages (possible score is 0–3). A 2×2 ANOVA (Condition $\times$ Version-Set) showed that the mean number of shared-system predictions chosen in the analogy group ($M = 1.875$) was significantly greater than the number chosen in the target-only group ($M = 1$), $F(1, 44) = 13.15$, $p = .001$. No other effects were significant. Considering the passages individually, again the predicted pattern does not hold for Passage 2 (the proportion of subjects choosing the shared-system fact is .54 and .58 in each group). However, the analogy and target-only groups differ significantly for the other two passages. For each of these passages, .66 of the analogy subjects chose the shared-system fact, in contrast to only .08 and .33 of the target-only subjects (Fisher exact, $p = .001$ and $p = .05$).

Choice Task Justifications

As in the previous two experiments we were interested in subjects' explicit reasons for their choice task responses. Justifications for choices and rejections of key predictions were coded in the same manner as in Experiment 2. Table 10 shows that the distribution of responses replicates the findings of the previous experiment: Subjects who made the systematic selection focused on the similarity or dissimilarity of causal information for the key predictions. Few of these subjects were concerned only with whether the prediction itself corresponded to the base domain. Thus, subjects' explicit criteria for a good analogical prediction concerned the connection of the prediction to a causal structure found in both domains.

Analogy Group Responses as a Function of Comprehension of the Base Passage

After analogy subjects read and memorized the base passage, they answered multiple-choice questions about the central events in the passage. Analysis of the relation between performance on this test and performance on the analogy tasks helps clarify the inconsistent findings for Passage 2. Performance on the multiple-choice test for Base Passage 2 was poor relative to the other passages. Specifically, on the test for Base Passage 2, 42% of the

TABLE 10

Experiment 3. Justifications by Analogy Subjects for Predictions Selected in Choice Task: Proportion and Frequency of Shared-System and Different-System Choices (and Rejections)

	Chosen Prediction			
	Shared-System (Correct)		Different-System (Incorrect)	
Justification Type	Prop.	Freq.	Prop.	Freq.
Similar cause for chosen fact	.69	(31)	.30	(8)
Similarity of chosen fact itself	.09	(4)	.30	(8)
Likelihood of chosen fact in target	.09	(4)	.18	(5)
Other	.13	(6)	.22	(6)
Total	100	(45)	100	(27)
	Rejected Prediction			
	Different-System (Correct)		Shared-System (Incorrect)	
Justification Type	Prop.	Freq.	Prop.	Freq.
Dissimilar cause for rejected fact	.64	(29)	.11	(3)
Dissimilarity of rejected fact itself	.11	(5)	.15	(4)
Likelihood of rejected fact in target	.07	(3)	.30	(8)
Other	.18	(8)	.44	(12)
Total	100	(45)	100	(27)

Note. Out of 72 opportunities, subjects chose the shared-system prediction (the predicted response) 45 times and the different-system preoiction 27 times. For each choice, justifications are shown both for the chosen prediction and the rejected prediction.

subjects made an error, but on the tests for Base Passages 1 and 3, only 8% and 29% of the subjects made an error. (Most subjects made only one error out of five to eight questions.) Furthermore, 12 of the 13 Passage 2 errors were to questions that specifically addressed the cause for one of the key facts. Consistent with this, Table 11 shows that the subgroup of subjects who were error-free on the multiple-choice tests were more likely to respond to the analogy tasks according to the systematicity predictions than was the analogy group as a whole. (Compare Tables 8 and 9.) This is especially true for Passage 2. These results must be interpreted with some caution, because with the removal of subjects, version-set is no longer fully counterbalanced. However, the results indicate that the unexpected analogy task results discussed above for Passage 2 are due to some subjects' failure either to recall or understand the causal structure of the base domain.

TABLE 11
Experiment 3: For each Analogy, Performance on the Prediction, Rating and Choice Tasks Among Analogy Subjects Who Were Error-Free on the Base Passage Test

	Prediction Task: Proportion of Subjects Making Each Prediction		Rating Task: Mean Ratings to Each Prediction		Choice Task: Proportion of Subjects Choosing the Shared-System Prediction
Passage	Shared System	Different System	Shared System	Different System	
1	.50	.18	6.22	4.27	.73
2	.50	.28	6.50	4.93	.71
3	.88	.06	5.59	4.35	.76

Note. The number of error-free subjects for each passage was 22 for Passage 1 (half in each Version-Set); 14 for Passage 2 (8 in Version-Set 1 and 6 in Version-Set 2); and 17 for Passage 3 (7 in Version-Set 1 and 10 in Version-Set 2).

