The Role of Textual Coherence in Incremental Analogical Mapping

Tate T. Kubose

Columbia University

Keith J. Holyoak   John E. Hummel

University of California, Los Angeles

Address for proofs: Tate Kubose
                        Department of Medical Informatics
                        Columbia University
                        622 W. 168th St. VC-5
                        New York, NY 10032

Email: tate.kubose@dmi.columbia.edu

Abstract

The LISA model of analogical reasoning (Hummel & Holyoak, 1997) assumes that mapping is performed incrementally within limited-capacity working memory, and that processing is guided by principles of text coherence. Predictions of the model derived by computer simulations were tested in four experiments, using both semantically-impoverished structural analogs and semantically-rich story analogs. Results for college students revealed a grouping effect. Processing multiple propositions together generated more accurate mappings than did processing individual propositions, but only when propositions that jointly provided strong structural constraints were grouped together. Other experiments revealed asymmetries in mapping and inference accuracy: mappings and inferences generated from a more coherent analog to a less coherent analog were more accurate than those made in the reverse direction. Implications for computational models of analogical reasoning and for education are discussed.

Keywords: analogy, mapping, coherence, computational model, working memory

Human intelligence is distinguished by a breadth of scope and flexibility of operation that is unrivalled by any other form of intelligence, either biological or artificial. Yet human thinking must operate within severe constraints imposed by limited working memory (WM; see Baddeley, 1992) and attention. For example, empirical evidence indicates that analogical mapping—comparing two structured situations to identify systematic correspondences—places substantial demands on WM. Waltz, Lau, Grewal and Holyoak (2000) have shown that adding a secondary task to reduce available WM capacity causes a shift from the use of relations to reliance on simpler element similarity as the basis for mapping (see also Tohill & Holyoak, 2000). How can a complex form of reasoning such as analogy be performed within a system with limited WM?

Working-Memory Limits and Analogical Mapping

In recent years a number of computational models of analogical mapping have been developed, including ACME (Analogical Constraint Mapping Engine; Holyoak & Thagard, 1989), SME (Structure Mapping Engine; Falkenhainer, Forbus & Gentner, 1989), IAM (Incremental Analogical Machine; Keane & Brayshaw, 1988; Keane, Ledgeway & Duff, 1994), STAR (Structural Tensor Analogical Reasoning; Halford et al., 1994), and LISA (Learning and Inference with Schemas and Analogies; Hummel & Holyoak, 1997). Several of these models (or variations of them) serialize analogical processing by considering smaller portions of the analogs at a time, allowing the overall mapping to emerge incrementally. The incremental approach to mapping was first introduced in IAM (Keane & Brayshaw, 1988), and has also been incorporated into Copycat (Hofstadter & Mitchell, 1994) and AMBR (Kokinov, 1994). An incremental version of SME called I-SME (Forbus, Ferguson, & Gentner, 1994) has been developed, and models similar to ACME have also been applied in incremental fashion (Hoadley, Ranney & Schank, 1994; Ranney & Thagard, 1988).

But although incremental analogical mapping is psychologically plausible, many of the incremental models do not provide a strong basis for the underlying assumption that WM capacity is inherently limited. For example, although I-SME can extend initial mappings if additional information is provided, it operates on whatever information about the analogs is given at once. Thus if entire analogs are initially provided (as in most experiments in the literature), I-SME becomes identical to SME, a model that has no particular capacity limit. IAM is more constrained, as it breaks down the information in the source analog into subsets, and begins by mapping only the subset with the most higher-order structure. However, IAM does not specify any limit on the size of that subset (i.e., the subset with the most structure can be arbitrarily large). The STAR model (Halford et al., 1994) does posit inherent capacity limits on analogical reasoning; however, that model does not handle mappings between hierarchically-structured analogies of the sort that commonly arise in everyday problem solving and comprehension (Holyoak & Thagard, 1995).

The focus of the present paper is on the LISA model Hummel & Holyoak, 1997). This model can map hierarchical structures, and is based on a type of architecture that provides a basis for estimating the maximum capacity available for mapping. LISA distinguishes WM from a broader *active memory* (cf. *long-term working memory*; Ericsson & Kintsch, 1995), which is the active subset of long-term memory. During analogical mapping, the two analogs and the emerging mappings between them are assumed to reside in active memory. Following Cowan (1995), LISA assumes that active memory has no firm capacity limit, but rather is time-limited, with activation decaying in roughly 20 sec unless it is reactivated by allocation of attention.

LISA assumes that within active memory, at any time a very small number of role bindings are in WM, and that these constitute the immediate focus of attention. In LISA's WM, distributed representations of predicates and their arguments are dynamically bound into

propositional structures by synchrony of firing: Units representing the semantic features of predicate roles fire in synchrony with units representing the features of the fillers of those roles, and separate role-filler bindings fire out of synchrony with one another. For example, to represent the proposition Abe loves Cathy, units representing Abe (e.g., human, male, adult, etc.) fire in synchrony with units representing "lover" (e.g., emotion, positive, strong), while units for Cathy fire in synchrony with units for "beloved". Crucially, the Abe-as-lover set must fire *out* of synchrony with the Cathy-as-beloved set. The collection of active but mutually desynchronized role bindings is referred to as the *phase set*, and corresponds to the model's working memory— i.e., the collection of role-bindings it is thinking about "right now".

The size of the phase set—i.e., the capacity of WM—is determined by the number of role-filler bindings (phases) it is possible to have simultaneously active but mutually out of synchrony. This number is necessarily limited (assuming the underlying neural substrate is subject to random noise), and its value is proportional to the length of time between successive peaks in a single phase (the period of the oscillation) divided by the duration of each phase (at the level of small populations of neurons) and/or temporal precision (at the level of individual spikes; see also Lisman & Idiart, 1995). There is evidence that binding is accomplished by synchrony in the 40 hz (gamma) range, meaning a neuron or population of neurons generates one spike (or burst) approximately every 25 ms (see Singer & Gray, 1995, for a review). If the temporal precision of spike timing is in the range of 4 - 6 ms (Singer & Gray, 1995), then with a 25 ms period, the capacity of WM ought to be approximately 25/5 = 5 role bindings (see also Lisman & Idiart, 1995; Luck & Vogel, 1997; Singer & Gray, 1995). LISA's architecture, by virtue of its use synchrony as the basis for dynamic binding, implies a comparable limit on the maximum amount of information that can be processed together during analogical mapping: 4-6 role bindings. Since propositions typically contain 2-3 roles (corresponding syntactically to a

subject, direct object, and perhaps indirect object), it follows that 4-6 role bindings is equivalent

to roughly 2-3 propositions. Other researchers (e.g., Hummel & Biederman, 1992; Hummel &

Holyoak, 1993, 1997; Hummel & Stankiewicz, 1996; Lisman & Idiart, 1995; Shastri &

Ajjanagadde, 1993) have suggested similar explanations for the capacity limits of visual attention

and working memory. Behavioral work with human subjects suggests a remarkably similar

figure for the capacity of visual working memory (four to five items; Bundesen, 1998; Luck &

Vogel, 1997; Sperling, 1960) and working memory for relations (three to five items; Broadbent,

1975; Cowan, 1995).

Capacity Limits and Incremental Mapping in LISA

Because of the strong capacity limit on the phase set, LISA's processing of complex

analogs is necessarily incremental: LISA can only hold at most three propositions in WM

simultaneously, so it must map large analogies in small pieces. The resulting algorithm (which

serves as the basis of analog retrieval, mapping, inference and schema induction; Hummel &

Holyoak, 1997, in press) is analogous to a form of guided pattern recognition. At any given

moment, one analog is the focus of attention and serves as the *driver*. One (or at most three) at a

time, propositions in the driver become active, generating synchronized patterns of activation on

a collection of units representing the semantic features of objects and relational roles (one pattern

for each role-filler binding in WM). During analogical mapping and inference, these patterns

activate propositions in a *recipient* analog held in active memory. (Note that the distinction

between driver and recipient depends on which analog is currently the focus of attention, and

thus is not equivalent to the distinction between source and target analogs, which is generally

defined in terms of the greater familiarity of the source.) Units in the recipient compete to

respond to the semantic patterns in much the same way that units for words or objects compete to

respond to visual features in models of word recognition (e.g., McClelland & Rumelhart, 1981)

and object recognition (e.g., Hummel & Biederman, 1992). LISA learns analogical correspondences by updating weights on *mapping connections*, which link elements in the driver and recipient (see Holyoak & Hummel, 1997, for a detailed description of the LISA model and its operation).

The incremental nature of mapping, as imposed by LISA's limited-capacity working memory, is analogous to the sequential scanning of a visual scene, as imposed by the limits of visual working memory (see, e.g., Irwin, 1992). Just as the visual field typically contains much more information than can be simultaneously attended, so any given analogy will typically contain more information than can be simultaneously placed in WM. And just as eye movements systematically guide the flow of information through foveal vision, the environment and the reasoner's goals guide the flow of propositions through LISA's WM.

As a general default, LISA assumes that propositions in the driver will fire in roughly the order they would naturally be stated in a text—that is, in accord with the principles of text coherence (see Hummel & Holyoak, 1997)—and that the most important propositions (i.e., pragmatically central propositions, such as those on the causal chain of the episode, and those pertaining to goals; see Spellman & Holyoak, 1996) tend to fire earlier and more often than less important propositions.[1] Also by default, propositions are fired one at a time (i.e., with only one proposition in a phase set), reflecting the assumption that holding multiple propositions in WM requires greater cognitive effort than thinking about only one fact at a time. If multiple propositions need to be considered in the phase set simultaneously (e.g., in cases when the mapping would be structurally ambiguous based on only a single proposition), LISA assumes that propositions are more likely to be placed together into WM if they are connected in ways that create textual coherence (Kintsch & van Dijk, 1978). In particular, propositions are more

likely to enter WM together if they are arguments of the same higher-order relation (especially a causal relation), or if they share arguments or predicates, than if they are unrelated (Hummel & Holyoak, 1997). Such factors also serve as top-down constraints on strategic processing (e.g., allocation of attention) in human text and event comprehension (Fletcher, 1986; Fletcher & Bloom, 1988; van den Broek, 1988). For example, readers allocate more attention to statements following causal antecedents than to statements following causal consequents (Fletcher, Hummel & Marsolek, 1990).

LISA's ability to discover a mapping between two analogs is strongly constrained by the order in which it fires propositions in the driver, and by which propositions it puts into WM simultaneously. This is especially true for structurally complex or semantically ambiguous analogies. Mappings discovered early in the mapping process constrain the discovery of later mappings. If the first mappings to be discovered correspond to structurally "correct" mappings, then they can aid significantly in the discovery of subsequent mappings. But if the first mappings correspond to structurally incorrect mappings, then they can lead LISA down a garden path that causes it to discover additional incorrect mappings.

Mapping performance is similarly affected by which propositions LISA puts together into WM. Consider the following example:

> Analog 1:   Analog 2:
>
> *taller-than* (*Abe Bill*)   *taller-than* (*Alice Betty*)
>
> *taller-than* (*Bill Charles*)   *taller-than* (*Betty Cathy*)

Taken in isolation, *taller-than* (*Abe Bill*) maps equally well to either *taller-than* (*Alice Betty*) or *taller-than* (*Betty Cathy*). In order to decide that *taller-than* (*Abe Bill*) maps to *taller-than* (*Alice Betty*)—and therefore that *Abe* maps to *Alice* rather than *Betty*—it is necessary to consider both statements in Analog 1 as a pair. LISA is more likely to correctly solve this

problem and similar problems if it fires the two statements together (i.e., as part of a single phase set) and then updates mapping weights rather than if it only fires one at a time, updating the mapping connections in between (see Hummel & Holyoak, 1997).

Hummel and Holyoak (1997) reported a set of simulations demonstrating that LISA captures 14 empirical phenomena related to analog retrieval, mapping, and their interrelationships. Holyoak and Hummel (2001) report additional simulations that illustrate how LISA can successfully map large analogs (summaries of World War II and the Persian Gulf War) while operating within its limited WM capacity. LISA has been shown to capture human sensitivity to structural parallels between situations (preferring to map elements that consistently fill corresponding roles), semantic similarity (preferring to map elements drawn from closely-related categories), and pragmatic pressures (preferring mappings based on knowledge relevant to the reasoner's goals).

Aim of the Paper

Our aim in the present paper is to empirically test a series of novel predictions that LISA makes about the impact of text coherence on analogical mapping. These predictions go beyond the findings that were used to develop the model. Although the full computational model is necessarily used to derive simulation results, the key predictions depend on a small number of general principles. To summarize the major principles on which we will focus: (1) Due to the capacity limits of LISA's WM, structure-mapping is necessarily an incremental process. (2) Also due to capacity limits, mapping is directional, proceeding from the driver analog to the recipient. (3) As a result of (1) and (2), mapping is sensitive to factors that affect the order in which propositions in the driver are mapped, and which propositions enter WM together. (4) Principles of text coherence determine which propositions a reasoner is likely to think about together, and in what order.

We report four experiments testing LISA's predictions concerning factors that should control the processing of analogs during mapping. The major factors are (1) the likelihood that multiple propositions are grouped together in a phase set, (2) the specific groupings that are made, and (3) the selection of which analog will serve as the driver, which may lead to asymmetries in the mapping process (i.e., driving analog A onto analog B can produce different mappings than driving B onto A). Experiments 1 and 3 use relatively artificial materials in which a unique mapping must be discovered using structural constraints alone, whereas Experiments 2 and 4 use richer and more meaningful texts. Although most of the materials (with the exception of Experiment 4) do not involve deep semantic analogies, all of them require relational reasoning for their solution, and hence provide tests of a core component of any model of analogy.

 For each experiment we will compare LISA's behavior (i.e., its performance on specific simulations) to human behavior (see "Overview of LISA Simulations" in the report of Experiment 1). Where possible, the predictions of LISA are compared with those derived from IAM (Keane et al., 1994), the model most similar to LISA in its use of incremental mapping.

