Constraints on Analogical Inference

ARTHUR B. MARKMAN
Columbia University

The ability to reason by analogy is particularly important because it permits the extension of knowledge of a target domain by virtue of its similarity to a base domain via a process of analogical inference. The general procedure for analogical inference involves copying structure from the base to the target in which missing information is generated, and substitutions are made for items for which analogical correspondences have already been found. A pure copying with substitution and generation process is too profligate to be useful, and so constraints must be placed on what information is to be carried over. In this paper, the importance of systematicity as a constraint on inference is explored in four studies in which subjects find correspondences between domains and also make inferences. This work suggests that people prefer to make inferences of information connected to systematic correspondences between domains. A second important theme of this paper is that violations of one-to-one mapping can lead to inconsistent object substitutions in inference. The data reveal no such inconsistent substitutions in people’s inferences, suggesting that they do respect one-to-one mapping in analogical inference. These findings are discussed relative to four prominent computational models of analogical mapping and inference.

INTRODUCTION

Analogy is a powerful cognitive mechanism that permits two domains to be seen as similar on the basis of connected systems of relations that hold between them (Gentner, 1983, 1989; Hesse, 1966). For example, the atom and the solar system may be seen as similar, because both involve satellites revolving around a central object. The use of analogies allows us to strip away compelling properties (e.g., the size and heat of the sun) and focus only on the common relations. If analogy stopped at the point where it found the relational commonalities of a pair of domains, it would be a valuable but limited process. However, analogies also permit inferences from one domain (often called the base) to another (often called the target). Thus, analogy permits one domain to be extended by its similarity to the
other. For example, if we know that the planets revolve around the sun because of a gravitational attraction associated with the difference in mass between them, we might infer that the electrons revolve around the nucleus because of a similar gravitational attraction. As this example demonstrates, analogical inferences can generate testable predictions about the target domain, though as in this case, the inferences may be factually incorrect.

This paper is focused on factors that determine what information is carried from base to target. This issue is important, because many sophisticated computational models of analogy have been developed that make predictions about the way people make candidate inferences. Some of these models have also made proposals about the way analogical inferences are made (e.g., Falkenhainer, Forbus, & Gentner, 1989; Holyoak & Thagard, 1989; Keane & Brayshaw, 1988; Keane, Ledgeway, & Duff, 1994). Further, there are interesting new models of analogical reasoning for which the candidate inference mechanisms are still being developed (Hummel & Holyoak, 1997, in press). Understanding the way that candidate inferences are constrained will assist in the development of these new models.

This paper focuses primarily on structural constraints on analogy. I will begin with a description of the analogical mapping process. Then, I will discuss a general way of viewing analogical inference—as copying with substitution and generation (CWSG). Next, I will focus on the constraining effects of systematicity on analogical inference (Clement & Gentner, 1991; Gentner, 1983, 1989), illustrating the role of systematicity with computer simulations using Falkenhainer, Forbus, & Gentner’s (1989) Structure-Mapping Engine (SME). In addition, I will examine the impact of many-to-one mappings (i.e., homomorphisms) on CWSG. Following this discussion, I will present four studies that examine the importance of systematicity in constraining analogical inference, and assess the impact of many-to-one mappings on analogical inference. Finally, I will discuss the implications of this work for current and developing computational models of analogy.

The Structural Alignment View of Comparison

Analogy may be thought of as one possible similarity relationship between a pair of items (Gentner, 1983, 1989). There is general agreement that understanding analogy requires a number of processes including retrieval and representation, mapping, inference, and adaptation (Gentner & Markman, 1997; Goldstone, 1994; Holyoak & Thagard, 1995; Keane, 1996; Keane, Ledgeway, & Duff, 1994; Markman & Gentner, 1993; Medin, Goldstone, & Gentner, 1993). These processes may interact, but it has been helpful to separate them for the purposes of empirical study and modeling. This paper focuses on the processes of mapping and inference. It is assumed that inferences are derived from a mapping between a base and target domains, and so the mapping process is described first.

There is general consensus among analogy researchers that mental representations contain information about properties of items, as well as the connections between elements in the representations. The representations themselves consist of entities, which are the objects in the domain (e.g., sun, nucleus), attributes, which describe properties of items [in the propositional notation commonly used by models of analogy, an attribute could be written as hot\{sun\}], relations, which relate two or more elements in a representation [e.g., revolve\{planets, sun\}], and functions, which are predicates that map onto values other than
truth values, and are often used to represent psychological dimensions (e.g., mass(sun), which would map to a value of the mass of the sun rather than a truth value). Relations may connect entities, attributes, functions or other relations. The representation of a domain at any given time is assumed to reflect the person's psychological construal of the situation. Different people, or the same person at different times, may represent a given situation in different ways.

A pair of structured representations can be compared to determine their similarity via a process of structural alignment. This process is governed by the fundamental constraints of semantic similarity and structural consistency (see Gentner, 1983, 1989; Gentner, Markman, & Medin, in preparation; Gentner & Markman, 1997; Holyoak & Thagard, 1989, for detailed discussions of these constraints). Semantic similarity requires that any correspondence between domains be rooted in a similarity of the representational elements. Not all elements need to be the same (e.g., object similarities differ in the case of analogy), but at least some of the elements must be the same. According to structure mapping theory, the similarities involve identities in representational elements (Gentner, 1983, 1989), though other theories of analogy relax this constraint (Holyoak & Thagard, 1989).

The core constraint on comparison is structural consistency, which is made up of the dual requirements of parallel connectivity and one-to-one mapping. Parallel connectivity requires that if a pair of elements in a representation are placed in correspondence, then the arguments of those elements must also be placed in correspondence. For example, if the atom and the solar system are seen as being similar because both have something revolving around something else, then the planets and the electrons must be matched, because both are revolving, and the sun and the nucleus must be matched, because both are being revolved around. One-to-one mapping requires that each element in one representation match to at most one element in the other representation.

One way that models of analogy differ is in their adherence to these constraints. For the Structure Mapping Engine (SME; Falkenhainer, Forbus, & Gentner, 1989), structural consistency is inviolable and interpretations must be rooted in matching predicates in the base and target that have identical names. Keane's IAM also enforces structural consistency without allowing violations. In contrast, Holyoak and Thagard's (1989) ACME, and Hummel and Holyoak's (1997) LISA treat the constraints on analogy as pressures that guide the formation of interpretations. The constraints of structural consistency and semantic similarity may be violated when other factors suggest inconsistent matches.

One other aspect of the analogy process that is important for the present discussion is systematicity. Systematicity suggests that, all else being equal, matches preserving connected systems of relations are preferred to matches preserving isolated matches. This preference for systematicity is rooted in the assumption that connected systems of relations are often the elements that provide a deep coherent interpretation of an analogy (see Clement & Gentner (1991), Gentner & Toupin (1986) and Markman & Gentner (1993) for evidence that systematicity influences analogy and similarity). This overview of the mapping process sets the stage for a discussion of analogical inference.
Analogical Inference

Analogical inference has not been the object of as much empirical work as many other aspects of the mapping process. There has been extensive related work on the role of analogy in problem solving (Gick & Holyoak, 1980; Novick, 1988, 1990; Novick & Holyoak, 1991; Reed, Ernst, & Banerji, 1974; Ross, 1987, 1989). When solving a problem, people may have a prior example that they can use to solve a new problem. In this case, people typically do not have an extensive representation of the target problem. The base example is used to structure the target problem so that it can be solved. In a problem-solving setting, there is pragmatic information about the task suggesting that the known solution procedure from the example is to be transferred. The present paper focuses on a situation in which there are relational correspondences between base and target, but there is no obvious pragmatic constraint on what information should be carried from base to target.

The interpretation of an analogy consists of a set of correspondences between representational elements in the base and representational elements in the target. Inferences can then be carried from the base domain to the target domain in order to extend knowledge about the target. The most general way to describe potential candidate inferences is copying with substitution (and generation), or CWSG (Holyoak, Novick & Melz, 1994; see Gentner, 1982, for a foundational discussion of this issue). Copying with substitution and generation involves taking any element in the base domain for which there is a correspondence and carrying over to the target all representational structure attached to that element. Whenever a correspondence between base and target exists for an element being inferred, that correspondence is substituted into the information being inferred. Relations in the base that are not in the target are carried over identically. Finally, new target entities can be posited when their existence is required to complete a structure from the base.

For example, Fig. 1 shows simple descriptions of departments at a base and target college. These descriptions are presented in a simplified propositional notation. Looking at just the first statements about the English and Computer Science departments, there is a correspondence in that both have obtained research grants. Thus, the following correspondences hold between the English department and the Computer Science department:

English faculty $\rightarrow$ Computer Science Faculty

grants $\rightarrow$ grants

obtain \((x,y)\) $\rightarrow$ obtain \((x,y)\)

Based on these correspondences, the CWSG algorithm would carry over the higher-order ‘cause’ relation, as well as the ‘hire’ relation from the base. The English faculty argument to the hire relation would be transformed into the Computer Science faculty by substitution, and a hypothetical new entity of ‘research assistants’ would be posited for the target domain.

This kind of candidate inference mechanism is very general, and places few constraints on what information gets carried over from the target. Any correspondence between base and target can lead to information being carried over. For example, if the base domain con-
ANALOGICAL INFERENCE

Old School

Sample Base and Target

English department

CAUSE
- (obtain Eng_faculty grants)
- (hire Eng_faculty rsch_assts)

CAUSE
- (infighting Eng_faculty)
- (avoid Eng_faculty offices)

New School

Computer Sci. Department

CAUSE
- (obtain CS_faculty grants)

CAUSE
- (hire CS_faculty rsch_assts)

CAUSE
- (infighting CS_faculty)
- (avoid CS_faculty CS_offices)

NOTE: CS_rsch_assts and CS_offices are entities created by the inference.

Figure 1. Sample base and target to illustrate the copying with substitution and generation inference process.

consists of the whole description of the English department as it appears in Fig. 1, and the target domain consists only of the fact that the Computer Science department faculty obtained grants, the CWSG candidate inference mechanism will carry over the relation that faculty in the department may fight and that this infighting causes the faculty to avoid their offices.

There is a danger in making too many candidate inferences. Most facts about a base domain are not likely to be true in the target domain. For example, it might be the case that the English faculty read a lot of Chaucer. This inference is unlikely to be true about the Computer Science department, though it is an inference that could be made solely on the basis of copying with substitution and generation. Making many untrue inferences defeats the purpose of reasoning by analogy. Candidate inferences are supposed to be an interesting source of new inferences about a domain. If most candidate inferences turn out to be false, then analogy is not being helpful.
Thus, there must be constraints on the analogical inference process that focus on information that is likely to be useful to infer. There are many ways that analogical inference could be constrained, such as focusing on pragmatically relevant facts or salient facts (Holyoak & Thagard, 1989; Keane, 1996). This paper will focus on the constraint of systematicity (Clement & Gentner, 1991), though other constraints will be considered in the general discussion. Systematicity suggests that the process of copying with substitution and generation takes place when the information to be inferred consists of relational structure that is attached to matching predicates between base and target. That is, it is not enough simply to have corresponding entities, there must also be some corresponding representational structure in order for the inference to be licensed. In the example in Fig. 1, the inference connected to the matching relation that the faculty obtains grants leads to the inference that these grants may cause the faculty to hire research assistants. However, because there is no fact connected to the second relational structure about the faculty fighting, this structure is not carried over as a potential candidate inference. In this paper, facts connected to matching predicates are called shared system facts, and facts that are not connected to matching predicates are called nonshared system facts. In this language, systematicity constrains analogical inference to shared system facts. Nonshared system facts are not carried over as inferences.