Discussion

Overall, the prediction, rating, and choice task responses, and subjects' explicit justifications, replicate the findings of Experiment 2. Thus, the previous findings are apparently not the result of subjects' reliance on surface, textual features of the passages, rather than on the causal structure of the passages. Furthermore, subjects' use of systematicity as a selection constraint is apparently not restricted to the situation in which a written base domain is available for inspection. In this experiment, as in typical cases of analogical reasoning, subjects made analogical inferences by drawing on their memory representations of the base domain. As before, the selection of analogical inferences appeared to be guided by a concern for the connection of these inferences to a shared causal structure. The effects are somewhat weaker than observed in the previous experiments, apparently because subjects did not always possess an accurate record of the base domain. However, in general, subjects followed the same rules which constrain good analogical predictions to those that follow from a larger matching system.

GENERAL DISCUSSION

An analogy is a selective form of comparison. Of the indefinitely large number of common features that two situations may share, only certain kinds of commonalities count in analogical mapping. These experiments examined one possible selection constraint, namely, *systematicity* (Gentner, 1983, 1989): The principle that people (a) preferentially include in an analogical mapping those matches that are embedded in a higher-order structure containing other matching elements, and (b) make new inferences in the target domain by mapping across further elements of the base domain that belong to a largely matching structure.

Our first experiment concerned the selection of which matches belong to the interpretation of an analogy. Subjects were required to judge which of two possible matching facts contributed best to an analogy. Their mapping preferences and their explicit justifications indicate that they were guided by systematicity: Subjects preferred the match that was linked by a higher-order relation to neighbors that also matched. Experiments 2 and 3 concerned how people draw predictions from an analogy. We gave subjects a substantially less constrained task than was used in Experiment 1: Subjects were required to make their own analogical predictions about a target domain. The base and target passages included many existing matches; however, two facts were always present in the base but absent in the target. Again, interconnections among the facts appeared to guide mapping: Subjects consistently predicted a fact in the target that was connected to a matching antecedent in the base. Thus, subjects made predictions that reflected the same relational interdependency in the base and target. The results of Experiment 3 showed that these findings do not depend on having a written description of the base domain available for inspection. We found the same pattern of results in a more typical situation in which subjects simply relied on their memory representations of the base domain.

Subjects' mapping choices cannot be accounted for by differences in how well the facts themselves matched between domains, or differences in the plausibility of the facts in the target. First, since two different versions of a target domain were used, the same key fact was sometimes embedded in a matching causal structure, but other times embedded in a nonmatching causal structure. Thus, effects of the quality of the match itself, or of other properties of the specific key facts, were controlled. Second, the validity (truth value) or plausibility of the key facts as applied in the target domain could not have determined choices. All analogies described fictional domains, and in Experiment 1 both facts, as given, were asserted to be true about these domains. In Experiments 2 and 3, antecedents for both key predictions were present in the target, and furthermore, responses of target-only control subjects indicated that both consequent predictions were equally plausible.[10]

[10] It might be argued that subjects preferred the shared-system match simply because it was *preceded by* other matching information, and the higher-order link between the key match and preceding matches was irrelevant. In contrast, on our account the higher-order link is needed. That is, if we had preceded a matching key fact with totally unrelated information that also matched between domains, this would not have been sufficient to determine subjects' choices. For example, preceding the key matching fact, *robots shut down their probes*, with *eggplants come in many varieties* would be unlikely to increase subjects' preference for the key fact. Subjects' response justifications also show that the higher-order link between a matching fact and neighboring matches is important. Subjects often cited the similarity in causal information as the reason for their choice.

Thus, the systematicity principle constrains both which *matches* between a base and target domain are included in the interpretation of an analogy and which *inferences* are drawn from an analogy. These results confirm that the choice of lower-order relations to map is not determined just by the independent relations themselves, but by the interconnections among these relations.