Experiment 1

Experiment 1 was based on a mapping task adapted from that used by Keane (1997, Experiment 3). Keane asked participants to identify corresponding people and relations between two structurally isomorphic, four-proposition analogs. The source analog either had causal content (i.e., a meaningful theme) or else was noncausal; the target analog was always noncausal. Table 1 presents the analogs Keane used, with the propositions in each analog numbered for ease of reference later in the present paper.

Insert Table 1 about here

_____

Each proposition describes a relation between two actors. Together, the four propositions introduce three actors and three relations between them. Because the analogs are isomorphic, there is only one structurally consistent set of correspondences between them. That is, each actor (or relation) in one analog corresponds to one and only one actor (or relation) in the other analog. The analogs vary in coherence, depending on the level of causal content in each. The causal source analog contains (implicit) causal content that adds coherence to the story told by the four propositions: In addition to the information explicitly stated in each of the individual propositions, there also exists a higher-order structure based on a "love triangle" theme. The noncausal source and target analogs provide no such implicit higher-order structure.

One group of subjects mapped the causal source onto the target, and a separate group mapped the noncausal source onto the target. Keane (1997) found that causal content facilitated mapping: mappings between the source analog with causal semantic content and the target analog with noncausal semantic content were more accurate than were mappings between the source analog with noncausal content and the target with noncausal content. This result is inconsistent with several major models of analogy. For example, both ACME and SME predict that higher-order structure can only facilitate mapping when it is present in both analogs (so that the higher-order propositions in the two analogs can map to each other). But in Keane's experiment the noncausal target lacked higher-order structure, so the implicit causal structure in the "jealousy" analog had no corresponding structure to map onto in the target. In this sense, the two noncausal analogs were actually more isomorphic to one another than were the causal and noncausal analogs.

Keane (1997) explained the advantage of the causal source in terms of his IAM model, which maps incrementally starting with the largest internally coherent group of propositions (the "seed group"). When mapping with a noncausal source (in which there is no higher-order structure), there can be no seed group larger than an individual proposition, since the four propositions each contain the same degree of internal coherence; the seed group chosen with a noncausal source thus consists of a single proposition. On the other hand, when mapping with a causal source (in which there is higher-order structure), the seed group will include the higher-order causal proposition and its constituent lower-order propositions, which together provide more internal coherence than any single proposition alone. Mapping with this larger seed group produces more accurate mappings than does mapping with a single-proposition seed group because where a single proposition in source analog will map equally well to any of the four propositions in the target (thus generating very ambiguous mappings), the higher-order structure imposes more constraints on which propositions in the target analog are more likely to match those in the source analog. Although IAM does not directly predict the observed differences in human mapping accuracy as a function of causal content (as the model invariably finds the complete set of correct mappings regardless of what seed groups it starts with), its model-specific measure of backtracking to find the optimal solution ("groups complexity") could, with some modifications, be used to infer such differences in percent correct mappings.

LISA provides a different explanation of Keane's findings, and generates additional predictions, as well as a providing a more direct measure of mapping accuracy. LISA assumes that the implicit causal content serves to increase textual coherence. Specifically, the causal relations increase the probability that multiple propositions (describing the cause and the effect) will enter the phase set together, rather than being processed separately. In this manner, LISA models the increased coherence that results as a function of causal semantic content in two ways:

explicitly, by the addition of higher-order causal propositions (which when present in the target as well as the source can bootstrap the discovery of the correct element mappings), and implicitly, by which propositions are processed together in the same phase set. The grouping of multiple propositions in the phase set will often provide greater structural constraints on mapping, especially when there is no corresponding causal structure in the recipient.

While higher-order causal structure strongly promotes the firing of multiple propositions in the same phase set, this grouping of propositions can also occur without such higher-order structure: LISA predicts that *any* cue that increases optimal grouping (not just causal content) will improve mapping accuracy for problems that depend on sensitivity to structure. In an initial experiment with 107 participants serving in roughly equal numbers in each of three conditions, we compared the beneficial impact of causal content to that of a more direct cue to group together certain propositions in the source—simply placing a box around two sentences in a noncausal analog and instructing participants to "think of the sentences together." Mapping accuracy in this noncausal "box" condition was approximately equal to that observed for a replication of Keane's (1997) condition in which the source analog had causal structure (74% versus 71%, respectively), and both conditions yielded significantly greater accuracy than when neither analog was causal and no box was provided (53%).

LISA makes other predictions about the impact of groupings on mapping accuracy. Whereas some groupings (such as that used in the "box" condition discussed above) provide strong structural constraints that aid in mapping, other potential groupings do not. As we show below by simulation, in LISA the amount of structural constraint provided by firing multiple propositions in WM simultaneously is a function of the number of elements—i.e., roles and fillers—that remain constant over the set of propositions. Experiment 1 was performed to directly compare the impact of two different groupings on mapping performance. We compared

mapping performance, in LISA and human participants, between a driver with noncausal content and a target analog with noncausal content. The noncausal drivers differed in whether and which propositions were grouped together.

Method

    Participants. The participants were 149 UCLA undergraduate students who took part in the experiment either as part of a course requirement for an introductory psychology class or as part of a class demonstration for an introductory cognitive psychology class. Approximately equal numbers of participants were assigned to each experimental condition (49 in Good Grouping, 50 in Bad Grouping, and 50 in No Grouping).

    Design, materials, and procedure. The task and materials were based on those used by Keane (1997; see Table 1). The driver (the noncausal source in Table 1) contained no obvious causal semantic content. The grouping of propositions in the driver was varied across three between-subjects conditions by placing a box around one pair of propositions (in the Good and Bad Grouping conditions) or not (the No Grouping condition). In the Good Grouping condition (a replication of the "box" condition of the initial experiment discussed above), we drew a box around sentences 2 and 3 in the noncausal driver. The Bad Grouping condition was identical to Good Grouping except that sentences 3 and 4 were placed within a box instead of sentences 2 and 3. LISA predicts that the Bad Grouping will produce less accurate mappings than the Good Grouping (see Simulation Results below). In all conditions the recipient was the noncausal target analog shown in Table 1.

    In all conditions, the analogs were presented side by side on a standard sheet of paper, with the driver on the left side and the recipient on the right side. The four sentences in the driver were always presented in the order 1-4; the sentences in the recipient appeared in one of two orders (varied across participants): either 3, 4, 1, 2 or else 4, 3, 2, 1. Instructions referred to

the driver analog as "Set A" and the recipient as "Set B". In order to focus attention on Set A and establish it as the driver, it was printed in boldface and was entirely surrounded by a black box. In the Good Grouping condition, sentences 2 and 3 in Set A were set off within an additional box. In the Bad Grouping condition, sentences 3 and 4 in Set A were set off within an additional box. Set B was printed in regular font and was not surrounded by a box. Below Set A, its three objects and three relations were listed (objects first, followed by relations); under Set B appeared six blanks, which were aligned with the six elements listed below Set A. Arrows pointed from the elements under Set A to the blanks under Set B. Instructions for the No Grouping condition read as follows: "Please read the two sets of sentences below. The situations described are analogous (each person in Set A corresponds to one person in Set B; likewise, each relation in Set A corresponds to one relation in Set B). Notice that the people and relations in Set A are listed below Set A. Below Set B is a set of blank spaces. For each person or relation in Set A, please write down the corresponding person or relation from Set B. There will be a five minute time limit." Instructions for the Good Grouping and Bad Grouping conditions were identical except for the addition of one sentence: "Pay special attention to the boxed items—thinking about them together will help you solve the problem." Although participants were told there would be a five minute time limit, this was not enforced because the majority of the participants finished within the given time allotment. The instruction was included simply to motivate the participants to complete the problem within a reasonable amount of time.

The experiment was embedded within an unrelated experiment. After reaching the midpoint of the other experiment, participants were randomly handed a test sheet and told to follow the written instructions. Upon completion of the task, participants were thanked and debriefed.

Overview of LISA Simulations

A full description of the operation of LISA is provided by Hummel and Holyoak (1997). For our current purposes, it is sufficient to note that LISA is an artificial neural network that performs analogical mapping as a form of guided pattern recognition, as described previously: One (or at most three) at a time, propositions in the driver become active, generating patterns of activation on a collection of semantic units. In turn these semantic patterns drive the activation of units representing objects, relations and propositions in the recipient. For example, when the proposition *loves* (Jim, Mary) fires in the driver, it activates semantic units representing Jim (e.g., adult, human, male, banker, etc.) in synchrony with units representing the lover role (e.g., emotion, positive, strong, etc.), and it activates units representing Mary in synchrony with units representing the beloved role. (As noted previously, the units for Jim and lover fire out of synchrony with those for Mary and beloved.) In a natural analogy, the semantic patterns generated by the driver will tend to preferentially activate some propositions over others in the recipient. For example, the semantic pattern generated by *loves* (Jim, Mary) is more likely to activate *likes* (Bill, Sally) than *drives-to* (Bill, beach). The resulting patterns of coactivity across analogs (between Jim and Bill, Mary and Sally, and corresponding roles of *loves* and *likes*) bootstrap LISA's discovery of the correspondences between the analogs. Based on their coactivity, LISA updates the weights on *mapping connections* between units in the driver and units in the recipient. Thus, when *loves* (Jim, Mary) activates *likes* (Bill, Susan), LISA increments the weight on the connections from Jim to Bill, Mary to Susan, and the corresponding roles of *loves* and *likes*.

The resulting connection weights serve both to store the mappings LISA discovers, and to constrain its discovery of future mappings. For example, if LISA learns a positive mapping connection from Jim to Bill in the context Jim's loving Mary, then this connection will allow Jim

to directly activate—and therefore map to—Bill in all future contexts. It is for this reason that LISA is so sensitive to the order in which it fires propositions: If the mapping from Jim to Bill is consistent with the structure of the analogy as a whole, then LISA's discovering it early in the context of Jim's loving Mary will aid LISA in discovering other (perhaps more semantically ambiguous) mappings subsequently. In this way, finding a good mapping early can help LISA discover of a globally coherent mapping. But if the mapping of Jim to Bill is inconsistent with the analogy's global structure, then LISA's "discovering" it early on can lead the model to find additional incorrect mappings. Placing multiple propositions into WM simultaneously helps LISA to find globally consistent mappings by allowing it to consider those propositions in (quasi-) parallel (see Hummel & Holyoak, 1997). If LISA fires proposition A before it fires proposition B, then the mappings it discovers in the context of A will influence the mappings it discovers for B; if it fires B first, then B will bias what it discovers for A. But if it puts A and B into WM together, then it will map them as a pair, and neither proposition will bias the mappings discovered for the other.

An "easy" mapping—such as *loves* (Jim, Mary) to *likes* (Bill, Susan) versus *drive-to* (Bill, beach)—can be discovered on the basis of the semantic relations between the objects and predicates of individual propositions; hence it is not necessary to consider multiple propositions together. More difficult are mappings that are not directly revealed by the semantics of objects and predicates. Such are the mappings underlying the analogies used by Keane (1997). These mappings require the reasoner to consider the analogy at a much more abstract, structural level. In such cases, it matters a great deal more which propositions LISA considers together, and in what order. For example, if Jim loves Mary and Mary loves Sam (in the driver), and Bill likes Susan and Susan likes Alex (in the recipient), then it is not obvious, based on any single proposition in isolation, whether Jim (who loves Mary) corresponds to Bill (who likes Susan) or

Susan (who likes Alex). To discover that Jim corresponds to Bill, it is necessary to consider both the fact that Jim loves Mary and the fact that Mary loves Sam at the same time (Halford et al., 1994; Hummel & Holyoak, 1997).

To a model such as LISA, which is sensitive to both semantic constraints (e.g., the fact that *loves* is more like *likes* than like *drives-to*) and working-memory limitations, some problems are much harder than others: The less the semantic constraint and the greater the WM load (i.e., the greater the number of role-filler bindings it is necessary to consider together), the harder the problem. By these criteria, Keane's (1997) problems are very hard: There is no semantic similarity between the relations in one analog and those in another, and the semantic relations among the characters (e.g., who is male and who female) provide no help whatsoever. LISA is therefore exquisitely sensitive to which propositions are considered together, and in what order. The behavioral prediction is that human reasoners will be similarly sensitive.

To test this prediction, we ran LISA (and tested human subjects) in a variety of conditions with the Keane (1997) analogy problems. We varied which (if any) propositions LISA (and human subjects) were encouraged to consider together. In all the simulations reported here, mapping in different conditions is modeled solely by varying the order in which propositions are fired, and which propositions fire together (i.e., in a single phase set). The experiments thus test LISA's detailed assumptions about the nature of the incremental mapping process, holding constant the basic representations and processes used by the model. The hand-coded representations for the textual materials necessarily involve somewhat arbitrary decisions (as is universally the case for the inputs given to process models applied to complex texts). However, we tried to make the representations plausible for a hypothetical "typical" reasoner who had read the experimental texts. Examples of the specific representations are provided in the Appendices

(copies of the simulation representations and firing orders for this, and those used in the rest of the experiments, are available from the first author).

Simulation Results

An example of the representations used in LISA simulations of Experiments 1 and 3 is provided in Appendix A. The differences in predicted mapping accuracy for the three conditions were derived solely by varying the groupings of propositions that LISA fired together in a single phase set. In the No Grouping condition each proposition in the driver was fired individually. The Good Grouping condition used the same noncausal driver analog as in the No Grouping condition, but fired the second and third propositions in the same phase set, followed by the sequential firing of the remaining two propositions. Simulations for the Bad Grouping condition were identical to those for Good Grouping, except that propositions 3 and 4 (rather than 2 and 3) were fired first and together.