A study by Clement and Gentner (1991) provided initial evidence that shared system facts are preferred to nonshared system facts as inferences. Subjects were given pairs of stories that were quite clearly analogous. In one stimulus set, one story described a robot using its probes to take in data, and its companion story described a space organism using its underbelly to ingest minerals. The base passage contained two causal statements. The story about the organisms had two key statements.

1. The organisms exhausted the minerals in one spot causing them to stop using their underbelly; and
2. The underbelly of the organism adapts to particular rocks causing them to be unable to function on new rocks.

The target passage also contained two facts. One of these facts matched the antecedent of one of the causal statements in the base story (i.e., it formed the basis of a shared system). For example

3. The robot exhausted the data in one area.

This fact corresponds to sentence (1) in which the organism exhausted the minerals on the rock. A second fact did not match either of the causal antecedents in the base. For example,

4. The robots cannot pack their probes to allow them to survive the flight to a new planet.
This statement does not match either statement (1) or statement (2).

When subjects were presented with these story pairs and told simply to use the analogy to make any one new prediction about the target, 53% of the predictions were of the fact connected to the shared-system, as compared with 22% for facts in the base not connected to the shared system, and 18% for all other predictions. In contrast, the control subjects selected one of the two key facts only 18% of the time, making predictions based on other factors 82% of the time. Similar evidence was obtained by Spellman and Holyoak (1996), who gave subjects pairs of soap opera plots that shared quite a bit of plot line. Subjects were asked to extend one of the plots based on the other. The plot extensions nearly always preserved the relationships between corresponding characters established by the matching aspects of the plots. Taken together, these results suggest that candidate inference is constrained to carry over shared system facts from base to target.

Candidate Inference and Many-to-one Mappings

Some recent results raise an interesting question about inference. In studies by Spellman and Holyoak (1996), subjects were given base domains that could have had many possible correspondences with the target. For example, in one of their studies, the stimuli were descriptions of countries containing both economic and military information. The target country might correspond to one country on the basis of its military power, but to a second country on the basis of its economic strength. While subjects who made inferences seem to have carried over shared system facts, subjects asked to find correspondences between the countries said that the target country corresponded to more than one country in the base. Spellman and Holyoak give a number of possible explanations for this finding. One possibility they raise is that the mapping process might search for homomorphisms, that is, the mapping process might allow interpretations with many-to-one correspondences between elements in the base and target domains. A second possibility they suggest is that the mapping process might be restricted to structurally consistent mappings (or isomorphisms), and that potential many-to-one mappings are resolved by forming many different isomorphisms. Spellman and Holyoak do not explicitly endorse either hypothesis.

The possibility of many-to-one mappings is interesting, because a problem arises with the copy with substitution and generation process given a many-to-one mapping [see Gentner (1982) for a similar discussion]. The problem will be illustrated with the inconsistent condition in Fig. 2. In this figure, the English department faculty could correspond to the Computer Science department faculty on the basis of the matching relation that both obtained research grants. The English department faculty could also correspond to the Music Department faculty on the basis of both being excellent teachers. In a model of analogical mapping that permits violations of one-to-one mapping, a homomorphism could be constructed that places the faculty of the English department in correspondence with both the faculty of the Computer Science department and the faculty of the Music department.

When copying with substitution and generation is run to generate inferences in this case (even assuming that the process is constrained by systematicity), the relation that obtaining grants causes the faculty to hire research assistants will be carried over as an inference. Because the English department has correspondences in the target domain, the CWSG pro-
<table>
<thead>
<tr>
<th>Old School</th>
<th>New School</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Consistent Condition</strong></td>
<td></td>
</tr>
<tr>
<td><strong>English department</strong></td>
<td><strong>Computer Sci. Department</strong></td>
</tr>
<tr>
<td><strong>CAUSE</strong></td>
<td><strong>CAUSE</strong></td>
</tr>
<tr>
<td>(obtain Engfaculty grants)</td>
<td>(obtain CSfaculty grants)</td>
</tr>
<tr>
<td>(hire Engfaculty rsch_assts)</td>
<td></td>
</tr>
<tr>
<td><strong>CAUSE</strong></td>
<td><strong>CAUSE</strong></td>
</tr>
<tr>
<td>(infighting Engfaculty)</td>
<td>(infighting CSfaculty)</td>
</tr>
<tr>
<td>(avoid Engfaculty offices)</td>
<td></td>
</tr>
<tr>
<td><strong>Biology department</strong></td>
<td><strong>Music Department</strong></td>
</tr>
<tr>
<td><strong>CAUSE</strong></td>
<td><strong>CAUSE</strong></td>
</tr>
<tr>
<td>(excellent Biofaculty teach)</td>
<td>(excellent Musfaculty teach)</td>
</tr>
<tr>
<td>(oversubscribed classes)</td>
<td></td>
</tr>
<tr>
<td><strong>CAUSE</strong></td>
<td><strong>CAUSE</strong></td>
</tr>
<tr>
<td>(small-number Biofaculty)</td>
<td>(small-number Musfaculty)</td>
</tr>
<tr>
<td>(not (perform Bio-faculty advising))</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Inconsistent Condition</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>English department</strong></td>
<td><strong>Computer Sci. Department</strong></td>
</tr>
<tr>
<td><strong>CAUSE</strong></td>
<td><strong>CAUSE</strong></td>
</tr>
<tr>
<td>(obtain Engfaculty grants)</td>
<td>(obtain CSfaculty grants)</td>
</tr>
<tr>
<td>(hire Engfaculty rsch_assts)</td>
<td></td>
</tr>
<tr>
<td><strong>CAUSE</strong></td>
<td><strong>CAUSE</strong></td>
</tr>
<tr>
<td>(excellent Engfaculty teach)</td>
<td>(infighting CSfaculty)</td>
</tr>
<tr>
<td>(oversubscribed classes)</td>
<td></td>
</tr>
<tr>
<td><strong>Biology department</strong></td>
<td><strong>Music Department</strong></td>
</tr>
<tr>
<td><strong>CAUSE</strong></td>
<td><strong>CAUSE</strong></td>
</tr>
<tr>
<td>(infighting Biofaculty)</td>
<td>(excellent Musfaculty teach)</td>
</tr>
<tr>
<td>(avoid Biofaculty offices)</td>
<td></td>
</tr>
<tr>
<td><strong>CAUSE</strong></td>
<td><strong>CAUSE</strong></td>
</tr>
<tr>
<td>(small-number Biofaculty)</td>
<td>(small-number Musfaculty)</td>
</tr>
<tr>
<td>(not (perform Bio-faculty advising))</td>
<td></td>
</tr>
</tbody>
</table>

*Figure 2.* Illustration of the design of Experiment 1. These base and target domains were used for the SME simulation as well.
cess will make a substitution in the inference. However, in the face of a many-to-one mapping, which correspondence should be substituted? There is no principled way to make this decision, and so a many-to-one mapping makes it possible to get inconsistent substitutions within an inference. For example, given the inconsistent condition in Fig. 2, a mapping like

\[
\text{CAUSE}
\]

\[
\text{(obtain (CS\text{\_}faculty, grants),}
\]

\[
\text{hire (Mus\text{\_}faculty, rsch\text{-}assts))}
\]

is possible. Notice that in this inference, the Computer Science faculty substitutes for the English faculty in the obtain relation, but the Music faculty substitutes for the English faculty in the hire relation. This substitution is inconsistent, because in the base domain a single entity (the English faculty) is the first argument to both the ‘obtain’ and ‘hire’ relations, but two different entities are substituted for the English faculty in this inference. When Holyoak and Thagard’s (1989) ACME program is run on examples that yield interpretations with many-to-one mappings, it permits this kind of inconsistent object substitution (Holyoak, personal communication).¹

These inconsistent object substitutions are a direct result of the relaxation of the one-to-one mapping constraint. The only way to avoid them is to enforce one-to-one mapping strictly during the inference process, as can be seen in a simulation using the Structure Mapping Engine (SME; Falkenhainer, et al., 1989; Forbus, Ferguson, & Gentner, 1994; Forbus & Gentner, 1989; Forbus & Oblinger, 1990). As a brief introduction, SME can be presented with a pair of structured representations written in predicate-calculus notation. The correspondence between domains is found using a local-to-global process (see also Goldstone & Medin, 1994; Holyoak & Thagard, 1989). First, SME implements semantic similarity by proposing correspondences between all identical predicates without regard to the structure of the match. The constraints of structural consistency are then enforced by ensuring that matching predicates have matching arguments and by weeding out non one-to-one matches. At this time, correspondences between nonidentical elements may be proposed when these elements are arguments to matching predicates. Consistent bits of independent structure are gathered together into global interpretations. SME forms one or a few global interpretations, where each interpretation is ranked by a structural evaluator that prefers interpretations preserving deep connected (i.e., systematic) matches to those preserving isolated matches (Forbus & Gentner, 1989). Within these interpretations, candidate inferences are generated by carrying over information from base to target when the information is connected to a matching system of facts and is structurally consistent with the existing match between base and target. At this stage, SME may posit the existence of objects not present in the target that are required to transfer the inference from the base.

For this simulation, the base domain consists of descriptions of the Biology and English departments at a fictitious university as shown in Fig. 2. The target consists of descriptions of the Music and Computer Science departments. The base and target descriptions submitted to SME were designed to highlight the aspects of SME that are responsible for the predictions, and are much less complex than those that subjects are likely to have generated
(see Ohlsson (1996) for a similar use of production system models). As shown in Fig. 2, the descriptions of the departments in the base school consist of causal statements (e.g., The high level of teaching quality of the Biology faculty causes their courses to be oversubscribed). The descriptions of the Music and Computer Science departments in the target school each consist of two statements that are not connected to causal relations. In the simulation of the consistent mapping condition, both statements about the Computer Science department correspond to statements about the English department, and both statements about the Music department correspond to statements about the Biology department. In the simulation of the inconsistent condition, one statement about the Computer Science department corresponds to a statement about the English department, and the other corresponds to a statement about the Biology department. Likewise, one statement about the Music department corresponds to a statement about the English department and the other corresponds to a statement about the Biology department.

For this simulation, SME was run in its literal similarity mode using the greedy-merge algorithm that generates a small number of interpretations (Forbus & Oblinger, 1990). For the simulation of the consistent condition, a single interpretation (GMAP) is constructed. It maps the English department in the base to the Computer Science department in the target and the Biology department in the base to the Music department in the target. This interpretation also carries over four candidate inferences corresponding to the four causal statements in the base domain. For example, a causal inference is made that the Computer Science faculty will hire many research assistants based on the shared system fact that the Computer Science faculty obtain many research grants. Similarly, the other analogical inferences are attached to their matching causal antecedents.

For the simulation of the inconsistent condition, a different pattern emerges. Here, two interpretations are generated. One maps the English department in the base to the Computer Science department in the target and the Biology department in the base to the Music department in the target. This interpretation proposes only two candidate inferences, one of the causal statement connected to the fact that both the English and Computer Science departments obtain grants and the other attached to the fact that both the Biology and Music departments are small. In the inferences, the Computer Science department always substitutes for the English faculty, and the Music faculty always substitutes for the Biology faculty. A second interpretation maps the English department to the Music department and the Biology department to the Computer Science department. This interpretation also contains two candidate inferences. One inference is of the causal statement connected to the matching fact that both Biology and Computer Science faculty fight with each other. The other inference is of the causal statement connected to the matching fact that both the English and Music departments have excellent teachers. Again, consistent substitutions are made in this inference. In the inconsistent condition, finding more than one correspondence for a given department in the base requires looking at the correspondences involving that department across interpretations. Likewise, finding more than one inference for each shared system fact in this case requires collecting inferences from more than one interpretation.
To summarize the simulation results, for the consistent correspondence only one mapping for each department is found. Candidate inferences for all matching causal antecedents are generated. For the inconsistent condition, multiple interpretations are constructed, which can be used in concert to generate multiple correspondences between departments and to generate more than one candidate inference. This simulation provides two main demonstrations: (1) multiple correspondences for an object can be generated by integrating information across structurally consistent interpretations (see also Markman & Gentner, 1993), and (2) enforcing one-to-one mapping ensures that consistent substitutions are made in inference in the face of potential many-to-one matches by segregating inconsistent correspondences into different interpretations.