How Systematicity Might Operate as a Selection Constraint: Evidence from Computer Simulation

The results here indicate that systematicity must be part of a descriptive theory of analogy processing. Can we go beyond the descriptive theory to suggest processing principles? The challenge here is to come up with a set of relatively simple processes which, working in combination, can accomplish the sophisticated feat of attending to higher-order interconnections during comprehension of an analogy. We now describe a computer simulation that accomplishes this task in a psychologically plausible manner.

The Structure Mapping Engine (SME; Falkenhainer, Forbus, & Gentner, 1989), explicitly employs structure-mapping principles to simulate the process of interpreting and making predictions from an analogy. SME begins by finding all possible local matches between the base and target. At this stage the program may have a large number of mutually inconsistent local matches. It next collects these local matches into *global mappings:* consistent sets of matching predicates. SME finds the largest possible systems of matched predicates with consistent object mappings. These global mappings are the possible interpretations of the analogy. As a further step, SME uses the analogy to suggest predictions about the target domain. For each interpretation, if there is a predicate connected to the base system, but not found in the target system, this predicate becomes a candidate inference in the target. Finally, each interpretation is given a structural evaluation. This is based partly on the number and kind of local matches, but, more interestingly, it also depends on the depth of the system of matches. (This preference for systematicity is currently implemented in a trickle-down algorithm in which the evidence for a given match is increased if there are also matches among its parent predicates; see Forbus & Gentner, 1989.)[11] Thus, in SME, systematicity enters into analogy processing both in evaluating an interpretation and in deriving inferences.

From a psychological point of view there are several interesting features of this model. First, the same processes that enter into forming mappings also lead to analogical inferences. Second, the model begins in a blind and local fashion by finding local identities and derives its final (potentially rather sophisticated) interpretation simply by combining these local matches

[11] SME can be run in an exhaustive mode in which all possible interpretations of an analogy are generated and evaluated, or it can be run in a selective mode in which only one interpretation is produced (Forbus & Oblinger, 1990).

into connected systems. Third, preexisting goals are not required to form a coherent matching structure or to generate inferences; candidate inferences can be drawn solely on the basis of shared structure. Thus, this model is applicable to analogies based on common human goals, and also to analogies based, for example, on common patterns of logical or geometrical relations. SME has successfully simulated human performance in judgments of analogical soundness (Skorstad, Falkenhainer, & Gentner, 1987), similarity ratings (Rattermann & Gentner, 1987), performance in a picture-mapping task (Markman & Gentner, 1990), and children's performance on an analogical mapping task (Rattermann & Gentner, 1990) and a similarity choice task (Kotovsky & Gentner, 1990).[12]

Connections to Text-Processing Models

There are interesting links between this research and models of comprehension of single texts. According to several models, readers focus on interrelations among information within a text, such as higher-order causal or explanatory relations, to develop coherent text representations (Kintsch & van Dijk, 1978; Meyer, 1984; Trabasso & Van Den Broek, 1985). This research indicates that it is reasonable to suppose that the inputs to the process of seeking analogies between texts are structured representations. Furthermore, there is some similarity in spirit between SME's local to global process, and Kintsch's (1988) construction-integration model of discourse comprehension. This model begins locally and then uses connections to develop a final text interpretation. However, models of text processing also underscore the need for further selectivity in analogical processing. As Kintsch and van Dijk (1978; van Dijk, 1980; van Dijk & Kintsch, 1983) make clear, text representations are complex and multidimensional, involving several levels of representation. Because analogy, by its nature, is a comparison between two situations, it requires processes of selection, mapping, and alignment between two representations, in addition to the processes involved in the comprehension of single texts. An account of analogical processing must specify

[12] The effects of systematicity may extend beyond those demonstrated in this research in two ways. In Experiment 1, we found that embedding in a shared structure determined which of two *matching* relations contributed to an analogy. Future research should examine the extent to which shared structure may allow *nonmatching* lower-order relations to be put into correspondence. That is, in the same way that dissimilar objects are mapped onto one another by virtue of their similar roles in a larger system, perhaps dissimilar lower-order relations may be mapped by virtue of their similar roles. For example, *differing* antecedent relations for the same consequent may be put into correspondence on the grounds that they lead to this same consequent (perhaps the relations are rerepresented so that they have a similarity defined by their shared role). A meaningful analogy must include many matching relations. However, mismatching relations that are components of a much larger matching structure may be put into correspondence, despite their intrinsic dissimilarity, exclusively on structural grounds.

how subjects, in the absence of advance knowledge, select among the many possible mappings between two texts. The results here suggest that one principle subjects use is to evaluate local matches according to their embedding in a larger matching structure.