In LISA, the amount of structural constraint provided by firing multiple propositions in WM simultaneously is a function of the number of elements—i.e., roles and fillers—that remain constant over the two propositions.[2] The model therefore predicts an advantage for the Good Grouping condition due to the fact that propositions 2 and 3 provide strong structural constraints when considered together, as they hold constant the relation and the filler of the patient role, while varying solely the filler of the agent role. In contrast, the Bad Grouping of the third and fourth propositions provides weak constraint because these two sentences have different relations with different fillers of the object roles; only one element (the filler of the agent role) is constant. LISA predicts that focusing joint attention on propositions that fail to provide structural constraints will not facilitate mapping relative to considering each proposition separately.[3]

Because LISA's mapping algorithm has a stochastic component, 10 simulations of the mapping task were conducted for each firing order within the three conditions. For each run, a mapping between an element in the driver and one in the recipient was considered correct if the structurally-consistent mapping achieved an asymptotic connection weight greater than 0.5 (weights range from -1.0 to 1.0) and was at least 0.2 greater than any other connection from that driver element to any recipient element. For the relation mappings, both case roles had to be correctly mapped in order for the entire relation to be considered correctly mapped. The number of correct mappings out of a possible six (three people and three relations) was then converted to a percentage, allowing direct comparison between the dependent measures used in the simulations and human data. No attempt was made to fit the human data beyond the level of statistically-reliable ordinal differences, which is typical in process modeling (as opposed to mathematical modeling, where the goal is to fit individual data points, but there is no attempt to account for the algorithm underlying human performance). Mean mapping accuracy averaged over ten runs for each condition was 25% for No Grouping, 86% for Good Grouping, and 26% for Bad Grouping, confirming LISA's prediction that for human reasoners the Good Grouping condition will yield greater mapping accuracy than either the Bad Grouping or No Grouping conditions. The latter conditions are predicted to yield no difference in mapping accuracy.

Results and Discussion

The number of people and relations correctly mapped was scored for each participant. A two-way analysis of variance revealed a main effect of grouping, $F(2, 143) = 4.11$, MSE = 4.49, $p = .02$. Planned comparisons showed that mapping accuracy was significantly higher in the Good Grouping condition (60%) than in the Bad Grouping (40%) and No Grouping conditions (47%) combined (with assigned weightings of 1, -.5, -.5; $p < .01$); the latter two conditions did not differ from one another ($p > .20$). There was no significant effect of sentence order, $F(1,$

143) = 1.73, MSE = 4.49, *p* > .15, and no interaction between that variable and grouping, *F* (2, 143) < 1.

The advantage of the Good Grouping condition over No Grouping replicates the finding obtained in the initial experiment summarized in the Introduction to Experiment 1. The fact that the Bad Grouping condition failed to yield any facilitation supports LISA's prediction that the critical factor affecting mapping accuracy is joint processing of information that provides strong structural constraints. The IAM model (Keane et al., 1994) predicts a small advantage of the Good Grouping over the Bad Grouping condition (assuming that explicit representations of higher-order propositions connect boxed sentences); however, IAM predicts (incorrectly) that the Bad Grouping condition should yield more accurate mappings than the No Grouping condition.[4] In IAM, the Bad Grouping condition produces an advantage because the presence of the explicit higher-order proposition connecting the sentences results in a larger seed group, which facilitates mapping. In other words, the mere presence of a higher-order proposition facilitates mapping in IAM, regardless of the nature of the propositions grouped by that higher-order proposition.

To understand why this is so, we need to summarize how IAM operates. First, IAM selects a seed group from the base domain (which in the studies presented in this paper is the driver analog). The seed group consists of whatever group of interconnected propositions has the most systematic structure. In Experiment 1, the seed group consists of the two boxed propositions, along with a higher-order proposition linking them together. Once the seed group is chosen, IAM determines the seed match and uses it as a starting point for the rest of the mappings. To do so, one element of the seed group is chosen, and all legal matches to that element are found; among these, that which best meets a combination of pragmatic, semantic, and structural constraints is chosen as the seed match. Based on this seed match, IAM finds all other legal matches to the remaining elements in the seed group (operating under the same

criteria, pragmatic, semantic, and structural, while taking into account previous mappings). After completing the mapping of the seed group, IAM evaluates the quality of the mapping. Generally speaking, if more than half of the predicates in the seed group have been matched successfully, IAM performs no more mappings (of the seed group); if less than half are matched successfully, IAM backtracks and considers other sets of mappings based on a different seed match. As mentioned previously, IAM invariably finds the correct mappings; thus the advantages of one group over another are expressed in terms of the degree of backtracking necessary to solve the analogy.

With tasks such as that used in Experiment 1, after IAM maps the seed group to satisfaction, it completes the mappings for the rest of the analog; in most situations, IAM does not map beyond the seed group, as the seed group usually provides sufficient information about the analogy, as well as providing the most interesting inferences. (Note that this description of IAM excludes its inference/transfer function, which is not relevant to these simulations.) With the materials used in Experiment 1, which lack semantic overlap to help constrain the mapping process, there are many possible matches for the rest of the elements in the seed group after a seed match is determined. Thus, in order to determine the correct mappings, IAM must backtrack. When first determining the correct mappings with the particular seed group (either propositions 2 and 3 and a higher-order proposition joining them in Good Grouping, or propositions 3 and 4 and a higher-order proposition joining them in Bad Grouping), IAM's backtracking continues until a set of mappings is found that satisfies its evaluation criteria. With the seed group consisting of two basic propositions to be mapped, there is substantially less backtracking necessary than when mapping with a single-proposition seed group (as in the No Grouping condition). In the No Grouping condition, IAM's backtracking is not limited to a two-proposition seed group (plus a third, higher-order proposition), but instead has to accommodate

all legal matches for all four propositions in the analog. Thus, the presence of higher-order structure, regardless of the relational content of the lower-order propositions, facilitates mapping due to the nature of IAM's backtracking strategy.

Experiment 2

The materials used in Experiment 1 are semantically impoverished and hence not representative of realistic analogies (Gentner & Markman, 1997). These mapping problems hinge primarily on structural constraints (role correspondences), with minimal support from semantic similarity of either relations or objects. In Experiment 2 we sought to broaden the domain in which we tested the implications of LISA's account of analogical mapping by using meaningful story problems. Accordingly, we constructed materials based on analogs that shared semantic similarity and that were presented in a more naturalistic context (summaries of movie plots), in order to expand upon the findings of Experiment 1.

As in Experiment 1, Experiment 2 varied the nature of the propositions that were grouped together. The grouped propositions either provided high structural constraint or else low structural constraint. When the grouped propositions provided high structural constraint, they contained relational information that would help to disambiguate the possible correct mappings, thus increasing mapping accuracy. Groupings that provided low structural constraint did not adequately disambiguate the possible mapping, and thus should fail to facilitate mapping. The LISA model therefore predicts that mapping accuracy when there is high constraint should be greater than when there is low constraint.

In addition, this difference in constraint should also produce a difference in the degree of transfer between driver and recipient analogs. When mappings are more accurate, inferences should be more accurate as well. Given that mapping accuracy will be greater from a driver

providing high constraint than a driver providing low constraint, inferences should also be more accurate when the driver provides high constraint than when it provides low constraint. Thus the goals of Experiment 2 were to replicate the findings of Experiment 1 with more naturalistic materials, and to show that in addition to facilitating mapping, higher levels of constraint also facilitate inference.

Method

Participants. The participants were 64 UCLA undergraduates who completed the experiment to fulfill the requirements of an introductory psychology course. The participants were equally and randomly distributed across the experimental conditions.

Design, materials and procedure. In order to examine the possible differences in mapping accuracy between grouped propositions that provide high constraint and those that provide low constraint, we varied the grouping of four focal propositions that were embedded within similar stories (see Appendix B for one example of the stories used as driver and recipient analogs). The focal propositions appearing in the stories used in Experiment 2 are given in Table 2. The sets of propositions are isomorphic, as are those used in Experiment 1, but for the materials used in Experiment 2 the isomorphism is dependent on the shared semantic similarity of predicates (i.e., "waves to" is more similar to "winks at" than to "gives" and "throws" is more similar to "gives" than to "winks at"). Because of this dependency on shared semantic similarity, different groups of propositions can provide different amounts of structural constraint. The High Constraint condition grouped together propositions 1 and 2, and 3 and 4. This manner of grouping provides high structural constraint because the relational information contained in the groups, coupled with the semantic similarity of the predicates, allows the correct mappings to be determined. For instance, "waves at" in proposition 1 of the Bank Context encourages mapping to propositions 1 and 2 of the Spy Context because of its semantic similarity to "winks at" (in the

case where the Bank Context is serving as the driver and the Spy Context as the recipient);

however, if proposition 1 were fired alone, it would not be clear whether it mapped to

proposition 1 or 2 in the Spy Context (i.e., taken alone, proposition 1 in the Bank Context maps

equally well to either proposition 1 or 2 in the Spy Context). By grouping together propositions

1 and 2 in the Bank Context, the relational information shared by these two propositions can

disambiguate the mappings. The grouping of propositions 3 and 4, along with the semantic

similarity between "throws" and "gives", also allows a disambiguation of mappings. In contrast,

the Low Constraint condition grouped propositions 1 and 3, and 2 and 4. Despite the semantic

similarity of predicates, these groupings do not provide a high level of structural constraint. As

with the High Constraint condition, while semantic similarity promotes grouping proposition 1 of

the Bank Context to either proposition 1 or 2 of the Spy Context (as opposed to propositions 3 or

4), it remains unclear whether "Danny waves to Lillian" maps to "Sharon winks at Rich" or to

"Rich winks at Julia". In contrast to the High Context condition, the grouping of proposition 1

with proposition 3 in the Low Context condition does not provide additional constraint for the

mapping problem, mainly because proposition 3 in the Bank Context maps ambiguously to either

proposition 3 or 4 in the Spy Context. Grouping propositions 2 and 4 similarly does not provide

sufficient structural constraint to determine the correct mappings. The grouping of propositions

in the stories was encouraged by providing a causal explanation for why the given two

propositions occurred together in time. The groups were separated both by these causal

explanations, as well as by a temporal gap in the storyline when the propositions appeared.

Insert Table 2 about here

There were two main contexts, Bank and Spy, which each provide an account of the main character's viewing of a film. As they read each story, participants in Experiment 2 were instructed to imagine that they were playing the role of this main character. The four focal propositions were presented as events that occurred within the film viewed by the main character. In constructing the materials great care was taken to ensure that the stories differed only in the level of structural constraint provided by the grouped focal propositions, and not in their global level of clarity or ease of comprehension. To assess the comprehensibility of the stories, 128 additional UCLA undergraduates rated each story along three dimensions of clarity (plausibility, meaningfulness, and coherence) as part of a requirement for an introductory psychology course. Stories were rated on a 7-point Likert scale, with low scores indicating a lack of clarity and high scores indicating a high degree of clarity. Plausibility was defined as "how realistic the story is, given that it is supposed to take place in 'the real world'"; meaningfulness was defined as "how easy the story is to understand"; coherence was defined as "how well the pieces of the story fit together." High Constraint stories did not differ significantly from Low Constraint stories in their rated plausibility ($M = 5.72$ and $M = 5.62$, respectively, $F < 1$), meaningfulness ($M = 5.53$ and $M = 5.48$, respectively, $F < 1$), or coherence ($M = 5.42$ and $M = 5.09$, respectively, $F(1, 124)$ $= 1.71$, MSE $= 2.02$, p $> .05$). The Bank context did not differ significantly from the Spy context in rated plausbility ($M = 5.46$ and $M = 5.87$, respectively, $F(1, 124) = 2.57$, MSE $= 2.06$, p $> .05$), meaningfulness ($M = 5.59$ and $M = 5.42$, respectively, $F < 1$), or coherence ($M = 5.22$ and $M = 5.30$, respectively, $F < 1$). The stories thus did not differ in their overall level of clarity, across either level of constraint (High or Low) or context (Bank or Spy).

The experimental task, which assessed the accuracy of both mapping and inference between a driver and recipient, was adapted from that used by Spellman and Holyoak (1996). Participants were told to role-play being called to court by a judge, and that they were to

determine whether one movie company was stealing storylines from another movie company. The driver analog contained information about the original movie, and the recipient contained information about the potentially plagiarized movie. The analogs were presented in the following manner: Participants were first introduced to the cover story, where they are told to imagine they had seen the original movie. They then read the driver analog, which contained an account of their viewing of the movie. Because the participants were asked to imagine that they had seen the original movie (and thus had the requisite knowledge to participate in the court decision about the plagiarism case), the driver analogs were all written in the second person, presenting an account of the movie itself and the circumstances surrounding their viewing of the movie. After reading the driver analog, they were then asked questions about the content of the movie (i.e., the focal propositions) described in the analog. These questions asked about the objects and predicates in the focal propositions, and also required the participants to state why the pairs of propositions occurred together (i.e., what the higher-order causal relationship between the two propositions was, or why/how the movie's plotline required these two propositions to occur together). After giving their answers, participants received the correct answers, and were told to verify that the answers were in fact correct by reading the account again. Both the questions and feedback were intended to increase the participants' ability to recall the stories during the mapping task. Participants then read the recipient analog, which contained a brief description of the potentially plagiarized movie's plot consisting of the four focal propositions. The recipient analog did not contain other information about the movie, and merely presented some events (i.e., the four focal propositions) occurring within it. The participants' task was to predict the order in which the events described by the four statements would take place, using the original (driver) movie plot as a guide. We chose this task to make the participants' role-playing task of determining plagiarism more plausible; if they can predict the correct order of the events within

the potentially plagiarized movie without additional contextual knowledge surrounding the events (as was given in the driver analog), then the judge would have strong evidence to assume that the movie in the recipient analog was plagiarized from the movie in the driver analog.

The focal propositions in the recipient analog were presented in one of two orders, chosen to minimize superficial correspondences between the presentation orders of the structurally-parallel focal propositions in the driver and recipient analogs (i.e., to avoid placing structurally-parallel propositions in the same ordinal position in the two analogs). Participants answered questions about the focal propositions and then received feedback in a similar manner as with the driver analog. The only difference was that the questions about the recipient only asked about the content of the focal propositions. Because the focal propositions in the recipient were not grouped together, there were no questions about the higher-order relations between propositions.