An Overview of the Studies

There are two key aspects of analogical inference of interest in the present studies. First, these studies will focus on broadening the support for the role of systematicity as a constraint on analogical inference. Second, these studies will examine the influence of many-to-one mappings on analogical inference to assess the strength of the one-to-one mapping constraint in inference. In these studies, the subjects play the role of a college dean who is moving from an old school (Fallsburg University) to a new school (Gordmont University). The departments in the old school will be described by pairs of causal statements in which an antecedent leads to some consequent. The departments in the new school consist of pairs of facts that contain only causal antecedents from the previous school. Subjects will provide correspondences between departments in the new and old schools, and will also make predictions (i.e., inferences) about what might happen given facts about departments at the new school.

The role of systematicity as a constraint on inference will be assessed by having facts in the target domain that match some causal antecedents in the base but not others so that there will be both shared system facts and nonshared system facts. The influence of many-to-one mappings on inference will be assessed in two ways. In Experiment 1, potential many-to-one mappings will be created by having departments in the old school share some facts with one department in the new school and other facts with a second department in the new school. In the remaining experiments, potential many-to-one mappings will be created by having facts about the new school that appear in descriptions about more than one department in old school.

Subjects will first be asked which departments in the old school correspond to each department in the new school in order to demonstrate that they are sensitive to the presence of many potential mappings for each department. Subjects should be able to report more than one correspondence for each department in the new school. This result alone has two possible explanations. One possibility is that the mapping process allows violations of one-to-one mapping, and people are reporting a single interpretation of the match between base and target. A second possibility is that the mapping process forms more than one structurally consistent match between base and target, and people are reporting correspondences from more than one interpretation. Violations of one-to-one mapping can be detected by looking at people's inferences. If the mapping process permits violations of one-to-one mapping, then some inferences should contain inconsistent object substitutions.
EXPERIMENT 1

The first experiment addresses the two key aspects of analogical inference using an experimental setup like the one presented in the computer simulation above. As in the simulation, the departments in the new school are either mapped consistently or inconsistently to the departments in the old school. In the consistent condition, both facts about the new school match the causal antecedents of a single department in the old school. In the inconsistent condition, each department in the new school matches two departments in the old school. In this case, one fact about each department in the new school matches a causal antecedent from one department in the old school, and the second fact matches to a fact in a different department in the old school. The base and target descriptions for both conditions are given in Appendix A.

After reading the descriptions of the departments in the base school (and taking a short quiz), subjects read the descriptions of the departments in the new school. Then, they are given a correspondence test and an inference test. Subjects are permitted to look back at the descriptions of the old school during both tests. In the correspondence test, subjects are asked which department or departments in the old school correspond to each department in the new school. Following Spellman and Holyoak (1996), the instructions of this task emphasize that many departments in the old school may correspond to each department in the new school, and many lines are given for subjects to write their answer. If subjects are able to integrate their responses across multiple interpretations (or if they are constructing a single interpretation that violates one-to-one mapping), then subjects shown the inconsistent mappings should find more correspondences for each school than should subjects shown the consistent mappings.

After the correspondence task, subjects are given an inference task. They are told that the dean must use her past experience to determine what might happen at the new school in an effort to avoid any potential problems or to nurture any potentially beneficial occurrences. This task allows the examination of two aspects of analogical inference. First, if systematicity is a strong constraint on analogical inference, then the inferences should include many more shared system facts than nonshared system facts. Second, if one-to-one mapping is a strong constraint on analogical inference, then the inferences should have consistent argument substitutions. In contrast, if one-to-one mapping is not enforced, then there should be cases in the inconsistent condition in which subjects give inconsistent substitutions within the same inference.

Method

Subjects

Subjects in this study were 32 subjects (16/condition) drawn from the Columbia University community. Subjects were paid $6.00 for their participation.
Materials

The primary experimental materials consisted of descriptions of academic departments at two Universities. The old school consisted of three departments, each described by two paragraphs (so that six total paragraphs were written). Each department appeared on a separate page. Each paragraph contained some filler information and also described a causal relationship between an antecedent and a consequent fact. Two versions of the departments in the old school were written that paired the six statements in different ways. A six question quiz was also developed. Each question on the quiz corresponded to one of the causal statements. The quiz either required the participant to identify the department containing a particular causal statement (2 questions), required the participant to identify the causal antecedent given the department containing it and the consequent (2 questions) or to identify the causal consequent given the department containing it and the antecedent (2 questions). The quiz was included to ensure that participants were reading the descriptions about the old school.

The descriptions of the two departments at the new school each consisted of one paragraph, each on its own page. Each paragraph contained some filler information as well as two facts. These facts corresponded to antecedents in the old school. In the consistent condition, the facts in the new school were both antecedents from one department in the old school. In the inconsistent condition, the facts in the new school came from different departments in the old school. Appendix A contains the descriptions from this study. The structure of these conditions is illustrated in Fig. 2, though the actual descriptions in Appendix A are somewhat different than what is shown in this Figure.

The correspondence test consisted of the name of a department from the new school followed by three lines on which to write responses. The inference task consisted of causal antecedents from the new school (along with the department they came with) followed by a page with lines on it for participants to write their response.

The order of pages describing departments at the old school was determined randomly for each subject. Likewise, the ordering of descriptions of the departments in the new school were determined randomly, as were the orderings of the questions in the correspondence and inference tests.

Procedure

Subjects were told they would be reading about a college dean moving from an old school to a new school. They first read descriptions of three departments. Then, they took a six question quiz about the descriptions. Next, they were asked to read descriptions of departments at the new school. After reading these descriptions, participants did the correspondence and inference tests (in that order).

In the correspondence test, participants were asked to say which department or departments in the old school were like each department in the new school. It was stressed that a department in the new school may not correspond to any departments at the old school, but it may also correspond to many departments. Three lines were given for their responses to encourage subjects to write down more than one response.
In the inference test, participants were told that it is often helpful to use past experience as a guide to future events. They were shown particular facts about departments at the new school and were asked to predict what might happen on the basis of the dean’s experience at the old school. A whole page was given on which the responses could be written. For both the inference and correspondence tests, participants were told that they could look back at the descriptions of departments in both the old and new schools. Participants were given as long as they required to finish the experiment. The study took about 30 minutes to complete.

**Design**

The primary independent variable was consistency of mapping, which had two between-subjects levels (Consistent/Inconsistent). For the correspondence test, the number of mappings made for each department was analyzed. For the inference test, inferences were scored as shared-system facts, non-shared system facts and other.

**Results**

**Quiz**

Subjects in the consistent ($m = 0.84$) and the inconsistent condition ($m = 0.80$) did not differ significantly in their performance on the quiz, $t(30) = 0.99, p > .10$.

**Correspondence Test**

In the correspondence test, the number of mappings made for each department was counted. It was expected that more mappings would be made on average in the inconsistent condition than in the consistent condition. As predicted, subjects in the inconsistent condition made more mappings ($m = 1.41$/department) than did subjects in the consistent condition ($m = 1.06$). $t(30) = 3.34, p < .01$. Also as expected, the number of mappings made in the inconsistent condition was reliably greater than 1, $t(15) = 4.32, s.d. = 0.38, p < .001$, but the number of mappings made in the consistent condition was not $t(15) = 1.47, s.d. = 0.17, p > .10$.

These data can be examined in more detail. In the consistent condition, 15/16 subjects said that the Music department in the new school corresponded to the Biology department in the old school. This correspondence is supported by both key facts in the Music department. Likewise, 16/16 subjects said that the Computer Science department in the new school mapped to the English department in the old school, as supported by both key facts in this department.

In the inconsistent condition, the Music department shares key facts with both the Biology department and the English department. Consistent with this setup, 9/16 subjects placed the Music department in correspondence with the Biology department, and 7/16 participants placed the Music department in correspondence with the English department, although only 2/16 subjects placed the Music department in correspondence with both the Biology and English departments. Finally, 5/16 subjects placed the Music department in
correspondence with the Political Science department. The source of this mapping is not clear. The Computer Science department shared key facts with both the Political Science and English departments. In a pattern similar to that obtained for the Music department, 9/16 participants placed the Computer Science department in correspondence with the Political Science department, 14/16 participants placed this department in correspondence with the English department, and 5/16 participants placed the Computer Science department in correspondence with both the English and Political Science departments. Thus, the participants in the inconsistent condition were sensitive to the fact that departments in the new school could potentially correspond to many departments in the old school.

Inference Test

Each statement made by participants in the inference test was scored as an explanation, a course of action, an inference of a shared-system fact, an inference of a non-shared system fact or as some other inference. Explanations were discussions of why a particular key fact was true for that department. For example, saying “The faculty don’t get along because they are selfish” is an explanation. Courses of action were suggestions by the subject for how to resolve some perceived problem. For example, saying “The dean should bring the faculty together to work out their differences” is a course of action. Neither explanations nor courses of action were listed frequently, because subjects were explicitly asked to find potential outcomes rather than solutions. These two types of statements will not be discussed further. Statements were scored as causal inferences when they were events that would be caused by the key fact. Shared-system causal inferences were those facts causally connected to the key fact in the old school. Given the descriptions in Fig. 2, saying “If the faculty don’t get along they will avoid their offices” is a shared system fact. Nonshared system causal inferences were inferences of other facts mentioned in the description of the old school, but not causally connected to the key fact. For example, saying “If the faculty don’t get along, maybe they will not do administrative work” is a nonshared system fact. All other causal inferences were scored as ‘other’. These inferences were statements about what might follow from the antecedent that were not part of the description of any department in the base. Saying “If the faculty don’t get along, they will make it difficult to hire new faculty” is an other inference. Inferences were examined to see whether the argument substitutions were consistent or inconsistent within the inference. An inference was scored as inconsistent if two or more different substitutions were made for a given item in an inference.

The data were scored by a rater who did not know the hypothesis under study. For the first experiment, a second coder (the author) also scored the data. There were no disagreements in the scoring, as the inferences were straightforward to code. In subsequent studies, only a rater who did not know the experimental hypotheses was employed.

The mean number of causal inferences of shared-system, nonshared-system and other facts listed for each inference are shown in Table 1. These data are analyzed in a 2 (Condition) x 3 (Inference Type) mixed model ANOVA, where inference type was treated as a within-subjects factor. This analysis revealed a significant main effect of Inference Type, $F(2,60) = 6.28$, $p < .01$. Neither the main effect of Condition or the interaction of these fac-
TABLE 1

Causal Inferences Made in Experiment 1

<table>
<thead>
<tr>
<th>Condition</th>
<th>Number of Shared System Inferences</th>
<th>Number of Nonshared System Inferences</th>
<th>Number of Other Inferences</th>
</tr>
</thead>
<tbody>
<tr>
<td>Consistent</td>
<td>0.45</td>
<td>0.24</td>
<td>0.52</td>
</tr>
<tr>
<td>Inconsistent</td>
<td>0.58</td>
<td>0.20</td>
<td>0.61</td>
</tr>
</tbody>
</table>

tors was significant, both $F < 1$. Post-hoc tests revealed that, as predicted, significantly more shared-system inferences were made ($m = 0.52$) than nonshared inferences ($m = 0.22$), $t(31) = 3.97$, $p < .01$. Unexpectedly, more other inferences were also made ($m = 0.56$) than nonshared inferences, $t(31) = 3.31$, $p < .01$. The difference between the number of shared system and other inferences made was not reliable. In addition, as a search for potential many-to-one mappings in inference, the data were examined for inconsistent object substitutions in inference. None were found (and indeed, none were found in any of the studies in this paper).