Goal Relevance and Structural Constraints

Several investigators have stressed that analogical processes take into account the goals of the reasoner or the goals present in the analogous domains. There have been various ways in which the relationship between goals and analogy have been discussed. One view, as discussed earlier, is that an analogy is similarity defined with respect to the reasoner's current goals, and preexisting goals, rather than structural principles, are necessary to guide mapping choices (Holyoak, 1985). Such a view is incompatible with the results here, which indicate that the relational structure intrinsic to the analogous domains can determine mapping, even in the absence of an extrinsic goal. A second view is that analogies tend to be about goals and plans. That is, the higher-order relations that govern the analogy are often goal structures. Examples include many analogies that arise in problem-solving and in case-based reasoning (Gick & Holyoak, 1983; Holyoak & Koh, 1987; Keane, 1988; Kolodner, 1987; Kolodner, Simpson, & Sycara, 1985; Schank, 1982). This view is not inconsistent with structure mapping and, indeed, can be seen as a special case. If we consider a goal scheme as a particular case of a relational structure, then the processes we have discussed here are sufficient to interpret analogies with goal structures. This has the advantage of parsimony: We can assume the same processes across different contents and external contexts.[13]

Some recent models have explored the ways in which goal relevance can augment structural constraints when analogies are used to achieve a reasoner's goal. Several computational accounts of analogy have attempted to integrate structural and goal-relevance constraints on mapping (Burstein, 1983; Burstein & Adelson, 1987; Holyoak & Thagard, 1989; Kedar-Cabelli, 1985a, 1985b; Thagard & Holyoak, 1988). Goal relevance and structural constraints may interact in several interesting ways. One possibility is that goals directly influence the mapping process; for example, predicates may be weighted for goal relevance (Thagard & Holyoak, 1988). An alternative possibility is that

[13] It should be noted that the findings here cannot be accounted for by supposing that subjects simply adopted the goals of characters in the target passages and then selected key facts based on relevance to those goals. The analogies were constructed so that both key facts were equally goal-relevant: That is, each was a consequent of a causal chain. Thus, goal relevance by itself could not provide a basis for choosing one key fact over the other. A further indication that relevance alone could not have been sufficient is that target-only subjects rated the importance of the key facts as equivalent.

extrinsic goals influence mapping indirectly. For example, Gentner (1989) proposes a before/after influence of goals on the mapping process: A person's goals influence the construals (i.e., the current working-memory representations) of the base and target domains that are input to the analogical mapping process, and after an interpretation is constructed, goals influence the evaluation of this interpretation, especially the inferences derived from it.

Analogy and Discovery

The use of systematicity to constrain analogical mapping, as proposed by the structure-mapping account, allows us to understand the role of analogy in creative discovery. On this account, simply by forming systematic and consistent matches between a base and target domain, a coherent matching structure and candidate inferences can emerge. Thus, analogical thinking is not restricted to cases in which we possess prior expectations of what matches or inferences to seek. Inferences that are unanticipated may arise, and a structure may be revealed that was formerly implicit in the analogous domains. Thus, analogy can not only be a tool for instrumental reasoning, but also can be used to achieve the general human goals of explanation, and discovery, and the satisfaction of curiosity.

Implications for Ordinary Similarity

An important question is whether the matching processes observed here extend to literal (ordinary) similarity matches. We have seen that analogical mapping concerns corresponding systems of elements and not merely independent correspondences among individual elements. Our findings cannot be accounted for by models of similarity that fail to incorporate interrelations among features into their computations. If we focus on analogy as a subcase of similarity, then our findings pose significant problems for many applications of Tversky's (1977) influential contrast model of similarity. In these approaches, similarity is computed by calculating matches and mismatches among sets of independent features (e.g., Gati & Tversky, 1982; Tversky & Gati, 1982).[14] But the present results indicate people seek to match systems of interconnected features. Some recent evidence indicates that interdependence among features may also be important in judgments of perceptual similarity (Goldstone, Gentner, & Medin, 1989; Goldstone, Medin, & Gentner, in press; Markman & Gentner, 1990; Markman, Medin, & Gentner, 1990; Medin, Goldstone, & Gentner, 1990). For example, Goldstone, Gentner, and Medin (1989) found that similarity judgments among patterns of geometric stimuli violated the independence assumption in

[14] It should be noted that the independent features assumption is not crucial to Tversky's contrast model in its theoretical statements. However, empirical tests of the model have generally assumed independent features (e.g., Gati & Tversky, 1982; Tversky & Gati, 1982).

that the importance of a matching relation depended on matches in neighboring relations. Thus, in judgments of literal similarity, as well as in analogy, it may be the nonindependent (interconnected) features that matter most in the comparison.