Thus, a hypothetical participant would first read the driver analog, "High Constraint Driver, Spy Context," which describes how the participant (i.e., the second-person character in the story) came to view a late-night movie, and provides details about parts of the movie that the participant viewed where the main characters (Sharon, Rich, and Julia; see Table 2) are stealing a small box from a mansion. The focal propositions were introduced in the description of the process by which the characters steal the box. The driver analog then ends with a description of the participant going to bed after becoming too sleepy to continue watching. (See Appendix B for an example driver analog.) After reading the driver analog, the participant would then answer questions about the focal propositions appearing in that driver analog. After receiving the answers to the questions and re-reading the driver, the participant would be presented with the recipient, "Recipient Order 1, Bank Context." "Recipient Order 1, Bank Context" provides a brief plot outline of the movie, which involves a bank robbery, along with the names of the characters involved (Danny, Lillian, and Steve; see Table 2), and the four focal propositions

describing events in the movie. Also see Appendix B for an example recipient analog. The participant would then answer questions about the focal propositions appearing in the recipient analog, and re-read it. Thus, the participants viewed the focal propositions solely within the contexts of the driver and recipient analogs.

After the presentation of the analogs, the participants completed the inference task. They were told that the judge presiding over the plagiarism case wanted them to try to predict the correct order of events in the movie in the recipient, which events took place together, and why the events took place together, based on the content of the movie in the driver analog. The analog containing the original movie was presented again, to emphasize its role as driver. This task was designed to motivate participants to take the inference task seriously, and to elicit a high level of effort in the determination of correspondences between the driver and recipient analogs. Finally, participants were asked to determine which characters and relations in the recipient analog corresponded to, or "matched", those in the driver analog. In the prior inference task, the participants were not explicitly asked to generate the mappings between elements of the analogs, although they were inherently forced to do so in order to generate the correct inferences about the order of propositions in the recipient. The three characters and two predicates of the driver were presented on the left side of the page, and participants filled in blanks on the right side of the page. They also provided "confidence ratings" of how good they thought the match was between each element of the driver and the chosen matching element of the recipient. These ratings were made on a scale from 1-5 (with "1" being "slight match" and "5" being "excellent match"). These ratings were intended to motivate the participants to take the mapping task seriously, and were not actually scored.

There were a total of 16 experimental conditions generated from the 2x2x2x2 between-subjects design. The main independent variable was the level of constraint provided by the

grouped focal propositions in the driver (Constraint: High Constraint or Low Constraint). The other three variables were designed to counterbalance and/or control for specific item effects. Which story context served as driver and which served as recipient was varied (Story Type: Bank or Spy), as was the order of the focal propositions in the recipient analog (Recipient Order: Recipient-Order 1 and Recipient-Order 2), and the random order of the mapping questions (Question Order: Question-Order 1 and Question-Order 2).

Participants were randomly assigned to conditions, and completed the experiment either individually or in groups of 2-4. They were instructed that they were to complete the experimental materials in order, and that they were not allowed to look back at previous pages once they were completed. There was no restriction on the amount of time to complete the task, but most participants finished within 20 minutes, at which time they were thanked and debriefed.

Simulation Results

Simulations based on the four focal propositions were run in LISA. The representations included shared semantics for the predicates, as well as across the objects of the analogs (see Appendix C for example representations used for the Bank Context Analogs). To simulate the High Constraint condition, the proposition linking the first and second propositions was fired first, followed by the first and second propositions being fired together, then the proposition linking the third and fourth propositions, followed lastly by the third and fourth propositions, which were also fired together. To simulate the Low Constraint condition, the proposition linking the first and third propositions together was fired first, followed by the first and third propositions fired together, then the proposition linking the second and fourth propositions, and lastly the second and fourth propositions fired together. A mapping between elements of the driver and recipient analog was considered correct based on the same criteria as in Experiment 1.

Based on 20 runs for each condition, mapping accuracy was higher in the High Constraint condition ($M = 100\%$) than in the Low Constraint condition ($M = 32\%$).

Results and Discussion

   Mapping accuracy.   A 2x2x2x2 analysis of variance revealed no main effects of Story Type, Recipient Order, or Question Order (all $F < 1$).   However, a significant main effect of Constraint was obtained, $F(1, 48) = 6.42$, MSE $= 2.34$, p $< .02$, with greater percent mapping accuracy in the High Constraint condition ($M = 81$) than in the Low Constraint condition ($M = 62$).   Participants were thus more accurate when mapping from a driver with grouped propositions that provided high structural constraint than when mapping from a driver with grouped propositions that provided low structural constraint.   The pattern of mapping results found in Experiment 1 was thus extended in Experiment 2 to materials that are much more naturalistic, and that contain overlapping semantic content. Previous models (e.g., SME, ACME, IAM) do not predict this difference in mapping accuracy between High and Low Constraint conditions.   The LISA model can account for these findings because its directional, limited-capacity mapping algorithm is affected not only by the presence or absence of higher-order structure in the driver analog (as is IAM, but not SME or ACME), but also by the degree of structural constraint that the propositions grouped within the higher-order structure provide to the mapping process.   While IAM is affected by the presence or absence of higher-order structure in the driver analog (unlike SME and ACME), the degree of structural constraint provided by the propositions grouped by the higher-order structure does not influence mapping accuracy. Accordingly, IAM simulations show equal performance for both the High Constraint and Low Constraint conditions (see Footnote 4).

   Inference accuracy.   Inference accuracy was scored for Proposition Order (whether or not participants were able to predict the correct order of the propositions in the recipient analog,

based on the order of the propositions in the driver analog) and Causal-Relation Correspondence (whether or not participants were able to correctly infer why two propositions were grouped together in the recipient analog, based on the reasons that the corresponding propositions were grouped together in the driver analog). A proposition was considered correct only if it both contained the exact corresponding object/predicate relation and also was stated to occur in the corresponding temporal position that it took place in the driver analog. For example, if the participant's first predicted proposition did not contain the correct corresponding object/predicate as the first proposition from the driver analog, it was automatically considered incorrect. Likewise, if a predicted proposition contained the correct correspondences, yet appeared in the wrong temporal position (e.g., was wrongly predicted to occur second, instead of in the actual corresponding temporal position of, say, first), it was also considered incorrect. Since there were four focal propositions, Proposition Order had a maximum score of 4. A 2x2x2 analysis of variance revealed a significant main effect of Constraint, $F(1, 56) = 9.75$, MSE $= 2.82$, $p < .01$, with High Constraint drivers generating more accurate Proposition Order in the recipient analog than did Low Constraint drivers ($M = 2.67$ and $M = 1.36$, respectively). There was no significant effect of Story Type or Recipient Order (both $F < 1$). (Question Order was not included in the analysis of inference accuracy because that manipulation occurred after the inference task.)

Causal-Relation Correspondence had a maximum value of 2, and was based on the degree of correspondence between the inferences generated from the driver analog and extended to the recipient analog concerning why the groups of propositions were related to each other. Two inferences were required, one to explain how the first two propositions were related and one to explain how the second two propositions were related. The degree of correspondence of each inference ranged from 0 to 1 (0, .25, .5, .75 or 1), and was scored by two independent raters. The scores given by the two raters did not differ significantly (t (63) $= 1.18$, $p > .05$ for the first

inference; $t(63) = .85$, $p > .05$ for the second inference). In addition, the scores of the two raters agreed more than would be expected by chance for both the first and second inferences, Cohen's kappa = .54 and .45, respectively, $p < .001$. A subsequent 2x2x2 analysis of variance used the average of the scores given by the two raters. The analysis revealed a significant difference in Causal-Relation Correspondence as a function of Constraint, $F(1,56) = 13.66$, MSE = 0.51, $p < .001$, with inferences being more accurate (i.e., showing a greater degree of correspondence) with High Constraint drivers than with Low Constraint drivers ($M = 1.07$ and $M = 0.41$, respectively). There were no significant effects of Story Type or Recipient Order, both $F < 1$. The overall pattern of transfer thus showed that inferences, as well as mappings, were more accurate from a driver that provided high structural constraint than from one that provided low structural constraint. The results of Experiment 2 thus provide further support for LISA's predictions.

Relationship between mapping accuracy and inference accuracy. We examined the relationship between Mapping Accuracy and Inference Accuracy by performing two correlational analyses. We found strong positive correlations between Mapping Accuracy and both Proposition Order ($r = .76$) and Causal-Relation Correspondence ($r = .71$) (both $p < .01$). As predicted, greater mapping accuracy was associated with greater inference accuracy.

Experiment 3

LISA's capacity-limited mapping process is inherently directional, in that the driver serves as the attentional focus, generating patterns of activation to which the recipient responds. Although analogs may switch between the driver and recipient roles in the course of reasoning, we assume that manipulations that control attentional focus (such as highlighting one analog) can at least partially determine which analog acts as driver. Because of the inherent asymmetry between the driver and recipient roles, LISA predicts that mapping accuracy can be affected by the choice of which analog serves as driver: Given analogs A and B, mapping A onto B may

yield more accurate mappings than mapping B onto A. This kind of asymmetry is particularly likely to the extent that one analog is better understood than the other (which is usually the case: the source is typically better understood than the target), because a well-understood analog is more likely to contain higher-order structure that constrains the order in which propositions are chosen to fire than is a poorly-understood analog.

Experiment 3 was performed to test this prediction using materials based on those employed by Keane (1997) and in Experiment 1. When one analog is causal and the other is not, causal connections can lead to joint processing of structurally-constraining propositions in the causal analog when it serves as driver. However, the recipient analog always responds in parallel to the activation originating from the driver, so there is no comparable basis for grouping effects in the recipient. It follows that when a causal and a noncausal analog are mapped, mapping accuracy should be greater if the causal analog is the driver and the noncausal analog is the recipient, rather than the reverse (see Simulation Results below). This prediction was tested in Experiment 3.

There have been many demonstrations of asymmetries in similarity judgments (Rosch, 1975; Tversky & Gati, 1978) and metaphor interpretation (Glucksberg & Keysar, 1990; Ortony, 1979), as well as in analogical inference (e.g., Bassok & Holyoak, 1989; Bassok & Olseth, 1995; Burns, 1996; Reed, Ernst, & Banerji, 1974). Most such asymmetries have been obtained when one case or analog provides more information than the other (Bowdle & Gentner, 1997). (In analogical reasoning, more inferences can be drawn from a rich source to an impoverished target than vice versa.) However, all the previous demonstrations of asymmetries in analogical reasoning have used as dependent measures either transfer of problem solutions, generation of inferences, or preferences as to which analog should be the source (see Bowdle & Gentner, 1997, for a review). That is, previous demonstrations of asymmetries can be interpreted as occurring

during post-mapping processes such as analogical inference. It has not been clearly established that asymmetries can arise in the process of finding correspondences between two analogs. In fact, virtually all computational models of analogy, including ACME and SME, predict that the mapping process is inherently symmetrical. Only LISA (see Simulation Results below) and IAM (Keane et al., 1994) predict that the mapping process, itself, can be asymmetrical.

LISA also predicts that the presence of causal content in the recipient analog will only aid mapping when the driver analog also has causal content. When the driver has causal content, causal content in the recipient should facilitate mapping because the higher-order structure in the driver will map to the higher-order structure in the recipient. Once these higher-order propositions map, they will aid in constraining lower-order mappings (Hummel & Holyoak, 1997). When the driver is noncausal, causal information in the recipient will convey no advantage, as there will be no corresponding higher-order propositions to map to in the driver, and groupings are not determined by the recipient. Thus, mapping from a causal driver to a causal recipient (Causal-Causal) should be more accurate than mapping from a causal driver to a noncausal recipient (Causal-Noncausal), whereas mapping from a noncausal driver to a causal recipient (Noncausal-Causal) should not differ in accuracy from mapping from a noncausal driver to a noncausal recipient (Noncausal-Noncausal) (see Simulation Results and Figure 1A).

Method

Participants. A total of 497 UCLA undergraduate students participated in the experiment. Twenty completed the study as a course requirement of a cognitive psychology course; the rest completed it as part of a course requirement for an introductory psychology course. Thirty-one participants completed the experimental materials in a laboratory setting among other unrelated materials, 318 completed it as part of a class demonstration, and 128 completed it in a pretesting session along with other unrelated surveys during class. The central

predictions involved possible mapping asymmetries between the Causal-Noncausal and Noncausal-Causal conditions, so the bulk of the participants were assigned to these two critical conditions. Participants were assigned to condition by a scheme for stratified random assignment, so that approximately 80% were randomly assigned to the two asymmetrical conditions (203 in the Causal-Noncausal condition and 205 in the Noncausal-Causal condition). An additional 46 participants were assigned to the Causal-Causal condition, and 43 to the Noncausal-Noncausal condition.

Design, materials and procedure. Experiment 3 incorporated the same mapping task as Experiment 1, with slight modifications. Four analogs were used: the causal source and noncausal target presented in Table 1, plus an additional causal and noncausal analog:

Causal Analog	Noncausal Analog
Gina punches Abe.	Tom approaches Eddie.
Gina hates Abe.	Tom confuses Eddie.
Sally hates Abe.	Frank confuses Eddie.
Sally becomes friends with Gina.	Frank saves Tom.

The level of causal content was varied factorially in both the driver analog and the recipient analog to create four conditions: causal driver with causal recipient (Causal-Causal), causal driver with noncausal recipient (Causal-Noncausal), noncausal driver with causal recipient (Noncausal-Causal), and noncausal driver with noncausal recipient (Noncausal-Noncausal). Participants completed a two-page booklet. The first page presented the driver analog, along with two questions. Instructions on the first page directed the participant to read the story analog and answer the two questions. These questions were introduced into the mapping task because, unlike in the previous two experiments, the directionality of mapping was not only an assumption, but a critical independent variable. The questions were designed to both highlight

the causal structure in the causal driver and the noncausal content in the noncausal driver, and also to more firmly establish the direction in which the mappings were to be made.