The means for the number of nonshared system inferences are a bit misleading. For three of the four inferences, a total of 11 nonshared system inferences were made across all 32 subjects. However, when subjects were asked to make an inference based on the small faculty size, the same number of shared system inferences and nonshared system inferences were made (17 for each). As shown in Fig. 2, the causal consequent for small faculty size was that the department was unable to carry out advising. The other causal system in the consistent condition in Biology department from the old school was that excellent teaching ability led to oversubscription of courses. Many subjects appeared to have seen small faculty size as a second causal antecedent to the oversubscription of courses. One piece of evidence for this interpretation is that 11/16 subjects in the consistent condition gave the incorrect response that small faculty size caused the oversubscription of courses in the quiz that followed the initial description of the departments from the old school.

Finally, the relationship between the correspondence task and the inference task was examined by looking at the inferences as a function of whether the subject had mapped together the base and target departments in the correspondence task. In the consistent condition, all but 2 shared system inferences and 2 nonshared system inferences came from departments that the subjects had previously mapped in the correspondence task. This finding is not surprising, as there is really only one good correspondence in the old school for each department in the new school in the consistent condition. The relationship between the correspondence and inference tasks is weaker in the inconsistent condition. Only 23/50 (46%) of the inferences made in this condition came from departments mapped together in the correspondence task. This total reflects that 20/37 (54%) of the shared system inferences came from departments mapped in the correspondence task, and only 3/13 (23%) of the nonshared system inferences came from departments mapped in the correspondence task.
Discussion

These data demonstrate a dissociation between the ability to see which items correspond, and the ability to make inferences from a base domain to a target. People in the inconsistent condition were often able to see that a department in the new school corresponded with more than one department in the old school when multiple mappings were available. In contrast, people’s inferences were specific to the particular causal consequent that arose from the matching causal antecedent. This focus on shared system facts is striking, because there was one shared system fact on each trial, but there were five nonshared system facts. Thus, if subjects were randomly selecting facts in the base to carry over, there should have been five times as many nonshared system inferences as there were shared system inferences.

There was also no evidence that subjects were creating interpretations of the correspondence between schools that violated one-to-one mapping. In particular, despite their ability to see that many departments in the base could correspond to each department in the target in the inconsistent condition, people’s inferences all contained consistent object substitutions. Thus, this study provides evidence that subjects’ inferences respected one-to-one mapping.

The observed pattern of data is consistent with the SME simulation presented before the experiment. In the simulations of both the consistent and inconsistent conditions, the inferences carried from base to target were pieces of information connected to the matching causal antecedent. In the simulation of the consistent condition, the correspondence between departments and the shared system inferences could be derived from a single structurally consistent interpretation. In contrast, in the simulation of the inconsistent condition, the two correspondences for each department and the shared system inferences required accumulating information across structurally consistent interpretations.

Subjects’ responses in the inference task were not entirely consistent with their responses in the correspondence task. In particular, in the inconsistent condition, subjects did not always map both departments with a matching causal antecedent to each department in the new school. Despite this deficiency in the correspondence task, subjects focused primarily on the facts connected to the matching causal antecedent in the inference task. As a result of this focus, subjects in the inconsistent condition often made inferences of information that came from departments the subjects had not mapped together in the correspondence task.

While these data are generally consistent with the predictions of structure-mapping, there are two points that require additional attention. First, as discussed above, one of the causal antecedents gave rise to inferences of both the shared system fact, as well as a non-shared system fact that was consistent with the causal antecedent. Subjects’ responses to the quiz suggest that they may have represented this causal chain incorrectly. However, it is also possible that subjects integrated two potential candidate inferences into a single causal story. This hypothesis will be addressed in Experiment 3. First, however, a study is presented that tests the main hypotheses in a slightly different way.
Experiment 1 examined a many-to-one mapping in which departments (i.e., entities) in the target had multiple correspondences with departments in the base. These correspondences were formed on the basis of different pieces of information about the target departments. In the inconsistent condition, focusing on one causal chain would lead to focusing on one correspondence, and focusing on the second causal chain would lead to focusing on the second correspondence. For example, in Fig. 2, the Music department could correspond to the Biology department on the basis of one fact and to the English department on the basis of a second fact. Thus, people may have been able to resolve the many-to-one mapping between departments by attending to different relations. This selective attention may break down when the same causal antecedent has a many to one mapping from base to target. In this case, even if people attend selectively to shared system facts, there will still be a potential many-to-one mapping that could manifest itself through inconsistencies in object substitutions in inference. This issue is addressed in Experiment 2.

**EXPERIMENT 2**

For this study, many-to-one mappings were made at the level of causal antecedents, which led to potential many-to-one mappings at the level of entities (i.e., departments). As illustrated in Fig. 3, each causal antecedent in the target could match with two possible causal antecedents in the base. Each of these causal antecedents was in a different department, so that the entities were also mapped many-to-one in this study. Finally, the causal chains were made longer in order to provide more possible information that could be carried from base to target. As a concrete example of this design, the Music Department at the new school had excellent teachers and a faculty that was active in social causes. In the old school, both the Political Science and English departments had excellent teachers. Further, the English and Anthropology departments had faculty that were active in social causes. One set of stories used in this study is presented in Appendix B.

This design allows three key questions to be addressed. First, the degree to which analogical inferences are specific to particular matching causal antecedents can be examined. Subjects are expected to make inferences of shared system facts more often than of non-shared system facts. Second, because there are potential many-to-one mappings between the departments, it is possible that inconsistent substitutions will be made in inferences. The data will be examined for this kind of evidence of many-to-one mappings in inference. Third, this study allows the impact of other information to be assessed. As shown in Fig. 3, the Music department shares the matching causal antecedent of having good teachers with both the Political Science and English departments. The other causal antecedent in this department, that the faculty are socially active, is also shared with the English department. One possible strategy that people may use to resolve many-to-one mappings is to use other facts to suggest the best set of entity correspondences, and to use only that interpretation (see also Goldstone & Medin, 1994). This strategy can be illustrated with an example. In Fig. 3, for an inference based on having good teachers there should be more inferences of the shared system facts in the English department than of the shared system facts in the
<table>
<thead>
<tr>
<th>Fallsburg University</th>
<th>Gordmont University</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Political Science Dept.</strong>&lt;br&gt;Excellent teaching CAUSES&lt;br&gt;Many students to sign up CAUSES&lt;br&gt;Oversubscribed classes.&lt;br&gt;Many grants CAUSES&lt;br&gt;Hiring of Assistants CAUSES&lt;br&gt;Good research and good teaching.</td>
<td></td>
</tr>
<tr>
<td><strong>English Department</strong>&lt;br&gt;Excellent teaching CAUSES&lt;br&gt;<em>Faculty to become smug</em> CAUSES&lt;br&gt;<em>Resentment by other depts.</em>&lt;br&gt;Active in social causes CAUSES&lt;br&gt;<em>Testifying before Congress</em> CAUSES&lt;br&gt;<em>Faculty become famous.</em></td>
<td></td>
</tr>
<tr>
<td><strong>Anthropology Dept.</strong>&lt;br&gt;Active in social causes CAUSES&lt;br&gt;<em>Organization of fund-drives</em> CAUSES&lt;br&gt;<em>Strong sense of community</em>&lt;br&gt;Poor faculty relations CAUSES&lt;br&gt;<em>Faculty to get job offers</em> CAUSES&lt;br&gt;<em>Faculty to leave school</em></td>
<td></td>
</tr>
<tr>
<td><strong>Music Department</strong>&lt;br&gt;Excellent Teaching&lt;br&gt;Active in Social Causes</td>
<td></td>
</tr>
</tbody>
</table>

Figure 3. Illustration of design of Experiment 2.
Political Science department, because the English department shares its other causal antecedent with the Music department as well. An alternative strategy would be to construct more than one structurally consistent interpretation (as subjects seem to have done in Experiment 1). In this case, the inferences would be based on shared system facts regardless of other matches between base and target.

Method

Subjects

Subjects in this experiment were 64 members of the Columbia University community, who were paid $6.00 for their participation. The data from four participants was dropped because they failed to respond to at least one of the inference questions. This left 30 participants for each stimulus set.

Materials

The materials for this study were similar to those used in Experiment 1. Again, the base and target consisted of descriptions of departments at fictitious colleges. The description of each department at the old school consisted of two paragraphs. Each paragraph described a causal antecedent and a causal chain that followed from that antecedent. The causal chain consisted of two additional facts. For example, faculty in a department might be excellent teachers, leading them to look down on faculty from other departments, leading to resentment. In all, four departments were described. Each causal antecedent was used twice across the set of departments. The two causal chains arising from each causal antecedent were made as dissimilar as possible in order to avoid having them confused.

The target departments were each described by one paragraph. These paragraphs contained some filler information and two of the causal antecedents that appeared in the base descriptions. The correspondence between departments in the base and target schools was constructed so that each department in the target had one department in the base that contained both of its causal antecedents and two departments that contained only one of its causal antecedents. Thus, for each department in the target, there were three departments in the base for which it had at least one matching causal antecedent. The materials from one stimulus set are presented in Appendix B.

In order to ensure that the results obtained are not a function of the salience of the particular causal consequents rather than the structure of the match, two versions of the base departments were written. For half of the participants a causal chain with a particular antecedent appeared in a department that also shared the other antecedent with the target department. For half of the participants that chain appeared in a department that did not share the other antecedent with the target department. As for Experiment 1, order of pages describing the departments and order of the questions within each task was randomized for each participant. Finally, the quiz that followed the description of the base school had eight questions, one corresponding to each causal chain.
Procedure

The procedure for this study was the same as that for Experiment 1. Once again, participants performed a quiz after reading the descriptions of the departments in the base school. Then, after reading the description of the target school, they did both the correspondence and inference tasks (in that order).

Design

The primary dependent variables were number of mappings in the correspondence task, and inferences. As for Experiment 1, inferences were scored as shared system facts, non-shared system facts and other inferences. There were two stimulus sets.

Results

Quiz

The pattern of all dependent variables was the same for both stimulus sets, so the discussion of the results is collapsed across this factor. Participants performed well on the quiz ($m = 0.87$).

Correspondence Test

Each department in the target corresponded to three departments in the base. Participants were sensitive to these many mappings. On average, participants mapped the Music Department to 1.57 departments in the base and the Computer Science department to 1.50 departments in the base. Both of these averages are reliably greater than 1.0 ($t(59) = 7.36, p < .001$, for the Music dept., $t(59) = 5.75, p < .001$ for the Computer Science department), suggesting that, on average, subjects were able to see multiple correspondences for each target department.

Looking more carefully at the correspondences, for the Music Department, the English department shared both causal antecedents, the Anthropology and Political Science departments shared one causal antecedent and the Biology department shared no causal antecedents. Consistent with this stimulus construction, 47 people made correspondences between Music and English, 25 between Music and Anthropology, 17 between Music and Political Science and only 2 between Music and Biology. Overall, 29 subjects placed the Music department in correspondence with at least two departments in the base.