Conclusion

What factors constrain the choice of information to map in an analogy? Although one obvious factor is the individual feature matches between two domains, local similarity alone is insufficient to determine whether a feature is mapped. Other factors are needed to determine which commonalities are important. This research provides a direct examination of selection constraints in analogy, and the first test of the specific claim that systematicity can act as a constraint. The results indicate that, given a choice between component matches or predictions that are in themselves equally good or valid, subjects prefer those matches and make those predictions that maintain a highly systematic correspondence between the two analogous domains. In deciding how to compare two situations, people select information on the basis of its connection to a larger matching structure. Thus, a preference for coherent systems of common information appears to be a psychologically real constraint on analogical mapping. These findings indicate that an adequate theory of analogy and similarity cannot focus simply on independent feature matches, but must take into account the embedding of the matches in connected systems.

REFERENCES

Anderson, J.R. (1989). A theory of the origins of human knowledge. *Artificial Intelligence, 40*, 313–351.

Black, M. (1962). *Models and metaphors*. Ithaca, NY: Cornell University Press.

Burstein, M. (1983). Concept formation by incremental analogical reasoning and debugging. In R. Michalski, J. Carbonell, & T. Mitchell (Eds.), *Machine learning* (Vol. 2). Los Altos, CA: Morgan Kaufman.

Burstein, M., & Adelson, B. (1987). Analogical learning: Mapping and integrating partial mental models. In *Program of the Ninth Annual Conference of the Cognitive Science Society*, Hillsdale, NJ: Erlbaum.

Carbonell, J.G. (1981). Invariance hierarchies in metaphor interpretation. In *Proceedings of the Third Annual Meeting of the Cognitive Science Society*, Hillsdale, NJ: Erlbaum.

Carbonell, J.G. (1983). *Metaphor and common-sense reasoning* (Tech. Rep. No. 83–110). Pittsburgh, PA: Carnegie-Mellon University, Computer Science Department.

Catrambone, R., & Holyoak, K. J. (1989). Overcoming contextual limitations on problem-solving transfer. *Journal of Experimental Psychology: Learning, Memory and Cognition, 15*, 1147–1156.

Clark, H. (1973). The Language-as-fixed-effect fallacy: A critique of language statistics in psychological research. *Journal of Verbal Learning and Verbal Behavior, 12*, 335–359.

Clark, H.H., & Chase, W.G. (1972). On the process of comparing sentences against pictures. *Cognitive Psychology, 3*, 472–513.

Clement, J. (1983). A conceptual model discussed by Galileo and used intuitively by physics students. In A. Stevens & D. Gentner (Eds.), *Mental models*, Hillsdale, NJ: Erlbaum.

Collins, A., & Burstein, M. (1989). A framework for a theory of mapping. In S. Vosniadou & A. Ortony (Eds.), *Similarity, analogy, and thought*. New York: Cambridge University Press.

Falkenhainer, B., Forbus, K.D., & Gentner, D. (1989). The structure-mapping engine: an algorithm and examples. *Artificial Intelligence, 41*, 1-63.

Forbus, K., & Gentner, D. (1989). Structural evaluation of analogies: What counts? In *Program of the Eleventh Annual Conference of the Cognitive Science Society*, Hillsdale, NJ: Erlbaum.

Forbus, K.D. & Oblinger, D. (1990). Making SME greedy and pragmatic. In *Proceedings of the Twelfth Annual Conference of the Cognitive Science Society*. Hillsdale, NJ: Erlbaum.

Gati, I., & Tversky, A. (1982). Representations of qualitative and quantitative dimensions. *Journal of Experimental Psychology: Human Perception and Performance, 8*, 325-340.