When the driver had causal content, answering the questions required the participant to think about that causal content and to integrate the information presented in the analog. For the "Jim loves Mary" analog, the questions were "Why did Jim kiss Mary?" and "Why was Bill jealous of Jim?" For the "Gina hits Abe" analog they were "Why did Gina punch Abe?" and "Why did Sally become friends with Gina?" In contrast, the questions for the noncausal analogs did not require integration, instead asking for information that could be obtained from a single proposition. The questions for the "Ruth knows Debra" analog were "To whom does Laura wave?" and "Who does Ruth know?" For the "Tom approaches Eddie" analog, the questions were "Who does Frank save?" and "Who does Tom approach?"

The second page presented the mapping task, which was identical in instructions and form to those used in the first two experiments. Assignment of stories to conditions was counterbalanced as follows. For the Causal-Causal and Noncausal-Noncausal conditions, both possible driver assignments were used. For the asymmetrical conditions, Causal-Noncausal and Noncausal-Causal, all possible permutations of the four stories (given their driver and recipient assignments) were used. This design resulted in a total of 12 analog combinations across all driver/recipient combinations.

Simulation Results

The four conditions based on causal versus noncausal status of the driver and recipient were simulated using LISA (see Appendix A for an example of a causal analog). The representations of the two causal analogs included a higher-order proposition stating that propositions 2 and 3 were causally-related (e.g., proposition 2, "Jim loves Mary," and proposition 3, "Bill loves Mary," caused proposition 4, which stated the "jealousy" conflict). Of the several

potential higher-order relations, this one was chosen because it connects the two propositions that jointly provide the greatest structural constraint. In all LISA runs with a causal driver, the causal proposition was fired first, followed by the second and third propositions (grouped together in the phase set due to their causal interconnection); then the remaining two propositions were fired individually.

Multiple firing orders were used with each driver in order to approximate the possible processing strategies of human reasoners. Mean mapping accuracy over 20 runs per firing order, averaged by condition, is depicted in Figure 1A. LISA's mapping accuracy was greater in the Causal-Noncausal condition (80%) than in Noncausal-Causal (20%), generating a predicted asymmetry in human performance. In addition, mapping accuracy in Causal-Causal (100%) was greater than in Causal-Noncausal, whereas mapping accuracy in Noncausal-Noncausal (25%) did not differ greatly from that in Noncausal-Causal. LISA thus predicts (as discussed previously) that causal content in the recipient will only facilitate mapping when there is also causal content in the driver.

Results and Discussion

Human mean mapping accuracy in the four conditions is depicted in Figure 1B. An analysis of variance tested the prediction of asymmetry in mapping accuracy. The three variables examined were assignment of a causal analog to driver versus recipient (i.e., Causal-Noncausal versus Noncausal-Causal), and two variables used to control for variance attributable to the individual stories. Mapping was significantly more accurate in the Causal-Noncausal condition (54%) than in the Noncausal-Causal condition (47%), F (1, 400) = 4.07, MSE = 4.72, $p < .05$. The two causal stories did not differ significantly, but one of the noncausal analogs ("Tom approaches Eddie") was associated with more accurate mappings than the other ("Ruth knows

Debra"), F (1, 400) = 18.61, MSE = 4.72, $p < .01$. No significant two- or three-way interactions were found.

———————————————

Insert Figure 1 about here

———————————————

An additional analysis of variance was performed to determine whether the presence of causal content in the recipient facilitated mapping. The variables of interest were type of driver (causal or noncausal) and type of recipient (causal or noncausal).   The ANOVA showed significant main effects of both driver type, F (1, 493) = 12.91, MSE = 4.99, $p < .001$, and recipient type, F (1, 493) = 3.92, MSE = 4.99, $p < .05$, which were moderated by a significant two-way interaction, $F(1, 493) = 6.19$, MSE = 4.99, $p < .02$. Tests of simple main effects showed that when the driver was causal, causal content in the recipient analog yielded greater mapping performance (74%) than when the recipient had noncausal content (54%), F (1, 493) = 10.25, MSE = 4.99, $p < .002$. The advantage of the Causal-Causal condition, according to the LISA model, is that in this condition mapping is facilitated both by grouping of causally-related propositions in the driver and by mapping of higher-order propositions between the driver and the recipient. It could be argued that the two causal analogs also are relatively similar at the level of predicate semantics (since "love" and "hate", two mapped predicates for the causal stories, are opposites). In contrast, when the driver was noncausal, causal content in the recipient did not affect mapping accuracy (47% mapping accuracy with a causal recipient versus 49% mapping accuracy with a noncausal recipient). Thus as predicted, the presence of causal content in the recipient only facilitated mapping when the driver analog also had causal content.

LISA's predictions were supported by the data of Experiment 3. The central finding is that human analogical mapping performance can be asymmetrical when the roles of driver and

recipient are manipulated between analogs with causal versus noncausal content. Mapping was more accurate when the driver had causal content and the recipient did not, than when the recipient had causal content and the driver did not. The results also supported LISA's prediction that causal content in the recipient will facilitate mapping only when the driver is also causal, and its prediction that mapping will be most accurate when both the driver and recipient have causal content.

Some of the differences in the human data appear smaller than those predicted by LISA (see Figure 1). In particular, the asymmetry between Causal-Noncausal and Noncausal-Causal conditions was very small (a 7% difference in mapping accuracy) in the human data. As noted earlier, LISA's predictions are purely ordinal. One of the most important reasons we cannot make point predictions with LISA is that there are inherent differences LISA's performance and that of humans. In particular, although direction of mapping is easily controlled in LISA, we have no guarantee that humans will always map in the direction that we tell them to. More generally, LISA is a "pure" analogy machine, with no additional strategies at its disposal, and nothing on its mind but the task at hand. As such, human subjects may out-perform LISA by adopting various strategies that lie outside the domain of "normal" analogical mapping (such as counting how many times a person or relation is mentioned in an analog; see Hummel & Holyoak, 1997), or they may under-perform LISA by getting distracted or bored, or by simply being unmotivated to perform the task. If some participants in the Causal-Noncausal condition mapped in the direction opposite to that instructed (i.e., if they happened to map from the noncausal analog to the causal analog, rather than from the causal analog to the noncausal analog), then their mapping accuracy may have been be reduced. Conversely, if some participants in the Noncausal-Causal condition mapped in the direction opposite to that instructed, then their accuracy may have been increased. In contrast, violation of the directional

instructions would be expected to have no impact on performance in the symmetrical conditions (Causal-Causal and Noncausal-Noncausal). Such potential variations in performance across conditions due to violation of instructions are beyond the scope of LISA. Nonetheless, replication of the small asymmetry effect obtained in Experiment 3 would clearly be desirable. Experiment 4 was performed to assess whether similar asymmetries can be identified with semantically-richer analogs.

Experiment 4

The goal of Experiment 4 was to examine the influence of causal coherence on mapping using semantically richer, more natural materials. Experiment 3 showed that richer causal content in a driver can facilitate mapping. The distinction between causal and noncausal analogs can be viewed as an extension of the typical distinction between a well-understood source and a poorly-understood target. The analogs used in Experiment 4 were Duncker's (1945) radiation problem (the target problem to be solved) and an analogous story about a general wishing to capture a fortress (the source). These problems were first used to investigate analogical problem solving by Gick and Holyoak (1980, 1983). The radiation problem describes a situation in which a doctor is faced with a patient with an inoperable tumor, and must find a way to destroy the tumor with a special ray without damaging the surrounding tissue. The general story describes a situation that is superficially different, yet structurally similar, in which a general must overthrow an evil dictator situated in a fortress by using the entire force of his army while simultaneously avoiding discharging land mines planted in the roads that lead to the fortress (see Gick & Holyoak, 1983, p. 3 and pp. 35-36 for the full text of the radiation problem and general story, respectively). The analogs are structurally similar because the protagonists in each (the doctor and general) share similar goals and face similar constraints. Each protagonist must apply a force at full magnitude, yet cannot do so with one large force. In their original form, the radiation

problem is presented without a solution, whereas the general story is presented with a solution. Specifically, it describes how the general used small groups of soldiers to simultaneously converge on the fortress from different directions. In so doing, he was able to apply the full force of his army without violating the constraint that no single path can tolerate the entirety of his army (the "convergence" solution). The general story thus contains more causal information than does the radiation problem.

In the studies of Gick and Holyoak (1980, 1983) the radiation problem was always unsolved and the general story always solved. These assignments reflected the fact that the materials were originally created so that the general story would offer a well-understood solution to the poorly-understood and unsolved radiation problem. However, it may be possible to distinguish two possible determinants of causal coherence that might impact optimal driver assignment. First, whichever analog is presented with a solution will tend to be more coherent, as the presence of a solution will make it easier to identify the most important propositions in the problem statement (i.e., those that are causally relevant to solution attainment). Second, some analogs may have greater internal coherence in the problem statement itself. The general story would appear to be richer in its internal causal structure than the radiation problem, even when the solution to the fortress story is omitted. The critical insight required to achieve the convergence solution is that multiple forces can be used. Whereas the problem statement for the radiation problem does not provide any hint that multiple ray sources could be used, the statement of the general story mentions that multiple roads lead into the fortress, thus providing a clue as to how the general might capture it. This clue makes the unsolved general story much easier to solve than the unsolved radiation problem.

We confirmed this difference in ease of solution by giving either the unsolved general story or the unsolved radiation problem to 318 UCLA undergraduates (tested in a pretesting

session in an introductory psychology class) and asking them to provide solutions (without an analogy). One hundred sixty-three students received the general problem and 155 received the radiation story. As expected, the frequency of generating the convergence solution was far higher for the general problem (45%) than for the radiation problem (1%), X^2 (1) = 83.4, $p <$ .0001.

Although Experiment 4 examined issues similar to those considered in Experiment 3, different factors were varied, generating somewhat different predictions. In Experiment 4 we factorially varied whether the general analog or the radiation analog was presented with a solution, and which analog was the driver for the mapping task. Given the overall coherence difference between the two analogs, it was expected that mapping accuracy would be greater overall when the general analog is the driver. When the radiation analog is the driver, providing it with the solution should increase its internal coherence and hence mapping accuracy. When the general analog is the driver, providing it with a solution should be less important, as participants are likely to generate the convergence solution themselves. As in Experiment 2, LISA also predicts that if mapping is impaired by having an impoverished target as the driver, then subsequent analogical transfer (which is based on the computed mappings) may also be impaired.

Method

Participants. 279 UCLA undergraduates completed the experiment during a pretesting session in an introductory psychology program. Seventy-four participants received an unsolved radiation analog as driver, 53 received a solved radiation analog as driver, 79 received an unsolved general analog as driver, and 73 received a solved general analog as driver.

Design, materials and procedure. The 2x2x2x2 design included 16 between-subjects conditions. The original general story was modified to create a version in which the solution was

omitted, and the original radiation problem was modified to create a version in which the convergence solution was stated. The four factors varied were which analog was the driver (general or radiation), which analog was presented with a solution, which was presented first, and which of two orders of the mapping questions was used. The first two factors are central to the predictions in LISA; the latter two factors were used for counterbalancing. Each subject saw one solved and one unsolved analog. Table 3 shows the assignment of specific analogs to roles (driver versus recipient) in the experimental design.

_____

Insert Table 3 about here

_____

The task was similar to that of Experiment 3, with some important differences. The two analogs were first presented, followed by the mapping task, and finally the inference task. The first page of the 4-page packet presented either the driver or recipient analog, with instructions to read the story carefully and to try to remember as much of it as possible. The second page presented the remaining analog along with instructions to read and remember the story. The third page presented the mapping task, which required the participant to specify correspondences to six elements of the driver analog. These elements consisted of four objects and two relations. The four objects referred to the protagonist, antagonist, force, and the "innocent" object which must be protected (i.e., the doctor, tumor, ray, and healthy tissue in the radiation problem, and the general, fortress, army, and surrounding villages in the general story). The two relations referred to the desire of the protagonist and the damaging of the "innocent" object (in the radiation problem, wanting to destroy the tumor, and destroying the healthy tissue; in the general story, wanting to capture the fortress, and damaging the surrounding villages).

The participants were given instructions that mentioned that the two stories were analogous, in a manner that emphasized the direction of mapping (i.e., from driver to recipient). For example, if the first story was the radiation analog and the intended driver, then the

instructions pointed out that the elements in the first story corresponded to elements in the second story. Likewise, the mapping questions were of the form: "The doctor in the first story is like the _____ in the second story." The elements that were provided were always from the driver, with the blanks to be filled with the appropriate elements from the recipient. Since the order of presentation of the driver and recipient analogs was counterbalanced, the mapping questions served to determine the direction of mapping.

The fourth page asked for solutions to the unsolved problem. Instructions again varied according to which story was to be solved, and in which order it had been presented. For example, if the radiation analog had been presented first, with the solution, then the inference question would ask the participant to provide a solution for the second story (the unsolved general analog), based on the correspondences with the first story. This problem-solving (i.e., inference) task is typical of those used to assess analogical transfer in previous studies; however, some subtle differences should be noted. Gick and Holyoak (1980, 1983) found that the proportion of participants generating the convergence solution increased substantially when an explicit hint to consider the prior story was provided. Although the instructions for the inference task in the present experiment did not specifically provide a hint to use the solved analog as an aid to solving the unsolved analog (as has typically been done in previous studies), the participants had already been provided with an "implicit hint" to consider the similarities between the analogs. Specifically, the instructions for the mapping task pointed out the similarities between the analogs, thereby providing them with a hint that information in the solved story might help them solve the unsolved problem.

Simulation Results

Representations for the radiation analog and general analog were constructed based on the information presented in the stories. While the actual mental representation of either of the stories is likely to vary from one participant to the next, the objects, predicates, and propositions that we generated were chosen to illustrate one plausible representation. Representations for both analogs included propositions that pertained to the solution of the problem, which could be

either included or omitted, depending on what condition was being simulated. In accord with evidence that people are very likely to spontaneously find the solution to the general analog, the simulations were simplified by having LISA use the solved version of this analog regardless of whether it was nominally solved or unsolved. The firing order varied depending on whether the driver was solved or unsolved. When an analog is solved, we assume that the reasoner focuses more on the propositions that are critical to the story, and focuses on these propositions together in groups, therefore tending to process these propositions first and more often. Accordingly, when the driver analog is solved LISA fires the more critical propositions in groups, and fires these groups first. When an analog is unsolved, we assume that the reasoner does not focus as much on the critical propositions (as the reasoner does not know which these are). Accordingly, when the driver is unsolved, LISA either fires its propositions in a random order, or in the order that they appear in the story. (See Appendix D for the LISA representation of the radiation story.)