For the Computer Science department, the Biology department shared both causal antecedents, the Political Science and Anthropology departments again shared one causal antecedent with Computer Science and the English department shared no causal antecedents with Computer Science. Mirroring the pattern for the Music department, 44 people made the correspondence between Computer Science and Biology, 21 between Computer Science and Political Science, 17 between Computer Science and Anthropology and only 2 between Computer Science and English. Overall, 23 subjects placed the Computer Science department with at least two departments from the base.
Inference Test

Each inference was scored as a shared-system fact, a nonshared system fact or other. For each causal antecedent, there were as many as four shared-system facts that could be inferred, two from each causal chain emanating from a causal antecedent. A one-way ANOVA revealed significant differences in the number of facts of each type inferred, $F(2,116) = 30.49$, $p < .001$. Post-hoc tests revealed that, as expected, significantly more shared-system facts were inferred on each trial ($m = 1.26$) than nonshared system facts ($m = 0.30$), $t(59) = 7.32$, $p < .001$. More shared system facts were also inferred than other facts ($m = 0.58$), $t(59) = 4.35$, $p < .001$. In addition, more other facts were inferred than nonshared system facts, $t(59) = 3.37$, $p < .005$. Finally, the data revealed no cases in which the inferences contained inconsistent object substitutions, suggesting that one-to-one mapping was not being violated.

The shared system inferences can be examined to see whether they come from the department in the base that shared both causal antecedents with the target department, or from the base department that shared only one causal antecedent with the target department. Approximately the same number of shared system inferences came from the department sharing both causal antecedents ($m = 0.64$) as came from the department sharing only one causal antecedent ($m = 0.62$).

As for Experiment 1, subjects’ inference performance can be examined as a function of their performance in the correspondence task. In this study, the correspondence test indicates that people were more likely to find a correspondence between a department in the base and target if they shared two matching facts than if they shared one. To see whether the inference data parallel the correspondence test data, the shared system and nonshared system inferences were divided into groups depending on whether the inference came from departments the subject had mapped in the correspondence task. Table 2 shows the mean number of inferences for each trial as a function of whether the inferences were shared system or nonshared system facts, and whether they came from departments the subject had placed in correspondence in the earlier correspondence task. A 2 (Inference Type) x 2 (Mapped vs. Unmapped departments) ANOVA on these data reveal two significant main effects qualified by a reliable interaction, $F(1,58) = 16.01$, $p < .01$. The interaction is of central importance here. Tests of simple effects reveal that more shared system inferences came from departments that the subject mapped together in the correspondence test ($m = 0.77$) than from departments that the subject did not map together ($m = 0.50$), $t(59) = 2.87$, $p < .05$. No reliable difference was found in the number of nonshared system inferences from departments the subject mapped together in the correspondence test ($m = 0.15$) and from the departments the subject did not map together ($m = 0.14$), $t(59) = 0.36$, $p > .10$. This result demonstrates that there is some relationship between the departments placed in correspondence and those from which inferences were drawn. Further, this influence is related to shared system facts, which are based in a correspondence between departments. Nonetheless, a substantial number of shared system inferences were also drawn from departments that were not placed in correspondence during the mapping task.
Another interesting question that can be examined is whether the inferences arising from a common causal antecedent are kept distinct. In SME simulations, each causal chain that is inferred is posited as a separate inference. However, it is possible that when people make inferences about facts from different causal consequents coming from a particular antecedent that they mix these inferences together. No previous research has examined subjects’ performance in this situation before.

In order to determine whether the inferences were kept distinct, the cases in which subjects inferred at least one piece of information from both causal chains emanating from a single causal antecedent were isolated. There were 50 such cases. Then, a straightforward criterion was adopted. If the inferences were listed in the same sentence, and were not separated by a connective like ‘but’ or ‘however’, they were counted as mixed. If the inferences were listed in separate sentences, but one inference was described as following causally from the other, then it was counted as mixed. This scoring was done by a rater not familiar with the hypothesis under study, and was checked by the author. There was one disagreement, which was resolved conservatively by calling it a mixed inference. In total, the causal chains were mixed only 3 times (6% of the time). In addition, in one case, the inference was mixed with a nonshared system fact. In the remaining 46/50 (92%) of the inferences, the causal chains were kept distinct, suggesting that people were keeping them separate.

**Discussion**

The results of Experiment 2 further support the two main hypotheses. First, many more shared-system facts were listed than nonshared system facts. In this study, shared-system inferences also outnumbered other inferences, perhaps because the longer causal chains were more salient than were the shorter causal chains used in Experiment 1. Second, no inconsistent object substitutions were found in the inferences, despite the presence of potential many-to-one mappings, suggesting that one-to-one mapping base being maintained.

One unexpected finding was that the other facts about a department affected the correspondence task more strongly than the inference task. In the correspondence task, the most frequent mappings between departments were between departments that shared both causal antecedents. Mappings between departments that shared only one causal antecedent were about half as frequent as mappings between departments that shared both causal anteced-
ents. In contrast, approximately the same number of inferences were drawn from the base department that shared both causal antecedents with the target as from the base department that shared only one causal antecedent with the target. As a further dissociation of the mapping task and the inference task, subjects were more likely to draw their inferences from the departments that they had mapped together in the correspondence task, but there were still many shared system inferences drawn from departments they had not mapped together. These data support the idea that inference focuses on facts connected to matching information between base and target regardless of other matches between base and target. However, it is possible that an even stronger preference for inferences from departments sharing both causal antecedents would be observed if subjects were not able to look back at the base descriptions during the inference task.

Subjects in this study did not mix together information from separate causal chains, even when they made inferences of information from more than one chain. This finding is consistent with the general process of copying with substitution and generation. Nonetheless, there is no prior evidence that subjects keep candidate inferences separated. The importance of this finding will be explored in more detail in the general discussion.

The result that subjects kept their causal chains separated contrasts with the finding in Experiment 1 that subjects often said that small faculty size would lead courses to be over-subscribed, even though this fact had a different causal antecedent in the old school (excellent teaching ability). Subjects' performance on the quiz in that study suggested that they may have encoded the causal chain incorrectly. However, it is also possible that when two facts are connected to the same entity in the base and they are mutually causally consistent that they may be combined into a single inference that reflects some deeper underlying cause. On the surface, this explanation would appear to be contradicted by the finding that subjects in Experiment 2 rarely mixed their inferences. However, in this study, the two causal chains connected to a common causal antecedent were designed to be as dissimilar as possible, which probably reduced the likelihood that people would mix the inferences together.

In Experiment 3, two factors were examined that might promote mixing of causal chains. In this study, the materials once again consisted of descriptions of departments in a new and old school in which facts about the new school were mapped many-to-one to causal antecedents in the base. In this study, however, the two causal chains emanating from a single causal antecedent either came from the same school or from different schools. It may be easier for people to envision combining a causal chain when it appears in the description of a single department in the base. In addition, the consistency of the chains was manipulated by varying the valence of the potential outcomes in the consequent. The two consequents emanating from a causal antecedent either had the same valence (i.e., both were positive outcomes or both were negative outcomes) or they had opposite in valence (i.e., one outcome was positive and the other was negative). It may be easier to mix two causal chains when they have the same valence than when they have opposite valence.

The two central predictions for this study follow the predictions of the first two experiments. If systematicity is a strong constraint on inference, then subjects should make more shared-system inferences than nonshared system inferences. If one-to-one mapping is respected in the inference task, then subjects should not make inconsistent substitutions.
within an inference, even when there are potential many-to-one mappings. Finally, SME simulations suggest that causal chains are kept separate. However, when the causal chains from the same matching causal antecedent come from the same department or have the same valence, it may be easier to mix the causal chains. Thus, these manipulations provide conditions favorable to the detection of mixed causal chains.

EXPERIMENT 3

Method

Subjects
Participants were 48 members of the Columbia University community (12/condition) who were paid $6.00 for their participation. One subject was dropped from the study due to experimenter error.

Design
The main independent variables in this study were Valence of the causal chains (Same/Opposite) and Department containing the same causal antecedent (Same/Different). Both of these factors were run between subjects. Subjects in this study only participated in the Inference task.

Materials
The materials for Experiment 3 were similar to those of Experiment 2. Six causal chains were constructed, in which each antecedent was followed by a chain of two results. Four of these causal chains were central to the experimental manipulation. Half of these causal chains were judged by the experimenter to be of positive valence and half were judged to be of negative valence. There were three base departments, each described by two causal antecedents and causal chains. In the same valence condition the two causal chains with a common antecedent were of the same valence (either both positive or both negative). In the different valence condition, the two causal chains with a common antecedent were of opposite valence. In addition, the two paragraphs describing causal chains associated with a particular antecedent were placed in the same department for half of the subjects and in different departments for half of the subjects. These conditions are illustrated in Fig. 4.5

Once again, there was a six-question quiz that followed the descriptions of the three base departments. Following the quiz were descriptions of two departments in the new school. These departments each had one causal antecedent that matched a causal antecedent that appeared in the base school.

Procedure
The procedure for this study was the same as that for the previous experiments, except that subjects did not participate in the correspondence task.
**Target Domain**

<table>
<thead>
<tr>
<th>Gordmont University</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Music Department</strong></td>
</tr>
<tr>
<td>A</td>
</tr>
<tr>
<td><strong>Computer Science Dept.</strong></td>
</tr>
<tr>
<td>M</td>
</tr>
</tbody>
</table>

**Base Domains**

*Same Department, Same Valence*

<table>
<thead>
<tr>
<th>Fallsburg University</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>English Department</strong></td>
</tr>
<tr>
<td>A -&gt; B -&gt; C</td>
</tr>
<tr>
<td>A -&gt; D -&gt; E</td>
</tr>
<tr>
<td><strong>Political Science Department</strong></td>
</tr>
<tr>
<td>M -&gt; N -&gt; O</td>
</tr>
<tr>
<td>M -&gt; P -&gt; Q</td>
</tr>
</tbody>
</table>

*Same Department, Opposite Valence*

<table>
<thead>
<tr>
<th>Fallsburg University</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>English Department</strong></td>
</tr>
<tr>
<td>A -&gt; B -&gt; C</td>
</tr>
<tr>
<td>A -&gt; P -&gt; Q</td>
</tr>
<tr>
<td><strong>Political Science Department</strong></td>
</tr>
<tr>
<td>M -&gt; N -&gt; O</td>
</tr>
<tr>
<td>M -&gt; D -&gt; E</td>
</tr>
</tbody>
</table>

*Different Department, Same Valence*

<table>
<thead>
<tr>
<th>Fallsburg University</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>English Department</strong></td>
</tr>
<tr>
<td>A -&gt; B -&gt; C</td>
</tr>
<tr>
<td>M -&gt; P -&gt; Q</td>
</tr>
<tr>
<td><strong>Political Science Department</strong></td>
</tr>
<tr>
<td>M -&gt; N -&gt; O</td>
</tr>
<tr>
<td>A -&gt; D -&gt; E</td>
</tr>
</tbody>
</table>

*Different Department, Opposite Valence*

<table>
<thead>
<tr>
<th>Fallsburg University</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>English Department</strong></td>
</tr>
<tr>
<td>A -&gt; B -&gt; C</td>
</tr>
<tr>
<td>M -&gt; P -&gt; Q</td>
</tr>
<tr>
<td><strong>Political Science Department</strong></td>
</tr>
<tr>
<td>M -&gt; N -&gt; O</td>
</tr>
<tr>
<td>A -&gt; P -&gt; Q</td>
</tr>
</tbody>
</table>

**Figure 4.** Illustration of design of Experiment 3. In this illustration, each fact is denoted by a different letter. The symbol '→' refers to a causal relation, so that the statement A → B should be read as 'A causes B'.
Results

Quiz

The scores on the quiz were analyzed with a 2 (Valence) x 2 (Department) ANOVA. Subjects in the same department condition scored higher on the quiz (m = 0.92) than did subjects in the different department condition (m = .83). This difference was marginally significant, F(1,43) = 3.36, .05 < p < .10. Subjects in the Different valence condition (m = 0.90) scored better on the quiz than did subjects in the Same valence condition (m = 0.85), but this difference did not approach reliability, F(1,43) = 1.52, p > .10. The interaction of these factors was not reliable F < 1.