Gentner, D. (1980). *The structure of analogical models in science* (Tech. Rep. No. 4451). Cambridge, MA: Bolt Beranek and Newman, Inc.

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

Gentner, D. (1989). The mechanisms of analogical learning. In S. Vosniadou & A. Ortony (Eds.), *Similarity and analogical reasoning*. New York: Cambridge University Press.

Gentner, D., & Clement, C.A. (1988). Evidence for relational selectivity in the interpretation of analogy and metaphor. In G.H. Bower, (Ed.), *The Psychology of learning and motivation. Advances in research and theory* (Vol. 22). New York: Academic.

Gentner, D., & Gentner, D. (1983). Flowing waters or teeming crowds: Mental models of electricity. In A. Stevens & D. Gentner (Eds.), *Mental models*, Hillsdale, NJ: Erlbaum.

Gentner, D., & Jeziorski, M. (1989). Historical shifts in the use of analogy in science. In B. Gholson, W.R. Shadish, R.A. Neimeyer, & A. Houts (Eds.), *Psychology of Science: Contributions to Metascience*. New York: Cambridge University Press.

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

Gentner, D., & Schumacher, R.M. (1986). Use of structure-mapping theory for complex systems. *Proceedings of the IEEE International Conference on Systems, Man and Cybernetics*. New York: IEEE.

Gentner, D., & Toupin, C. (1986). Systematicity and surface similarity in the development of analogy. *Cognitive Science, 10*, 277-300.

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

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

Goldstone, R., Gentner, D., & Medin, D. (1989). Relations relating relations. In *Proceedings of the Eleventh Annual Meeting of the Cognitive Science Society*. Hillsdale, NJ: Erlbaum.

Goldstone, R., Medin, D., & Gentner, D. (in press). Relational similarity, and the non-independence of features in similarity judgements. *Cognitive Psychology*.

Goodman, N. (1970). Seven strictures on similarity. In L. Foster & J.W. Swanson (Eds.), *Experience and theory*. Amherst: University of Massachusetts Press.

Hall, R.P. (1989). Computational approaches to analogical reasoning: A comparative analysis. *Artificial Intelligence, 39*, 39-120.

Hesse, M.B. (1966). *Models and analogies in science*. Notre Dame, IN. University of Notre Dame Press.

Holyoak, K.J. (1984). Analogical thinking and human intelligence. In R.J. Sternberg, (Ed.), *Advances in the psychology of human intelligence* (Vol. 2). Hillsdale, NJ: Erlbaum.

Holyoak, K.J. (1985). The pragmatics of analogical transfer. In G.H. Bower (Ed.), *The Psychology of learning and motivation* (Vol. 1). New York: Academic Press.

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

Holyoak, K.J., & Thagard, P. (1989). A computational model of analogical problem solving. In S. Vosniadou & A. Ortony (Eds.), *Similarity, analogy, and thought*. New York: Cambridge University Press.

Keane, M. (1988). *Analogical problem solving*. Chichester, England: Ellis Horwood.

Kedar-Cabelli, S. (1985a). *Analogy—from a unified perspective* (Tech. Rep. No. ML-TR-3). New Brunswick, NJ: Rutgers University, Hill Center for the Mathematical Sciences, Laboratory for Computer Science Research.

Kedar-Cabelli, S. (1985b). Purpose-directed analogy. *Proceedings of the Seventh Annual Conference of the Cognitive Science Society*, Hillsdale, NJ: Erlbaum.

Kintsch, W. (1988). The role of knowledge in discourse comprehension: A Construction-integration model. *Psychological Review, 95*, 163-182.

Kintsch, W., & van Dijk, T.A. (1978). Towards a model of text comprehension and production. *Psychological Review, 85*, 363-394.

Kittay, E.F. (1987). *Metaphor: Its cognitive force and linguistic structure*. Oxford, England: Carendon Press.

Kolodner, J. (1987). Capitalizing on failure through case-based inference. *Proceedings of the Ninth Annual Conference of the Cognitive Science Society*, Hillsdale, NJ: Erlbaum.

Kolodner, J., Simpson, R.L., & Sycara, K. (1985). A process model of case based reasoning and problem solving. *Proceedings of the Sixth Annual International Joint Conference on Artificial Intelligence*. Los Altos, CA: Morgan Kaufman.