Because of the stochastic nature of the simulations, we ran twenty trials per firing order. LISA's mean percent correct on the four basic conditions is shown in Figure 2A. Note that we express the general story as either being "solved" or "unsolved", despite the simulations always including a solution, to maintain clarity regarding the nature of the recipient analog. That is, a "solved" general story driver was paired with an *unsolved* radiation story in one condition, while the "unsolved" general driver was paired with a *solved* radiation story in the other. Although the solution was actually provided for the general driver in both conditions in order to simulate the high level of spontaneous solution, the solution was only present for the radiation story in the latter condition. The presence of a solution facilitated mapping when the radiation story was in the role of driver analog, with mapping accuracy being greater when the radiation story was solved (75%) than when it was unsolved (59%). When the solved general story was the driver, mapping accuracy was greater overall, and did not vary with the presence (99%) versus absence of a solution (98%).

Results and Discussion

The dependent measures of interest were percent correct on the mapping task and percent of participants in each condition who provided the analogous convergence solution to the radiation problem. Percent correct mappings was out of a possible six correct. For the solution measure, a complete convergence solution to the ray problem consists of directing multiple, simultaneous, low-intensity rays at the tumor; an answer was scored as correct as long as it clearly stated the "multiple" aspect of the complete solution. The choice of a dichotomous scoring technique over one that would allow partial credit was based on a similar decision made by Gick and Holyoak (1980), who found that the results were similar regardless of whether scoring was all-or-none or whether partial credit was given.

Figure 2B provides the mean mapping accuracy for the major conditions. Mapping accuracy was analyzed using a 2x2x2x2 analysis of variance. Mapping was significantly more accurate when the general analog was the driver (93%) than when the radiation analog was the driver (73%), F (1, 263) = 94.53, MSE = 1.01, $p < .001$, consistent with the greater internal coherence of the former, and with LISA's predictions. There was an overall trend toward more accurate mappings when the solved rather than the unsolved analog was the driver (85% versus 81%), but this difference was not significant, F (1, 263) = 3.18, MSE = 1.01, $p = .075$. However, the interaction between solved versus unsolved driver and the analog assigned to be driver (general or radiation) was significant, F (1, 263) = 6.05, MSE = 1.01, $p < .02$. When the radiation analog was the driver, mapping accuracy was greater when it was presented with a solution than when it was unsolved (77% versus 68%), F (1, 263) = 7.98, MSE = 1.01, $p < .01$; by contrast, when the general analog was the driver, accuracy did not differ as a function of whether the solution was given or not (92% versus 93%), $F < 1$. This pattern is consistent with the fact that people are more likely to generate the convergence solution to the general analog on their own, whereas they are unable to generate it for the radiation problem, and with the simulation results generated by LISA. It should be noted, however, that the statistical interaction could be attributed to a ceiling effect created by the very high accuracy of mapping when the general analog was the driver. Accuracy did not vary significantly as a function of presentation

order of the driver and recipient, $F < 1$, or order of mapping questions, F (1, 263) = 1.68, MSE = 1.01, $p > .15$.

_____

Insert Figure 2 about here

_____

The conditions most comparable to those in which an asymmetry in mapping accuracy was observed in Experiment 3 were (1) solved general analog as driver with unsolved radiation problem as recipient versus (2) unsolved radiation problem as driver with solved general analog as recipient (because the general story is the more coherent analog; hence the first of these conditions places the more coherent analog in the role of driver). The advantage of having the more coherent analog in the role of driver was replicated with the semantically-rich materials of Experiment 4 (92% versus 68%), F (1, 263) = 73.51, MSE = 1.01, $p < .0001$.

The overall pattern of mapping accuracy supports the predictions of the LISA model. Mapping was relatively accurate whenever the analog with greater causal structure (the general analog) was the driver; if the analog with less causal structure (the radiation analog) was the driver, mapping was aided if the solution was provided to increase its causal coherence.

An analysis of solution frequencies revealed that the convergence solution was generated more often overall for the general story than for the radiation story (88% versus 39%, 116 out of 132 participants and 57 out of 147 participants, respectively), X^2 (1) = 71.18, $p < .001$. This difference provides further evidence that the convergence solution can be generated more easily (either with or without an analog) for the former story. When the radiation story was the unsolved analogy, the convergence solution tended to be produced more frequently when it was the recipient and the solved general was the driver as compared to when it was the driver and the solved general the recipient (44% versus 34%, 32 out of 73 participants and 25 out of 74 participants, respectively). Although this difference was not significant, X^2 (1) = 1.56, $p > .20$,

the frequencies are close to the comparable frequencies observed in data from another experiment with a greater number of participants in the two conditions (46% versus 31%), which yielded a reliable difference, X^2 (1) = 6.84, p < .01. (The difference in mapping accuracy for these two conditions was replicated in the other experiment.) The lack of statistical reliability for the difference in solution frequencies observed in Experiment 4 can be attributed to the much lower numbers of participants in these two conditions in Experiment 4 (147) than in the replication experiment (314).

Mapping accuracy and transfer were found to be positively related, with a point-biserial correlation of .32, p < .01. In addition, the average mapping accuracy for those participants who provided the correct solution to the unsolved problem was 88% (n = 106), which was significantly greater than for those who did not provide the correct solution (75%, n = 173), t(277) = 5.54, p < .001. Thus, as in Experiment 2, greater mapping accuracy was associated with greater transfer.

General Discussion

Implications for Analogy Models

The results of the four experiments strongly supported LISA's predictions concerning grouping effects, structural constraints, and directional asymmetries in analogical mapping. Experiment 1 demonstrated that mapping accuracy is greater when a direct cue is given to group specific sentences in a noncausal driver, but only when the grouped propositions jointly provide strong structural constraints on the mapping. This finding was replicated in Experiment 2, which used semantically richer, more meaningful materials, and which demonstrated the influence of grouping on inference as well as mapping.

Experiment 3 demonstrated an asymmetry at the mapping stage, with mapping being more accurate when the analog with greater causal coherence was in the role of driver (and the

less coherent analog was in the role of recipient) than when those roles were reversed. Experiment 4 provided a wider range of evidence that greater causal coherence in the driver increases mapping accuracy by using more meaningful materials, the radiation problem and the analogous general story (Gick & Holyoak, 1980). Mapping accuracy was greater when the more-coherent general analog (solved or unsolved) was the driver and the less-coherent radiation problem was the recipient, rather than vice versa. When the radiation problem was the driver, mapping accuracy was higher if it was solved rather than unsolved. The asymmetry effect observed in Experiment 3 was quite small (a 7% advantage when the more coherent analog was the driver); but the corresponding asymmetry observed in Experiment 4 was more robust (a 24% difference between the two asymmetrical conditions). More generally, although LISA predicts that the *process* of analogical mapping is inherently asymmetrical (with the driver analog being the focus of attention), the *products* of analogical mapping (the correspondences found between the two analogs) are only predicted to be asymmetrical in specific circumstances (in particular, when the two analogs differ in relative coherence).

All of our findings can be accounted for in terms of the role of text coherence in guiding LISA as it operates within its inherent WM limitations. In LISA, mapping is necessarily directional (with the driver being the current focus of attention), leading to the potential for mapping asymmetries. The mapping process is necessarily incremental, with no more than 2-3 driver propositions being activated at once. These WM limits imply that mapping is sensitive to whether (and which) driver propositions are activated together in the phase set. In LISA, any factor that affects groupings of driver propositions, including causal coherence, provision of a solution, and overt grouping cues, can influence mapping accuracy.

It should be emphasized that LISA's mapping performance need not require conscious or analytic decision-making. Rather, we assume that human analogical processing has evolved to

favor implicit algorithms that serve to optimize use of limited WM resources. Although the version of LISA used in the present simulations was told the order and grouping of propositions to use in mapping, we assume that this process is executed autonomously by the human analogy engine. Selection of propositions for firing may be accomplished implicitly by a process of activation-passing, which will favor the joint selection of propositions linked by causal relations or overlap of elements. The LISA simulation of the complex radiation-problem analogy (Experiment 4), as well as other simulations using even larger analogical representations (Holyoak & Hummel, 2001), demonstrate that the model, like humans, can work incrementally within its WM limits to map analogs that are far too large and complex to process as whole structures.

No other computational model of analogy can account for the full range of the present findings as well as related findings concerning the role of WM in mapping. Most previous models, including SME (Falkenhainer et al., 1989), I-SME (Forbus et al., 1994) and ACME (Holyoak & Thagard, 1989) imply that the mapping process is inherently symmetrical. The IAM model (Keane, 1997; Keane & Brayshaw, 1988; Keane et al., 1994) is the only other extant model that can potentially account for mapping asymmetries and grouping effects. In Experiment 1, IAM erroneously predicted that the Bad Grouping condition would yield more accurate mappings than the No Grouping condition, and in Experiment 2 it erroneously predicted that the Low Constraint condition would be just as effective as the High Constraint condition. In addition to accounting for the present findings, LISA can also explain why restricting WM capacity by adding a dual-task requirement (Waltz et al., 2000) or by increasing state anxiety (Tohill & Holyoak, 2000) impairs people's ability to map on the basis of relations.

The Basis for Asymmetries in Analogical Processing

The present findings raise some questions regarding previous interpretations of asymmetries in the literature (e.g., Bassok & Olseth, 1995; Bowdle & Gentner, 1997; Burns, 1996). Most previous analyses of asymmetry in analogical reasoning have assumed that any asymmetries necessarily arise at the inference stage (i.e., post-mapping). For instance, Bowdle and Gentner (1997) found an asymmetry in similarity comparisons based on the level of informativeness associated with each direction of comparison. That is, a comparison from a more coherent base to a less coherent target was preferred over a comparison from a less coherent base to a more coherent target, and the former led to a greater level of informativeness in the inferences drawn from source to target. Their findings were interpreted from the viewpoint of the structure-mapping theory, using SME to simulate many of their findings. As noted previously, the mappings generated in SME are symmetrical between source and target analogs, with potential asymmetries occurring in the number of candidate inferences drawn. SME has also been used to explain observed asymmetries in metaphor comprehension (Gentner & Wolff, 1997); as in the case of similarity judgments, this approach assumes that all asymmetries in comprehension arise during the projection of inferences, and not at the prior mapping, or alignment, stage. However, none of these studies directly measured mappings, so the data do not reveal whether the observed asymmetries arose at the mapping stage, the inference stage, or both. Given the present evidence for asymmetries at the mapping stage, the possibility that mapping asymmetries sometimes contribute to asymmetries in similarity comparisons and metaphor comprehension should be considered.

Some previous findings in the literature support the possibility that asymmetries may arise during mapping (e.g., Bassok & Holyoak, 1989; Bassok & Olseth, 1995). Bassok and Olseth (1995) found that transfer was greater from word problems that dealt with discrete

variables to word problems with continuous variables than vice versa. Although the solutions to both types of problems (those with discrete variables and those with continuous variables) were structurally isomorphic, Bassok and Olseth posited that the participants may have formed non-isomorphic representations of the two classes of problems, which in turn led to the observed asymmetry in transfer. Bassok and Olseth measured problem-solving transfer and not mapping accuracy; nonetheless, it is possible that the non-isomorphic representations led to asymmetries in mapping, which in turn led to asymmetries in transfer. Assuming that discrete variables are represented as points arranged in a linear fashion and continuous variables are represented as a continuous line (see Alibali, Bassok, Olseth, Syc, & Goldin-Meadow, 1995), mapping from the discrete representation to the continuous representation would be fairly straightforward (e.g., the discrete points could simply be mapped to points on the continuous line), whereas the mapping from the continuous to the discrete would be more ambiguous (e.g., it would be unclear how to map the continuous line to the discrete points). Such a difference in ease of mapping could lead to an asymmetry in mapping accuracy, which could then lead to an asymmetry in inference, and thus transfer (a linkage supported by results for solving algebra word problems reported by Novick, 1995). (See Burns, 1996, for additional evidence regarding asymmetries in analogical transfer.) The present findings suggest that more research is needed to determine the locus of asymmetries in other tasks that require similarity comparisons or analogical reasoning.

In addition to their theoretical implications, the present findings have potential relevance to the educational use of analogies. Given the advantage of processing from the more coherent analog to the less coherent analog, it follows that learners should be encouraged to perform mappings from the better-understood source analog to the less-understood target analog, rather than vice versa. It may be more effective, for example, to map a worked example problem to a novel target problem, rather than vice versa, if inclusion of a solution renders the worked

example more causally coherent. As naive problem solvers may tend to spontaneously focus on the target problem (which it is their explicit goal to solve), it is possible that instruction to map in the source-to-target direction (i.e., making the source rather than the target the focus of attention) could aid learners in their use of more familiar cases or worked examples to solve a novel problem.

In conclusion, LISA's inherent WM limitations generate a number of distinctive predictions for which the findings reported here provide empirical support. Moreover, the model suggests possible alternative explanations of previous findings of asymmetries in similarity comparisons and analogical reasoning, and suggests possible ways to use analogies more effectively as instructional tools. LISA offers a first step towards answering the basic question of how human reasoners are able to perform complex reasoning tasks, such as analogical mapping and inference, while working within the severe limits of human working memory.

Footnotes

1. The correspondence between the order in which propositions are stated in a story and the organization of these same propositions in terms of their textual coherence may vary. If there is a high degree of textual coherence among the propositions (as would be associated with a clear understanding of the relations within the text), then LISA may fire the more important propositions (e.g., the more critical or causally-relevant propositions) first, and more often, rather than as they appear in the story. On the other hand, if there is a low degree of textual coherence (as would be associated with poorer understanding of the relations within the text), LISA's default strategy would be to activate the propositions as they appeared in the story. Experiment 4 explores these issues further.