Inference Test

The inferences were scored as shared-system, nonshared system and other. These data are presented in Fig. 5. A 3 (Inference Type) x 2 (Valence) x 2 (Department) mixed model ANOVA was performed on these data. This analysis revealed a significant main effect of inference type, F(2,86) = 30.87, p < .001. Post-hoc comparisons revealed that, as expected, on average significantly more shared-system inferences were made to each item (m = 1.45) than nonshared system inferences (m = 0.39), t(46) = 6.23, p < .001. Further, significantly more shared system inferences were made than other inferences, (m = 0.61), t(47) = 4.79, p < .001. The difference between the number of nonshared and other inferences listed was only marginally reliable, t(46) = 2.42, .05 < p < .10. The only other reliable effect was an interaction between inference type and valence, F(2,86) = 4.83, p < .05. One contributing source of this interaction can be seen by considering 2 x 2 ANOVAs done on the three inference types separately. The main effect of valence is reliable for shared system inferences, F(1,43) = 5.87, p < .05, but not for nonshared system inferences or for other inferences, both F < 1. Once again, no inconsistent substitutions were found in this study suggesting that one-to-one mapping was being strictly enforced.

As for Experiment 2, a tally was made of the number of inferences involving at least one element of both causal chains that started with the same causal antecedent. For the Same department/Same valence condition there were 11 cases of inferences from both causal chains out of 24 inference trials for subjects in this condition (12 subjects x 2 inferences). For the Same department/Opposite valence condition, there were 8 inferences of both causal chains out of a possible 24. For the Different department/Same valence condition, there were only 2 inferences of both causal chains out of 22 trials made. Finally, for the Different department/Opposite valence condition, there were only 3 inferences of both causal chains of a possible 24. Thus, subjects were clearly more likely to infer parts of both causal chains when the causal chains came from the same department, than when the causal chains came from different departments.

Of the 24 inferences of elements of both causal chains, only 5 of them (21%) mixed the causal chains. Three of these mixed inferences were made by subjects in the Same department/Same valence condition, with one each being made by subjects in the Same department/ Opposite valence and Different department/Opposite valence conditions.
Figure 5. Results of Experiment 3.
Discussion

These data are consistent with those of Experiments 1 and 2. Overall, subjects made more inferences of shared system facts than of nonshared system facts. Indeed, once again nonshared system facts were less frequently inferred than other facts that did not appear in the description of the base school at all. Both the consistency of valence and the configuration of the information in the base school had an impact on the inferences made in this study. Overall, more shared system inferences were made when the two causal chains had the same valence than when they had opposite valence. This finding suggests that when selecting inferences to make, inferences that seem compatible (e.g., two positive outcomes of a fact) are more likely to be selected than are inferences that do not seem compatible (e.g., a positive and a negative outcome of a fact).

The configuration of information also had an impact on the inferences made. Participants were more likely to make inferences of both causal chains when the causal chains both came from one department than when they came from two departments. This finding is consistent with the idea that corresponding facts between one target department and one base department can be made in a single interpretation, but that corresponding facts between one target department and two base departments require using two interpretations. Assuming that there is some cost to switching interpretations, it seems reasonable that there would be fewer inferences of both elements of a causal chain when the interpretation must be switched than when it does not.

Despite the fact that many inferences made by participants in the Same department condition contained elements of both causal chains, these inferences were rarely mixed together. In general, inferences of distinct causal chains were kept separate. Thus, while people may try to integrate information into a single causal story, this does not seem to be the dominant mode of processing in analogical inference.

These data shed some light on one unanticipated result from Experiment 1. In that study, when given the causal antecedent that the faculty size in the Music department was small, people often inferred that the courses in that department would be oversubscribed, even though this fact was actually causally related to the causal antecedent that the faculty were good teachers. The data from Experiment 3 suggest that there is not a strong tendency to mix causal chains, even when the same causal antecedent is connected to two distinct causal chains in the same department and the causal consequents have the same valence. Subjects in Experiment 1 do not seem to have been making the inference based on course oversubscription by analogy to the base. Rather, they seem to have constructed a separate causal link that was not based on the analogy using a property that was both accessible because it appeared in the description of the base domain and was appropriate given the antecedent of the inference.

There is one final aspect of this experimental situation that warrants further attention. In the first three studies, the base domain and target domain were presented in the same booklet, and subjects were encouraged to look back at the prior information when making inferences. It is possible that carrying over shared system facts is easiest when the base and target are both available for inspection. On this view, a smaller difference between the
number of shared system and nonshared system facts would be expected if the base domain was not available throughout the experiment. In order to assess this possibility, Experiment 3 was repeated with the change that the description of the departments in the old school were taken away before the description of the departments in the new school were presented. In this study the departments in the base domain were not available during the inference task.

EXPERIMENT 3A

The primary motivation for Experiment 3a is to replicate the basic results of Experiment 3 in which the base domain is in memory. However, there is one other issue that will be addressed. In Experiments 1 and 2, all of the correspondences between departments that were relevant to the correspondence task were also antecedents for causal chains that were used in the inference task. However, it should be possible to further dissociate the correspondence task from the inference task. Placing a pair of departments in correspondence requires only a matching property, while making an inference requires a shared system fact. In previous work, Lassaline (1996) found that adding a matching property between a premise category and conclusion category in an inductive argument increased the perceived similarity of those categories, but not the strength of an inductive argument. In her studies, only relations that linked matching information to a piece of information to be inferred increased inductive strength. The setup of Experiment 3a parallels that of Lassaline’s study. In this study, a commonality between a department in the new school (the Computer Science department) and a department in the old school (the English department) was added, but this commonality was not related to any of the properties involved in the inference task. Given this setup, people should find the correspondence between the Computer Science and English departments in the correspondence task, but this correspondence should not affect the inference task, because it does not involve shared system facts.

Method

Subjects

Subjects in this experiment were 80 members of the Columbia University community (20 per condition) who were paid $6.00 for participating. Two subjects were dropped from the experiment for failing to follow instructions.

Materials

The materials for this study are identical to those of Experiment 3 with one small change. In the description of the Computer Science department in the new school, a sentence was added mentioning that the faculty in this department were socially active. This fact corresponds to a fact about the English department. This fact was not relevant to any of the inference trials.
Procedure and Design

This experiment is identical to Experiment 3, with two exceptions. First, the description of the departments from the new school and the inference task were placed into a separate booklet from the description of the departments from the old school and the quiz. Second, subjects in this study participated in the correspondence task before doing the inference task. Subjects were first given the book describing the old school, and they took the quiz. After completing the quiz, they exchanged the first booklet for a second containing descriptions of the target schools and the correspondence and inference tasks.

Results

Quiz

The scores on the quiz were analyzed with a 2 (Valence) x 2 (Department) ANOVA. The only reliable effect revealed by this analysis was a main effect of department, with subjects in the same department condition (m = .92) performing more accurately on the quiz than subjects in the different department condition (m = 0.82), F(1,74) = 8.50, p < .01.

Correspondence Test

In this study, we are primarily interested in whether subjects found the correspondence between the Computer Science department in the new school, and the English department in the old school. As expected, 66/78 (85%) subjects made this correspondence. For purposes of comparison, the Political Science department and Computer Science department, which also had a corresponding fact, were placed in correspondence by 57/78 (73%) subjects. In contrast, the Computer Science and Biology departments, which did not share any facts, were placed in correspondence by only 9/78 (12%) of subjects.

Inference Test

Once again, the inferences were scored as shared-system facts, nonshared system facts and other. These data are presented in Fig. 6. A 3 (Inference Type) x 2 (Valence) x 2 (Department) mixed model ANOVA was done on the data. This analysis revealed a reliable difference in the number of each type of inference, F(2,148) = 35.57, p < .001. As in the previous experiments, significantly more shared system inferences were made overall (m = 1.41) than nonshared system inferences (m = 0.36), t(77) = 7.20, p < .005. More shared system inferences were also made than other inferences (m = 1.02), but this difference was only marginally significant t(77) = 2.35, .05 < p < .10. Finally, more other inferences were made than nonshared system inferences, t(77) = 6.41, p < .005. Once again, no inferences with inconsistent argument substitutions were found, suggesting that one-to-one mapping was maintained during analogical inference.

The ANOVA illuminated a somewhat more complex pattern of data than was obtained in Experiment 3. Significantly more inferences overall were made when both chains emanating from a single causal antecedent were from the same department (m = 2.99) then
Figure 6. Results of Experiment 3a.
when they were from different departments \( (m = 2.55) \), \( F(1,74) = 4.99, p < .05 \). Surprisingly, more inferences were made overall for opposite valence items \( (m = 2.37) \) than for same valence items \( (m = 3.18) \), \( F(1,74) = 18.85, p < .001 \). These main effects must be interpreted in light of a pair of two-way interactions: one between Inference Type and Valence, \( F(2,148) = 5.71, p < .01 \), and a second between Inference Type and Department, \( F(2,148) = 19.88, p < .001 \).

The source of these interactions can be seen by examining separate 2 x 2 ANOVAs for each of the inference types separately. A 2 (Department) x 2 (Valence) ANOVA on the number of shared system inferences reveals main effects of both Valence, \( F(1,74) = 15.10, p < .005 \), and Department, \( F(1,74) = 29.49, p < .005 \). A similar ANOVA on the number of nonshared system inferences reveals only a main effect of Department, \( F(1,74) = 10.56, p < .05 \). Finally, the ANOVA on the number of other inferences reveals no significant effects. Thus, the number of shared system inferences made is influenced strongly by whether the causal chains are both from the same department as well as whether both causal chains have the same valence. In contrast, the number of nonshared system inferences is influenced strongly only by whether the causal chains with the same antecedent come from the same department. The number of other inferences made is not strongly influenced by either factor.

A central prediction in these studies has been that more shared-system inferences would be made than nonshared system inferences. Separate t-tests comparing the number of shared system and nonshared system inferences in the four experimental conditions (with alpha levels corrected by the Bonferroni adjustment) reveals reliable differences for all of the conditions except when the causal chains with the same antecedent come from different departments and have the same valence. It is not clear why this condition did not yield a reliable difference, particularly given that this condition yielded more shared system inferences than nonshared system inferences in Experiment 2.

Once again, a tally was made of the number of inferences of at least one fact from the causal consequents from two causal chains. In the Same Department, Same Valence condition, 13/40 inference trials yielded inferences of one element from each causal chain. Of these 2 (15%) could be interpreted as mixtures of the causal chains. In the Same Department, Opposite Valence condition, 23/40 inference trials yielded inferences of one element from each causal chain. Of these, 3 (13%) could be interpreted as mixtures of the causal chains. The remaining two conditions each yielded 4 inferences of elements from both causal chains, and none of these inferences could be interpreted as mixtures of the causal chains. Thus, overall, only 5/44 (11%) of the inferences could be interpreted as mixtures of two different causal chains in inference. This total is lower than the 23% of inferences that could be interpreted as mixtures in Experiment 3. As we found before, subjects were more likely to infer elements from both causal chains when they came from the same department (36 cases) than when they came from different departments (8 cases).