Kotovsky, L., & Gentner, D. (1990, March). *Pack light: You will go farther*. Paper presented at the Midwestern Artificial Intelligence and Cognitive Science Society Conference, Carbondale, IL.

Lakoff, G., & Johnson, M. (1980). *Metaphors we live by*. Chicago: University of Chicago Press.

Markman, A.B., & Gentner, D. (1990, March). *Robot arms go with cars: Evidence for analogical mapping in similarity*. Paper presented at the Midwestern Artificial Intelligence and Cognitive Science Society Conference, Carbondale, IL.

Markman, A., Medin, D., & Gentner, D. (1990). *An analysis of the perceptual-verbal distinction in similarity judgements*. Unpublished manuscript.

Medin, D., Goldstone, R., & Gentner, D. (1990). Similarity involving attributes and relations: Judgements of similarity and difference are not inverses. *Psychological Science, 1*, 64-69.

Meyer, B.J.F. (1984). Text dimensions and cognitive processing. In H. Mandl, N.L. Stein, & T. Trabasso (Eds.), *Learning and comprehension of text*. Hillsdale, NJ: Erlbaum.

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

Ortony, A. (1979). Beyond literal similarity. *Psychological Review, 86*, 161-180.

Ortony, A., Vondruska, R.J., Foss, M.A., & Jones,L.E. (1985). Salience, similies, and the asymmetry of similarity. *Journal of Memory and Language, 24*, 569-594.

Rattermann, M.J., & Gentner, D. (1987). Analogy and similarity: Determinants of accessibility and inferential soundness. *Proceedings of the Ninth Annual Conference of the Cognitive Science Society*. Hillsdale, NJ: Erlbaum.

Rattermann, M.J., & Gentner, D. (1990, March). *The development of similarity use: It's what you know, not how you know it*. Paper presented at the Midwestern Artificial Intelligence and Cognitive Science Society Conference, Carbondale, IL.

Reed, S.K. (1987). A structure-mapping model for word problems. *Journal of Experimental Psychology: Learning, Memory and Cognition, 13*, 124-139.

Rips, L. (1989). Similarity, typicality, and categorization. In S. Vosniadou, & A. Ortony (Eds.), *Similarity and analogical reasoning*. New York: Cambridge University Press.

Ross, B. (1984). Remindings and their effect in learning a cognitive skill. *Cognitive Psychology, 16*, 371-416.

Ross, B. (1989). Remindings in learning and instruction. In S. Vosniadou & A. Ortony (Eds.), *Similarity and analogical reasoning*. New York: Cambridge University Press.

Rumelhart, D.E., & Norman, D.A. (1981). Analogical processing in learning. In J.R. Anderson (Ed.), *Cognitive skills and their acquisition*. Hillsdale, NJ: Erlbaum.

Schank, R.C. (1982). *Dynamic memory*. New York: Cambridge University Press.

Schumacher, R.M., & Gentner, D. (1988). Transfer of training as analogical mapping. *IEEE Transactions on Systems, Man, and Cybernetics, 18*, 592-600.

Skorstad, J., Falkenhainer, B., & Gentner, D. (1987). Analogical processing: A simulation and empirical corroboration. In *Proceedings of the meeting of the American Association of Artificial Intelligence*, Los Altos, CA: Morgan Kaufmann.

Sternberg, R.J. (1977). *Intelligence, information processing, and analogical reasoning: The componential analysis of human abilities*. Hillsdale, NJ: Erlbaum.

Thagard, P., & Holyoak, K. (1988). Analogical problem solving: A constraint satisfaction approach. *Proceedings of the Tenth Annual Conference of the Cognitive Science Society*, Hillsdale, NJ: Erlbaum.

Trabasso, T., & Van Den Broek, P. (1985). Causal thinking and the representation of narrative events. *Journal of Memory and Language, 24*, 612-630.

Tversky, A. (1977). Features of similarity. *Psychological Review, 84*, 327-352.

Tversky, A., & Gati, I. (1982). Similarity, separability, and the triangle inequality. *Psychological Review, 89*, 123-154.

van Dijk, T.A. (1980). *Macrostructures*. Hillsdale, NJ: Erlbaum.

van Dijk, T.A., & Kintsch, W. (1983). *Strategies of discourse comprehension*. New York: Academic Press.