2. In broad strokes, the reason is that placing multiple propositions into WM simultaneously allows the repeated elements (i.e., those elements that appear in more than one of those propositions) to "sample" their multiple potential mappings (i.e., the different mappings suggested by the different propositions in which they appear) before committing to any one of them. Maximizing the number of repeated elements therefore maximizes the amount of constraint that is gained by placing multiple propositions into WM together.

 3. It is possible that human subjects in the No Grouping condition occasionally processed propositions in groups, and therefore potentially grouped propositions 2 and 3 (essentially transforming the No Grouping condition into the Good Grouping condition); however, such spontaneous groupings were not incorporated into LISA's simulations. No such grouping was encouraged by the instructions, and the lower mapping performance in the noncausal condition tested in an earlier experiment (which was identical to the No Grouping condition of Experiment 1) suggests that such spontaneous grouping effects were infrequent in our participants.

4. We thank Mark Keane for performing IAM simulations for the three conditions tested in Experiment 1 and the two conditions tested in Experiment 2.

References

Alibali, M. W., Bassok, M., Olseth, K. L., Syc, S., & Goldin-Meadow, S. (1995). Gestures reveal mental models of discrete and continuous change. In J. D. Moore & J. F. Lehman (Eds.), *Proceedings of the Seventeenth Annual Conference of the Cognitive Science Society* (pp. 391-396). Hillsdale, NJ: Erlbaum.

Baddeley, A. D. (1992). Working memory. *Science, 255,* 556-559.

Bassok, M., & Holyoak, K. J. (1989). Interdomain transfer between isomorphic topics in algebra and physics. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *15*, 153-166.

Bassok, M., & Olseth, K. L. (1995). Object-based representations: Transfer between cases of continuous and discrete models of change. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *21*, 1522-1538.

Bowdle, B. F., & Gentner, D. (1997). Informativitiy and asymmetry in comparisons. *Cognitive Psychology, 34*, 244-286.

Broadbent, D. E. (1975). The magical number seven after fifteen years. In A. Kennedy & A. Wilkes (Eds.), *Studies in long-term memory* (pp. 3-18). New York: Wiley.

Bundesen, C. (1998). Visual selective attention: Outlines of a choice model, a race model and a computational theory. *Visual Cognition*, *5*, 287-309.

Burns, B. (1996). Meta-analogical transfer: Transfer between episodes of analogical reasoning. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 22*, 1032-1048.

Cowan, N. (1995). *Attention and memory: An integrated framework.* New York: Oxford University Press.

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

Ericsson, K. A., & Kintsch, W. (1995). Long-term working memory. *Psychological Review*, *102*, 211-245.

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

Fletcher, C. R. (1986). Strategies for the allocation of short-term memory during comprehension. *Journal of Memory and Language, 25*, 43-58.

Fletcher, C. R., & Bloom, C. P. (1988). Causal reasoning in the comprehension of simple narrative texts. *Journal of Memory and Language, 27*, 235-244.

Fletcher, C. R., Hummel, J. E., & Marsolek, C. J. (1990). Causality and the allocation of attention during comprehension. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 16*, 233-240.

Forbus, K. D., Ferguson, R. W., & Gentner, D. (1994). Incremental structure-mapping. In A. Ran & K. Eiselt (Eds.), *Proceedings of the Sixteenth Annual Conference of the Cognitive Science Society* (pp. 313-318). Hillsdale, NJ: Erlbaum.

Gentner, D., & Markman, A. B. (1997). Structural vs. syntactic matching: Analogy entails common relations. In M. G. Shafto & P. Langley (Eds.), *Proceedings of the Nineteenth Annual Conference of the Cognitive Science Society* (p. 927). Mahweh, NJ: Erlbaum.

Gentner, D., & Wolff, P. (1997). Alignment in the processing of metaphor. *Journal of Memory and Language, 37*, 331-355.

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

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

Glucksberg, S., & Keysar, B. (1990). Understanding metaphorical comparisons: Beyond similarity. *Psychological Review*, *97*, 3-18.

Halford, G. S., Wilson, W. H., Guo, J., Gayler, R. W., Wiles, J., & Stewart, J. E. M. (1994). Connectionist implications for processing capacity limitations in analogies. In K. J. Holyoak & J. A. Barnden (Eds.), *Advances in connectionist and neural computation theory, Vol. 2: Analogical connections* (pp. 363-415). Norwood, NJ: Ablex.

Hoadley, C. M., Ranney, M., & Schank, P. (1994). WanderECHO: A connectionist simulation of limited coherence. In A. Ran & K. Eiselt (Eds.), *Proceedings of the Sixteenth Annual Conference of the Cognitive Science Society* (pp. 421-426). Hillsdale, NJ: Erlbaum.

Hofstadter, D. R., & Mitchell, M. (1994). The Copycat project: A model of mental fluidity and analogy-making. In K. J. Holyoak & J. A. Barnden (Eds.), *Analogical connections. Advances in connectionist and neural computation theory, Vol. 2* (pp. 31-112). Norwood, NJ: Ablex.

Holyoak, K. J., & Hummel, J. E. (2001). Toward an understanding of analogy within a biological symbol system. In D. Gentner, K. J. Holyoak, & B. N. Kokinov (Eds.), *The analogical mind: Perspectives from cognitive science* (pp. 161-195). Cambridge, MA: MIT Press.

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

Holyoak, K. J., & Thagard, P. (1995). *Mental leaps: Analogy in creative thought*. Cambridge, MA: MIT Press.

Hummel, J. E., & Biederman, I. (1992). Dynamic binding in a neural network for shape recognition. *Psychological Review, 99*, 480-517.

Hummel, J. E., & Holyoak, K. J. (1993). Distributing structure over time. *Behavioral and Brain Sciences*, *16*, 464.

Hummel, J. E., & Holyoak, K. J. (1997). Distributed representations of structure: A theory of analogical access and mapping. *Psychological Review, 104*, 427-466.

Hummel, J. E., & Holyoak, K. J. (in press). Analogy and creativity: Schema induction in a structure-sensitive connectionist model. In T. Dartnall & D. Peterson (Eds.), *Creativity and computation*. New York: Elsevier.

Hummel, J. E., & Stankiewicz, B. J. (1996). An architecture for rapid, hierarchical structural description. In T. Inui & J. McClelland (Eds.), *Attention and performance XVI: Information integration in perception and communication* (pp. 93-121). Cambridge, MA: MIT Press.

Irwin, D. E. (1992). Perceiving an integrated visual world. In D. E. Meyer & S. Kornblum (Eds.), *Attention and performance XIV: Synergies in experimental psychology, artificial intelligence, and cognitive neuroscience - A Silver Jubilee* (pp. 121-142). Cambridge, MA: MIT Press.

Keane, M. T. (1997). What makes an analogy difficult? The effects of order and causal structure on analogical mapping. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 23*, 946-967.

Keane, M. T., & Brayshaw, M. (1988). The Incremental Analogical Machine: A computational model of analogy. In D. Sleeman (Ed.), *European working session on learning* (pp. 53-62). London: Pitman.

Keane, M. T., Ledgeway, T., & Duff, S. (1994). Constraints on analogical mapping: A comparison of three models. *Cognitive Science, 18*, 387-438.

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

Kokinov, B. N. (1994). A hybrid model of reasoning by analogy. In K. J. Holyoak, & J. A. Barnden, (Eds.), *Analogical connections. Advances in connectionist and neural computation theory, Vol. 2* (pp. 247-318). Norwood, NJ: Ablex Publishing Corp.

Lisman, J. E., & Idart, M. A. P. (1995). Storage of 7 ± 2 short-term memories in oscillatory subcycles. *Science*, *267*, 1512-1515.

Luck, S. J., & Vogel, E. K. (1997). The capacity of visual working memory for features and conjunctions. *Nature*, *390*, 279-281.

McClelland, J. L. & Rumelhart, D. E. (1981). An interactive activation model of context effects in letter perception: I. An account of basic findings. *Psychological Review, 88*, 375-407.

Novick, L. R. (1995). Some determinants of successful analogical transfer in the solution of algebra word problems. *Thinking and Reasoning*, *1*, 5-30.

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

Ranney, M., & Thagard, P. (1988). Explanatory coherence and belief revision in naive physics. *Proceedings of the Tenth Annual Conference of the Cognitive Science Society* (pp. 426-432). Hillsdale, NJ: Erlbaum.

Reed, S. K., Ernst, G. W., & Banerji, R. (1974). The role of analogy in transfer between similar problem state. *Cognitive Psychology, 6*, 436-450.

Rosch, E. (1975). Cognitive reference points. *Cognitive Psychology, 7*, 532-547.

Shastri, L. & Ajjanagadde, V. (1993). From simple associations to systematic reasoning: A connectionist representation of rules, variables and dynamic bindings using temporal synchrony. *Behavioral & Brain Sciences, 16*, 417-494.

Singer, W., & Gray, C. M. (1995). Visual feature integration and the temporal correlation hypothesis. *Annual Review of Neuroscience*, *18*, 555-586.

Spellman, B. A., & Holyoak, K. J. (1996). Pragmatics in analogical mapping. *Cognitive Psychology*, *31*, 307-346.

Sperling, G. (1960). The information available in brief presentations. *Psychological Monographs*, *74* (whole no. 498).

Tohill, J. M., & Holyoak, K. J. (2000). The impact of anxiety on analogical reasoning. *Thinking & Reasoning*, *6*, 27-40.

Tversky, A., & Gati, I. (1978). Studies of similarity. In E. Rosch & B. B. Lloyd (Eds.), *Cognition and categorization* (pp. 79-98). Hillsdale, NJ: Erlbaum.

van den Broek, P. (1988). The effects of causal relations and hierarchical position on the importance of story statements. *Journal of Memory and Language, 27*, 1-22.

Waltz, J. A., Lau, A., Grewal, S. K., & Holyoak, K. J. (2000). The role of working memory in analogical mapping. *Memory & Cognition*, *28*, 1205-1212.

Appendix A

Example of LISA's Representation of Analogs and Firing Orders for Experiments 1 and 3

I. Representations of Analogs

For a full description of the operation of LISA, see Hummel and Holyoak (1997). LISA's representations of analogs are comprised of objects, predicates, and propositions. The input consists of propositions in a form of predicate calculus, combined with names for objects and predicates and a list of associated semantic units for each object and predicate. Predicates have a number of roles (indicated by a digit after its name), each of which has a set of semantic units. By default, each role has a distinct set of semantic units. (Only one set of semantic units is shown below for each predicate.) Names are only used for ease of reference; they have no significance for LISA's algorithm. The names of semantic units are also arbitrary labels. The units ending in a digit represent unique semantic features of that object or predicate.

"Jim kisses Mary" Analog

    Objects:

        Mary   person m1 m2 m3

        Jim   person j1 j2 j3

        Bill   person b1 b2 b3

    Predicates:

        kiss 2   pred k1 k2 k3

        love 2   pred l1 l2 l3

        jealous 2   pred je1 je2 je3

        causally-linked 2   cause c1 c2 c3

    Propositions:

        P1   (kiss Jim Mary)

        P2   (love Jim Mary)

        P3   (love Bill Mary)

P4   (jealous Bill Jim)

P5   (causally-linked P2 P3)

II. Firing Orders

Numbers indicate which proposition of the driver analog is fired, and hyphens indicate an updating of the mapping connections (e.g., with the firing order P1P2-P3-P4, propositions 1 and 2 are fired together in a phase set, then the mapping connections are updated, followed by firing of proposition 3 and another update, and finally firing of proposition 4 and an update).

	Condition	Firing Orders
Experiment 1		
	Good Grouping	P2P3-P1-P4, P2P3-P4-P1
	Bad Grouping	P3P4-P1-P2, P3P4-P2-P1
	No Grouping	P1-P2-P3-P4, P3-P4-P1-P2,
		P4-P1-P2-P3, P4-P3-P2-P1
Experiment 3		
	Causal-Causal	P5-P2P3-P1-P4, P5-P2P3-P4-P1
	Causal-Noncausal	P5-P2P3-P1-P4, P5-P2P3-P4-P1
	Noncausal-Causal	P1-P2-P3-P4, P3-P4-P1-P2,
		P4-P1-P2-P3, P4-P3-P2-P1
	Noncausal-Noncausal	P1-P2-P3-P4, P3-P4-P1-P2,
		P4-P1-P2-P3, P4-P3-P2-P1

Appendix B

Example Driver and Recipient Stories Used in Experiment 2 (Bank Context)

Driver, Bank Context

After a long day at work, you come home and decide to relax by watching some television. Your roommate, who is bilingual in English and Italian, is already watching a movie which is spoken entirely in Italian. Since you do not know Italian, in order to enjoy the movie, you must try to understand the storyline by the context. You watch and pay careful attention to the characters' actions, and try to figure out what is going on. There seem to be three main characters, Danny, Lillian, and Steve, who seem to form a small gang of criminals. They are at a bank, and seem to be planning on robbing it.

[*High Constraint:* You see Danny wave to Lillian, and then you see Lillian wave to Steve. You ask your roommate what's going on, and it's explained to you that Danny waved to Lillian to inform her that the security guard had left the building to patrol the parking lot, and that Lillian waved to Steve to signal him to start the robbery.]

[*Low Constraint:* You see Danny wave to Lillian, and then see Steve throw a small bag to Danny. You ask your roommate what's going on, and it's explained to you that Danny waved to Lillian to signal her that they were about to begin their robbery, so Lillian should be on the lookout for the security guard; the bag that Steve threw to Danny contained the gun that they were planning on using to rob the bank.]

The movie seems interesting, but you have to go to the bathroom.

After going to the bathroom, you come back and start watching the movie again. The three main characters seem to still be at the same bank. [*High Constraint:* You see Steve throw a bag to Danny, and then you see Danny throw the same bag to Lillian. Your roommate tells you that Steve had filled the bag with stolen money, and threw it to Danny to give to Lillian, who was closest to the door.] [*Low Constraint:* You see Danny throw a bag to Lillian, and then see Lillian wave at Steve. Your roommate tells you that Danny had thrown Lillian a bag full of money, and that Lillian waved to Steve to signal for him to start the getaway car, as the robbery was complete.]