Finally, in the inference task, subjects were asked to make inferences about what would happen based on the degree of research funding in the Computer Science department. These inferences were examined to see whether any inferences came from facts that appeared in the description of the English department. As discussed above, the English
department and Computer Science department corresponded on the basis of a property (social activism) that was not connected to information about research funding. Across the study, only 10 inferences were made in response to the inference question about research funding that came from the English department. Of these, only 3 were related to the social activism that was the corresponding fact between the two departments. Thus, it does not appear that simply placing a pair of departments in correspondence has a strong impact on subjects' performance on the inference task.

Discussion

The main purpose of this study was to replicate the basic pattern of results in a study where subjects could not look back at the departments in the base domain while making inferences. This basic goal was satisfied, as these results are broadly similar to those of Experiment 3. Indeed, on average, the number of shared system and nonshared system inferences in Experiment 3a is nearly identical to the number shared system and nonshared system inferences made by subjects in Experiment 3. In general, subjects made more shared system inferences than nonshared system inferences. Further, as before, more inferences were made when both causal chains with a common antecedent came from the same department in the base. One surprising result given the findings of Experiment 3 was that causal chains with opposite valence actually gave rise to more inferences in Experiment 3a, where they had given rise to fewer inferences in Experiment 3. It is not clear why this should be the case, but it is possible that causal chains of opposite valence are more distinctive in memory, and hence that there is less interference between causal chains. This topic, should be explored in further research. Finally, as in the previous studies, there was very little evidence that subjects mixed together distinct causal chains emanating from the same causal antecedent.

This study also helps clarify the difference between the correspondence task and the inference task. In this study, subjects in the correspondence task mapped departments together when they shared salient common features. However, simply placing a pair of departments in correspondence in this task did not lead subjects to make inferences based on that correspondence. Rather, subjects focused on shared system facts in the inference task. This finding is consistent with Lassaline's (1996) previous findings in inductive inference, and reinforces the importance of systematicity—as opposed to mere property overlap—in inference.

GENERAL DISCUSSION

It is generally accepted that analogical inferences involve a form of copying with substitution and generation. However, there must be mechanisms that keep people from making too many inferences that are likely to be false, otherwise the benefit of making inferences by analogy will be outweighed by the cost of checking for potentially false inferences. The studies presented here suggest that there is a bias to make shared system inferences, a finding that is consistent with other recent work (Bowdle & Gentner, in press; Clement & Gentner, 1991; Keane, 1996; Spellman & Holyoak, 1996). As discussed above, the focus on
shared system facts is particularly striking given that there were always far more nonshared system facts than shared system facts that could have been inferred. Thus, the data provide a strong demonstration that subjects focus on shared system facts in inference.

Not only were the inferences made up of shared system facts, but even when shared system facts from both systems connected to the causal antecedent were inferred, these inferences were kept separate. Indeed, in Experiments 3 and 3a, these inferences were kept separate even when they could easily have been integrated into a single causal story. Thus, the analogical inference process is quite conservative. It is a good strategy to be conservative in analogical inference, because analogical inferences are not guaranteed to be true and must be checked by some other process to assess whether they are factually correct.

These findings also help clarify the importance of one-to-one mapping in analogical inference. As discussed above, many-to-one mappings open the possibility for inconsistent substitutions in the process of copying with substitution and generation. In contrast to this possibility, in the four studies in which potential many-to-one mappings between departments were presented, there were no cases of inconsistent substitutions. Thus, subjects maintain one-to-one mapping in situations where they are making analogical inferences.

Potential many-to-one mappings may be resolved by enforcing one-to-one mapping and generating multiple interpretations of the match between base and target. If people are generating more than one interpretation, then there should be some difficulty in integrating information across interpretations relative to a case where only a single interpretation is required. In Experiments 3 and 3a, people generated more inferences and were more likely to generate both causal chains coming from a single causal antecedent when the antecedents were in the same department than when they were in different departments. The same department situation requires only a single interpretation of the match between base and target, while the different department situation, which creates a potential many-to-one mapping, involves two interpretations. Thus, this finding suggests that it was more effortful to generate and use two interpretations than to generate and use one.

**Constraints on Analogical Inference**

These data suggest that systematicity is an important constraint on what information is carried over from base to target. Nonetheless, it is not the only constraint on inference. Another factor that may influence analogical inference is the importance or salience of the information, as high salience information may be a more likely candidate to be carried from base to target than low salience information. One factor that may have influenced salience is causality. Lassaline (1996) demonstrated that causal relations are more salient in inductive arguments than are temporal precedence relations. In the present studies, all of the inferences involved causal relations. Fewer inferences might be observed in tasks that involve other relations. Salience may also be involved in the generation of nonshared system inferences. It is possible that nonshared system facts may be facts that are salient for subjects.

Another central factor that governs inference is pragmatics. As discussed above, students who learn to solve problems by example learn that the relevant thing to carry from an example to a new domain is the solution to the problem. Psychological studies of problem
solving reflect this understanding, as subjects typically carry over the solution procedure from the example to structure their knowledge of the target problem (Gick & Holyoak, 1980; Novick, 1988, 1990; Ross, 1987, 1989).

Pragmatic information may manifest itself in other ways as well. Scientists may enter into a situation in which they use analogy to find the answer to a particular problem (Dunbar, 1995). Modern science also places a premium on structural consistency in explanatory analogies (Gentner & Jerzioski, 1993). In science, inferences containing information relevant to the problem being solved are of paramount importance, and potential inferences unrelated to that information are probably ignored.

Another pragmatic factor is adaptability. Keane (1996) held the structure of the base and target domains constant, and varied only how easily the solution from the base situation could be adapted to the target. Subjects were more likely to make inferences for problem solutions that could be more readily adapted to the problem context. Interestingly, subjects in Keane's studies typically made only one inference on each trial, while subjects in the present studies often made inferences from more than one causal chain in a single trial. One difference between studies that may explain this discrepancy is that subjects in Keane's study were in a problem solving context in which they were to find a solution to a problem, while subjects in the present study were in a context in which they were supposed to find any benefits or problems that might arise as a result of a particular fact. Thus, the context in this study may have promoted attention to more than one inference.

Computational models of analogy

These data suggest two important constraints on analogical inference. First, models of analogy must focus on shared system facts when making inferences. Second, inferences must maintain structural consistency, and hence one-to-one mapping must be enforced. As discussed above, the Structure Mapping Engine, a computational model of structure-mapping theory implements both of these constraints (Falkenhainer, et al., 1989). Analogical inferences are only carried from base to target when they are shared system facts. Further, the model strictly enforces structural consistency, and so inferences that would lead to violations of one-to-one mapping are not carried over. Finally, a pragmatic marking procedure has been added to SME that allows it to focus only on those inferences that are pragmatically important for the current situation (Forbus & Oblinger, 1990).

Other models approach analogical inference differently. The basic mechanism used by ACME (Holyoak & Thagard, 1989) is one of copying with substitution and generation. ACME is a parallel constraint satisfaction algorithm in which structural constraints including one-to-one mapping are 'soft' constraints that can be violated if such a violation will provide a good match. Because it uses CWSG, ACME is capable of generating matches based on any correspondence between base and target. It is not constrained to make shared system mappings. However, ACME has incorporated other interesting constraints on analogy like pragmatic centrality. It can be set up so that only pragmatically important inferences are made. While this will not capture the preference for shared-system facts, it is another important constraint on inference. Finally, ACME treats one-to-one mapping as a soft constraint that may be violated. Hence, it will make inferences with inconsistent argu-
ment substitutions when its interpretation of a comparison involves many-to-one correspondences. Increasing the relative importance of one-to-one mapping when inferences must be made can potentially eliminate these inconsistent substitutions.

Another important model of analogy is the Incremental Analogy Machine (IAM; Keane & Brayshaw, 1988; Keane, Ledgeway, & Duff, 1994). IAM is designed specifically to model the incremental interpretation of analogies. In IAM, one structure from the base is selected as a seed and is mapped to the target. If necessary, more structures from the base are selected to augment the initial mapping. Once correspondences have been formed between structures in the base and target, a CWSG process is applied to fill in structure from the base that is not part of the target domain. This algorithm places an interesting constraint on analogical inference. Inferences are effectively limited to shared system facts by carrying over only the information from a relational group that was used as the basis of an analogical match. Further, because pragmatic information can be used to guide which aspects of the base and target take part in a match, the algorithm incorporates pragmatics and salience easily. Finally, like SME, IAM forms isomorphic matches, and hence does not have the problems with inference that many-to-one mappings can create.

Finally, a new and interesting model of analogy is LISA (Hummel & Holyoak, 1997). LISA uses a connectionist procedure of dynamic binding by temporal synchrony to represent relational structures. Base and target elements come to be mapped through the sequential firing of units in the base, which, through learned connections between base and target, can lead corresponding target elements to fire in phase as well. Candidate inference are made by recruiting units to the target domain to represent structure in the base not present in the target. Without any further constraints, this type of inference is a form of CWSG (Hummel & Holyoak, in press). Thus, without the addition of further constraints, LISA does not focus on shared system facts. Further, because the model permits violations of one-to-one mapping, it can lead to inconsistent object substitutions in inference. However, there are a number of ways that LISA can be extended. For example, by having some units fire more frequently than others, LISA can be focused on pragmatically important or salient information. Further, at the time that LISA starts making inferences, shared system elements in the base could have their firing frequency increased, thereby raising the likelihood that shared system inferences would be made. Because LISA is a newcomer to the field, the present data should be viewed as placing constraints on its candidate inference mechanism rather than as evidence against the model’s assumptions.

Uses of Analogical Inference

The study of analogy is a central area of cognitive science. Much of this work has been devoted to understanding the constraints underlying the comparison process. Other work, like that in analogical problem solving, is set in a domain in which there is a strong pragmatic goal that influences which information will be carried from base to target (Gick & Holyoak, 1980; Novick, 1988, 1990; Novick & Holyoak, 1991; Ross, 1987, 1989). However, it is also important to understand how analogical processes may operate in other domains. Much of the previous work on the influence of prior knowledge has focused on the effects of the statistical structure of the background knowledge. For example, in cate-
gorization Heit (1994) examined mathematical models of the use of background knowledge in categorization, in which the effects of background knowledge consisted of information about the distributions of attribute values in previous categories that were related to the category currently being learned. This work suggests that people use background knowledge in the acquisition of new categories by retrieving information about the distribution of features in previous similar cases and integrating that information with the distribution of attributes in the new category. In decision making, Ross and Creyer (1992) focused on correlations between attribute values both within and across options in order to understand how missing attributes of options might be filled in. When an attribute was missing, subjects could fill in that missing value by attending to the distribution of values of that attribute for other options.

The current studies suggest that prior knowledge may also have an impact by licensing candidate inferences that can be used to add information to other representations. On this view, prior knowledge provides some information about the distribution of features in particular categories, and it also provides conceptual relations that connect those features. For work in categorization, this possibility suggests that making inferences using new categories may focus subjects on relational information within a category. Some work consistent with this hypothesis has been done by Lassaline and Murphy (1996) and by Markman, Yamauchi, & Makin (1997). In this work, subjects were asked to guess feature values of different dimensions of novel stimuli. In both cases, this task focused participants on relationships between attributes within the categories. However, these studies used stimuli in which attributes only had a correlational structure. Further work must focus on the impact of inference in category acquisition using materials that embody conceptual relations.