While the action was fun to watch, the characters start to talk at length about something, and you find that you simply cannot follow the storyline. You decide that you would rather read a book instead of watching a movie you do not understand, so you go to your room and start reading the latest Tom Clancy novel.

Recipient Order 1, Bank Context

The Trojan Film Company's movie has three main characters, Danny, Lillian, and Steve, who seem to form a small gang of criminals. They are at a bank, and seem to be planning on robbing it. Below are events which take place while they are in the bank:

Lillian waves to Steve
Danny throws a bag to Lillian
Danny waves to Lillian
Steve throws a bag to Danny

Appendix C

LISA's Representations of Driver and Recipient Versions of Bank Context for Experiment 2

Recipient Version

      Objects:

Danny	person danny1 danny2 danny3 male
Lillian	person lil1 lil2 lil3 female
Steve	person steve1 steve2 steve3 male

      Predicates:

wave 2	pred wave1 wave2 wave3 communicate
throwbag 2	pred throw1 throw2 throw3 transfer

      Propositions:

P1	(wave Danny Lillian)
P2	(wave Lillian Steve)
P3	(throwbag Steve Danny)
P4	(throwbag Danny Lillian)

Driver Version (identical to Recipient Version, with the following additions)

      Predicates:

related1 2	related1 related2 related3
related2 2	related11 related22 related33

      Propositions:

      High Constraint:

P5	(related1 P1 P2)
P6	(related2 P3 P4)

      Low Constraint:

P5	(related1 P1 P3)
P6	(related2 P2 P4)

Appendix D

LISA's Representation of Radiation Story for Experiment 4

Radiation Story

Objects:

rays	object force power straight linear extend blowup
tissue	object biological tissue1 tissue2
tumor	object bad death tumor1
stomach	object biological tissue1 stomach1
doctor	person human alive doc1 doc2 doc3
patient	person human alive pat1 pat2 pat3
raysource	object artifact raysource1
hirays	energy radiation hirays1
lorays	energy radiation lorays1

Predicates:

gointo 2		approach go1 go2
vulnerable 2	vul1 vul2 vul3
surround 2		surround1 surround2 surround3
wantsafe 2	safe1 safe2 safe3
inside 2		state location in1
destroyed 1	state reduce destroyed1
want 2		mentalstate goal want1
notwant 2		mentalstate goal negative notwant1
canmake 2		trans potential change generate canmake1
candestroy 2	trans potential change reduce candestroy1
cannotdest 2	trans potential neutral powerless cannotdest1
ifthen 2		conditional
use 2		trans utilize use1

Propositions:

P1	(inside tumor stomach)
P2	(destroyed tumor)
P3	(want doctor P2)
P4	(canmake raysource hirays)
P5	(candestroy hirays tumor)
P6	(gointo rays tumor)
P7	(vulnerable tumor rays)
P8	(vulnerable tissue rays)
P9	(surround tissue tumor)
P10	(use doctor hirays)
P11	(ifthen P10 P2)
P12	(candestroy hirays tissue)
P13	(destroyed tissue)
P14	(notwant doctor P13)
P15	(ifthen P10 P13)
P16	(wantsafe doctor tissue)

Solution:

P17	(canmake raysource lorays)
P18	(cannotdest lorays tumor)
P19	(cannotdest lorays tissue)
P20	(use doctor lorays)
P21	(ifthen P20 P2)

Table 1

Materials Used by Keane (1997, Experiment 3) and in Present Experiments 1 and 3

(Adapted by permission from Keane, 1997, p. 966)

Causal Source	Noncausal Source	Noncausal Target
1) Jim kisses Mary.	Jim hugs Mary.	Ruth knows Debra.
2) Jim loves Mary.	Jim sees Mary.	Ruth motivates Debra.
3) Bill loves Mary.	Bill sees Mary.	Laura motivates Debra.
4) Bill is jealous of Jim.	Bill is beside Jim.	Laura waves to Ruth.

Table 2

Focal Propositions Used in Experiment 2

Bank Context	Spy Context
1) Danny waves to Lillian.	Sharon winks at Rich.
2) Lillian waves to Steve.	Rich winks at Julia.
3) Steve throws a bag to Danny.	Julia gives a black box to Sharon.
4) Danny throws a bag to Lillian.	Sharon gives a black box to Rich.

Table 3

Assignment of Analogs to Roles in Design of Experiment 4

Role of Analog

	Driver	Recipient
	General solved	Radiation unsolved
Analogs Used	General unsolved	Radiation solved
	Radiation solved	General unsolved
	Radiation unsolved	General solved

Figure Captions

Figure 1. LISA simulation (Panel A) and human mapping performance (Panel B) for each condition in Experiment 3.

Figure 2. LISA simulation (Panel A) and human mapping performance (Panel B) for each condition in Experiment 4.

A. LISA's Simulation of Mapping Performance
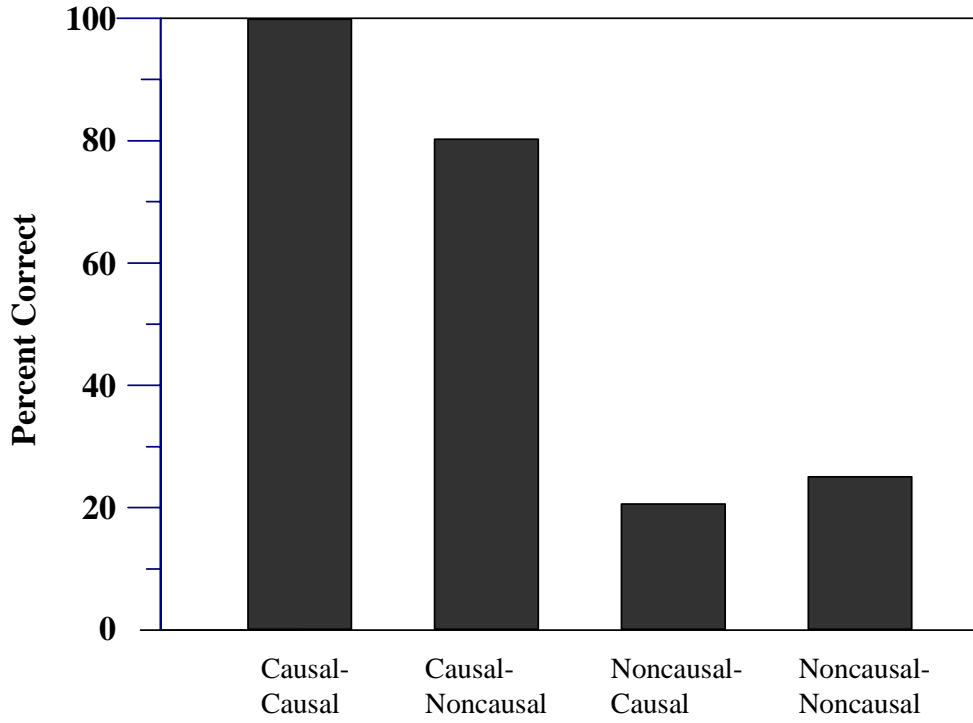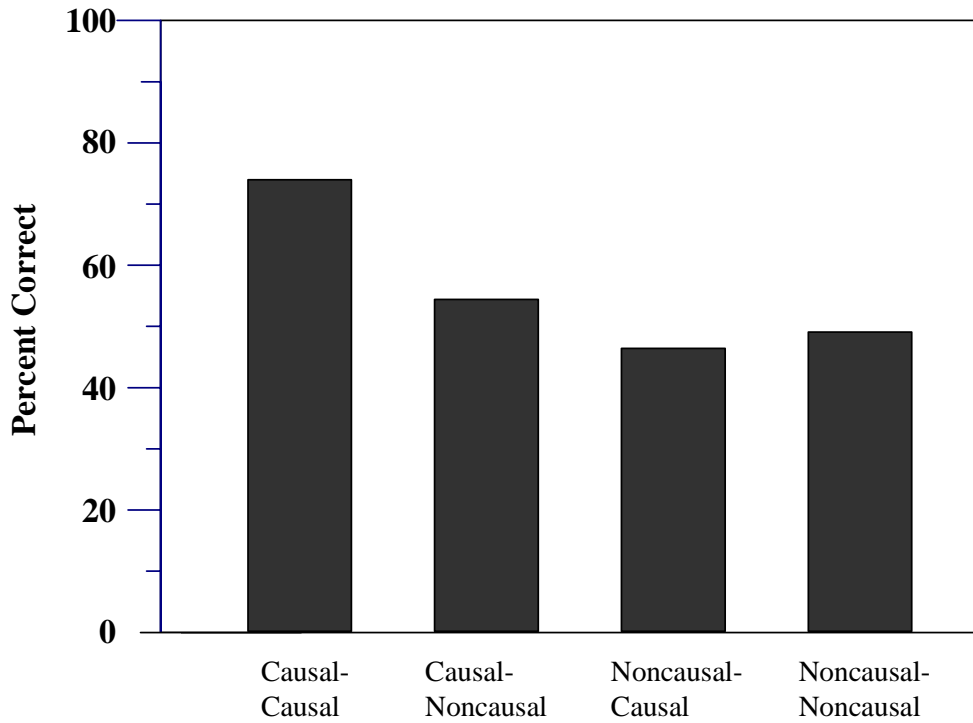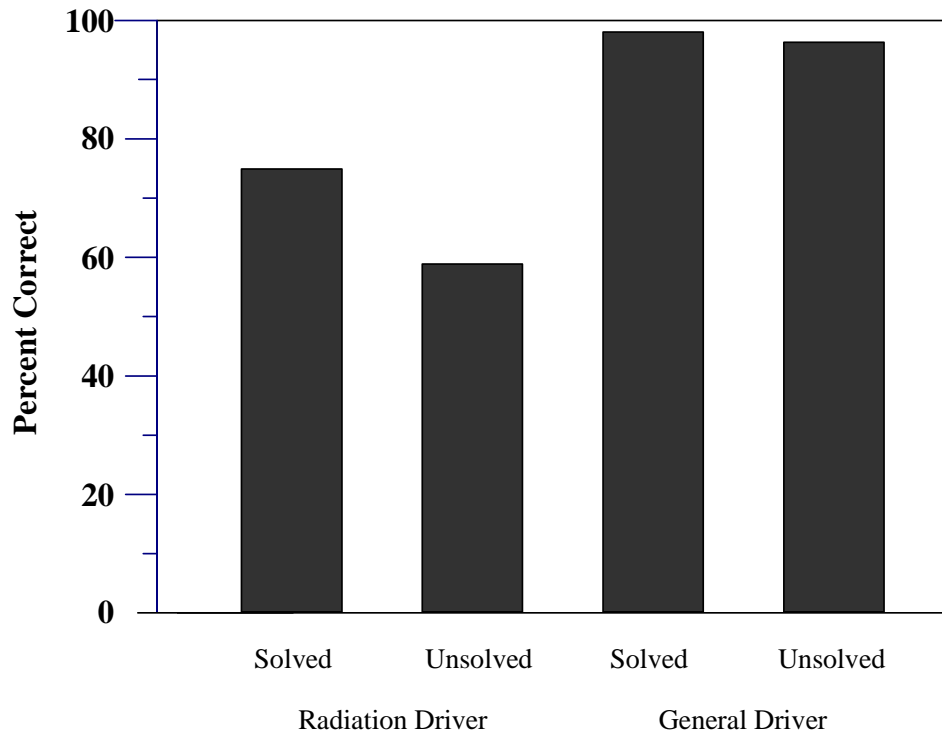

B. Human Mapping Performance

Figure 1.

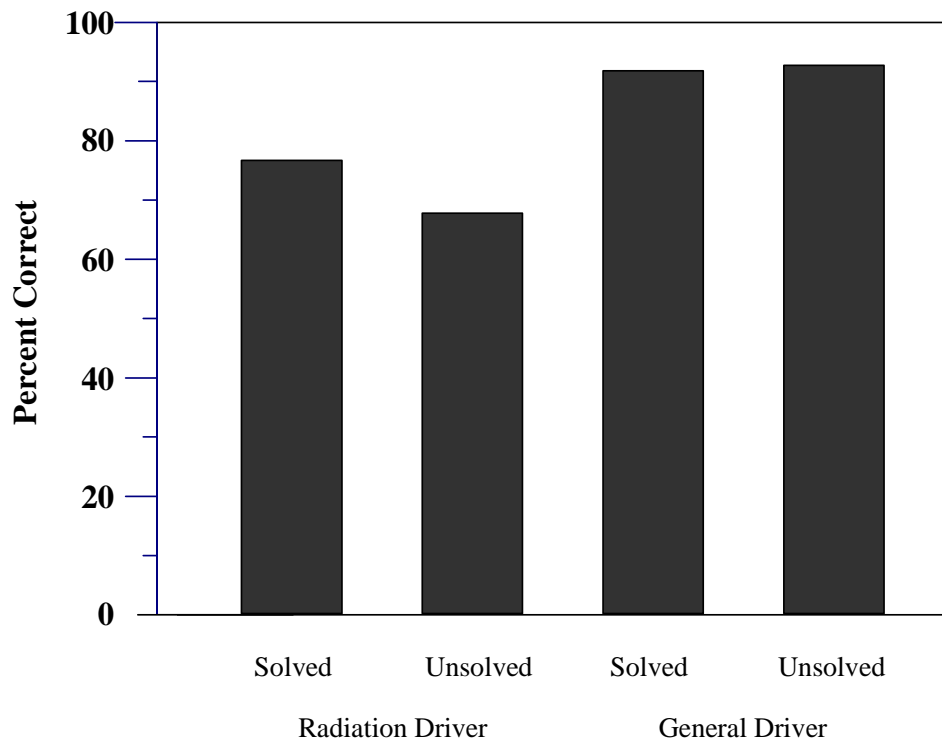## A. LISA's Simulation of Mapping Performance

B. Human Mapping Performance

Figure 2.