A similar extension may be profitable in other domains that involve comparisons, like decision making. Previous research has suggested that subjects rarely make spontaneous inferences of missing attribute values (Gardial & Biehal, 1991). However, like the categorization work, studies of choice in which this result has been obtained have used only statistical correlations between features. If conceptual relations between features were introduced into these materials, then spontaneous analogical inference might be observed. Further research should address this possibility.

**CONCLUSIONS**

To summarize, analogical inference is a selective process that carries information from a base domain to a target domain when that information is consistent with the information in the target and is connected to matching structure. In order for a copy with substitution and generation inference mechanism to avoid inconsistent substitutions, the mapping process must strictly enforce one-to-one mapping. The observed pattern of data is consistent with models of analogical reasoning that maintain isomorphic matches between base and target like SME and IAM. These data may also be useful for constraining the development of new models like LISA. Finally, the process of analogical inference may be profitably extended to other domains in which prior knowledge is expected to be used to augment the representations of new items.
APPENDIX A

Descriptions of the base departments from the consistent condition of Experiment 1.

Fallsburg University

Biology Department

The biology department at Fallsburg teaches many of the pre-medical courses for the students, and also prides itself on maintaining an active research program. Over the last few years, a number of the older faculty have taken advantage of an early retirement program offered by the state, and, three younger faculty were lured to another school with offers of higher salaries. These departures have left the Biology department with fewer faculty than they would like to have. Because there are fewer faculty than are needed in this department, it has been difficult for them to carry out all of the administrative work that needs to be done. There are not enough faculty to organize meetings for faculty business and to advise majors in biology.

The undergraduate students have consistently rated the faculty in this department as the best teachers on campus. Their reputation as fine teachers has spread throughout the campus so that almost every student wants to take at least one Biology course during their studies. This popularity has caused most of the Biology courses to be oversubscribed. Students must often wait over a year to get into the course of their choice.

Department of Political Science

The Political Science department has an odd mix of faculty. About half of the professors have a strong interest in the role of Socialist and Marxist ideology in world government. The other half of the faculty study the emergence of democratic republics, particularly in Third World nations. Despite these radically different orientations, most of the professors are quite open-minded. In fact, the department recently organized a conference at which both groups gave talks about their findings. The conference was quite well received, and a favorable story about the meeting appeared in The Washington Post.

Many of the faculty in the Political Science department got their degrees in the 1960s. This era has clearly left its mark on the faculty. They are vocal proponents of environmental causes, and they often encourage their students to speak out on these issues. Because of their well-known stance as activists, many members of the faculty have been called to testify before Congress

Department of English

The English Department is one of the largest departments at Fallsburg. Unfortunately, many of the faculty in the English department have strong personalities and big egos. Inevitably, this has led to personality clashes among the faculty. It is not uncommon for faculty meetings in this department to erupt into shouting matches. One result of these arguments
is that the faculty rarely go to their University offices, except during office hours to avoid running into their colleagues.

The faculty in English are quite active researchers, and almost all of them have research grants from large government agencies. These grants provide the department with money to hire personnel. Because of the large number of grants, the department has been able to hire research assistants for every faculty member.

Descriptions of the base departments from the inconsistent condition of Experiment 1

Department of Biology

The biology department at Fallsburg teaches many of the pre-medical courses for the students, and also prides itself on maintaining an active research program. Over the last few years, a number of the older faculty have taken advantage of an early retirement program offered by the state, and, three younger faculty were lured to another school with offers of higher salaries. These departures have left the Biology department with fewer faculty than they would like to have. Because there are fewer faculty than are needed in this department, it has been difficult for them to carry out all of the administrative work that needs to be done. There are not enough faculty to organize meetings for faculty business and to advise majors in biology.

Many of the faculty in the biology department got their degrees in the 1960s. This era has clearly left its mark on the faculty. They are vocal proponents of environmental causes, and they often encourage their students to speak out on these issues. Because of their well-known stance as activists, many members of the faculty have been called to testify before Congress.

Department of Political Science

The Political Science department has an odd mix of faculty. About half of the professors have a strong interest in the role of Socialist and Marxist ideology in world government. The other half of the faculty study the emergence of democratic republics, particularly in Third World nations. Despite these radically different orientations, most of the professors are quite open-minded. In fact, the department recently organized a conference at which both groups gave talks about their findings. The conference was quite well received, and a favorable story about the meeting appeared in The Washington Post.

The faculty in Political Science are quite active researchers, and almost all of them have research grants from large government agencies. These grants provide the department with money to hire personnel. Because of the large number of grants, the department has been able to hire research assistants for every faculty member.

Department of English

The English department at Fallsburg is one of the largest departments at the University. The undergraduate students have consistently rated the faculty in this department as the
best teachers on campus. Their reputation as fine teachers has spread throughout the cam-
pus so that almost every student wants to take at least one English course during their stud-
ies. This popularity has caused most of the English courses to be oversubscribed. Students
must often wait over a year to get into the course of their choice.

Unfortunately, many of the faculty in the English department have strong personalities
and big egos. Inevitably, this has led to personality clashes among the faculty. It is not
uncommon for faculty meetings in this department to erupt into shouting matches. One
result of these arguments is that the faculty rarely go to their University offices, except dur-
ing office hours to avoid running into their colleagues.

Descriptions of target departments in this study.

Gordmont University

The Music Department

The music department at Gordmont is a relatively new department, begun in 1987 in
response to student pressure for more music classes. Because of the focus on undergradu-
ate education at Gordmont, the University searched for the best instructors that it could
find. The success of this effort is evident in the fact that the faculty in this department have
been awarded the prestigious University Teaching Award in each of the last three years.
Interestingly, the music department is also the smallest on campus. Because it is a new
department, it has taken some time to hire as many faculty as were originally planned when
the department was formed.

The Computer Science Department

The computer science department at Gordmont is widely acknowledged to be one of the
finest groups of its kind in the world. Over the last five years, the department has been
ranked number one among computer science departments in terms of the number of publi-
cations by the faculty. This success has allowed the members of the faculty to secure a
number of large contracts that provide funds for the professors to carry out their research.
Unfortunately, the research reputation of this department is only outdone by the personal
reputations of the individual faculty. Almost all of the professors are known to be difficult
to get along with. Indeed, the individual members of the department seem to have little or
no respect for each other.

APPENDIX B

This appendix contains descriptions of the departments from the base school in Experiment
2. This is one stimulus set. In the second stimulus set, paragraphs with the same causal
antecedent were switched to form different pairings, as described in the methods section of
Experiment 2.
Fallsburg University

Department of Biology

Unfortunately, many of the faculty in the Biology department at Fallsburg have strong personalities and big egos. Inevitably, this has led to personality clashes among the faculty. It is not uncommon for faculty meetings in this department to erupt into shouting matches. One result of these arguments is that the faculty rarely go to their University offices, except during office hours to avoid running into their colleagues. Because the faculty are rarely in their offices, undergraduates and graduate students alike complain that the faculty are inaccessible.

One of the most striking aspects of the Biology department is the tremendous financial support for research. One result of this has been a high level of motivation for faculty members to excel at their research. Faculty members in this department often spend long hours on their research. This hard work has paid off, as the department is among the top 10 of all Biology departments in its research reputation nationally.

Department of Political Science

The faculty in the Political Science department have strenuously and successfully pursued research grants to allow them to conduct their research. As a result, this department has more money for the pursuit of research than any other department at the University. This funding has allowed the faculty to hire research assistants for all of the faculty members. Because every faculty member has a research assistant, the faculty have had enough time to maintain high-level research careers without sacrificing the quality of their teaching.

The Political Science department at Fallsburg is one of the largest departments at the University. The undergraduate students have consistently rated the faculty in this department as the best teachers on campus. Their reputation as fine teachers has spread throughout the campus so that almost every student wants to take at least one Political Science course during their studies. This popularity has caused most of the Political Science courses to be oversubscribed. Students must often wait over a year to get into the course of their choice.

Department of English

Many of the faculty in the English department got their degrees in the 1960s. This era has clearly left its mark on the faculty. They are vocal proponents of environmental causes, and they often encourage their students to speak out on these issues. Because of their well-known stance as activists, many members of the faculty have been called to testify before Congress. As a result, many members of this department have become recognizable faces on television news programs.
The English department faculty also have excellent reputations as teachers. Not only are the faculty in this department extremely bright, they are able to enter the "beginner's mind" and present material in a clear engaging manner. Because of this teaching skill, the English faculty have developed a rather smug attitude, and often look down on members of other departments. As a result, the members of the English department are resented by many faculty from other departments.

Department of Anthropology

The degree of animosity between members of the faculty in Anthropology is obvious to anyone who looks at the department. This has not gone unnoticed at other Universities, where it has been viewed as a potential opportunity, and at least once per semester, a faculty member in the department is offered another job at a prestigious University. The offers from outside schools has led a number of faculty members to leave in recent years.

Because of their work with people from other cultures, the Anthropology faculty tend to be politically active. In particular, the faculty in this department are on the boards of many charities. In their classes, they organize fund-raising drives for multi-cultural causes. This activity has fostered a spirit of giving at the University, and has led to a strong community involvement among students.

Description of target departments in Experiment 2.

Gordmont University

The Music Department

The music department at Gordmont is a relatively new department, begun in 1987 in response to student pressure for more music classes. Because of the focus on undergraduate education at Gordmont, the University searched for the best instructors that it could find. The success of this effort is evident in the fact that the faculty in this department have been awarded the prestigious University Teaching Award in each of the last three years. Interestingly, because it is a new department, there has been a long period of adjustment in faculty relations. The faculty do not seem to get along well at all, and the members of the department seem to have little or no respect for each other.

The Computer Science Department

The computer science department at Gordmont is widely acknowledged to be one of the finest groups of its kind in the world. Recently, the department has begun to attract research funding from both federal granting agencies and private research foundations. Currently, the department is one of the best funded departments on campus. Despite the focus on research, the faculty have bonded together in a spirit of social activism. The members of the department often organize campus-wide events like a recycling-awareness day.
Acknowledgments: Please address all correspondence to Arthur B. Markman, Department of Psychology, Columbia University, 406 Schermerhorn Hall, New York, NY 10027. This work was supported by NSF CAREER award SBR-95-10924. Very special thanks to Keith Holyoak for running ACME simulations to confirm predictions of that model. The author would also like to thank Ron Ferguson, Ken Forbus, Dedre Gentner, Jim Greeno, Tory Higgins, Mark Keane, Dave Krantz, Tomislav Pavlicic, Saskia Traill, Takashi Yamauchi, Shi Zhang, an anonymous reviewer, and the Similarity and Cognition Lab for comments and helpful discussions during the evolution of this project. Thanks to Adalis Sanchez for help running subjects. Thanks also to Kerry Kittles for making the best of a pitiful season, and Evander Holyfield for reaching the ACME of his career.

NOTES

1. Very special thanks to Keith Holyoak for running these ACME simulations in response to a query about the model’s behavior.
2. Another set of SME simulations was done in which the representations contained more of the information that is included in the stories that were actually presented to participants (shown in Appendix A). These simulations show the same pattern of results. However, for expository purposes, representations with less information in them were used, because it makes it easier to see the aspects of the simulation that cause its behavior.
3. The alpha-levels for all post-hoc tests have been corrected using the Bonferroni procedure.
4. Actually, as shown in the descriptions in Appendix A, the faculty were said to be unable to complete administrative duties This was shortened to advising for the figure.
5. As with the earlier figures, this figure is meant to illustrate the conditions of the study. The actual stories seen by participants differed somewhat from what is shown in this Figure.
6. One subject was dropped from this condition due to experimenter error.

REFERENCES


