The Analogical Mind

Keith J. Holyoak *University of California, Los Angeles*
Paul Thagard *University of Waterloo*

The use of analogy in human thinking is examined from the perspective of a multiconstraint theory, which postulates 3 basic types of constraints: similarity, structure, and purpose. The operation of these constraints is apparent in laboratory experiments on analogy and in naturalistic settings, including politics, psychotherapy, and scientific research. The multiconstraint theory has been implemented in detailed computational simulations of the analogical human mind.

Many parents know that young children take comfort in getting a kiss on an injury to "make it better." Little Aaron, aged 24 months, would routinely come to his mother saying things like, "I bump my head. Kiss it." But one morning, for the first time ever, the tables turned. While his mother was dressing him, she realized she had a bruise on her hand. Without really thinking, she said, "Ow, my hand hurts." Aaron immediately responded, "I kiss it." His mother then put her hand in front of Aaron's face and received a kiss from him.[1]

Aaron's reaction provides a small example of thinking by analogy—trying to reason and learn about a new situation (the *target* analog) by relating it to a more familiar situation (the *source* analog) that can be viewed as structurally parallel. Aaron's source was the knowledge that when he had been hurt in the past, his mother's kiss had made it better; this source was then evoked by the target situation of his mother's bruised hand (the *access* or retrieval step in analogy use). The child went on to find the correspondences between the source and the target (the *mapping* step). Note that he did not simply use the superficial mapping of his mother to herself (if he had, he would have simply told her to kiss her own hand!). Rather, Aaron mapped his mother to himself (for she was the injured one) and himself to her. On the basis of these mappings, he found a solution to the target problem—his kiss would ease her pain (the *inference* step). Although it isn't known for sure, it is quite possible that Aaron's use of analogy also led him to learn something more general, a kind of abstraction of the commonalities shared by the source and the target (the *learning* step). Roughly, he may have induced a schema or rule along the following lines: "If a person is injured, a kiss from a loved one will ease the pain." Our description of Aaron as analogizing from treatment of his own injury to treatment of his mother's injury assumes that he had not previously formed this general rule.

At two years of age, Aaron had an analogical mind. The remarkable thing about this example of a child's reasoning is that it is not especially exceptional. Young children, before they enter school, without any specialized tutoring from their parents or elders, develop a capacity for analogical thinking (e.g., Gentner, 1977; Goswami & Brown, 1989; Holyoak, Junn, & Billman, 1984; Inagaki & Hatano, 1987). The analogical mind is simply the mind of a normal human being. Indeed, to a limited but impressive degree, it is the mind of at least a few other primates, most notably chimpanzees that have received extensive training in symbol manipulation (Gillan, Premack, & Woodruff, 1981). Analogical thinking can be traced from these early phylogenetic and ontogenetic beginnings to an extraordinarily diverse range of uses by human adults, including generation of metaphors for the self; decision making in politics, business, and law; and scientific discovery.

Our aim in this article is to provide an overview of analogical thinking from a perspective we have termed the *multiconstraint theory* (Holyoak & Thagard, 1989, 1995). As its name implies, the multiconstraint theory assumes that people's use of analogy is guided by a number of general constraints that jointly encourage coherence in analogical thinking. First, we describe these constraints in qualitative terms, illustrating them with examples from psychological studies. Then, we survey additional examples of naturalistic uses of analogy that can be understood in terms of the constraints. Finally, we discuss approaches to implementing the multiconstraint theory in computational models that simulate the human analogical mind (for a more thorough discussion of these issues, see Holyoak & Thagard, 1995).

Keith J. Holyoak, Department of Psychology, University of California, Los Angeles; Paul Thagard, Department of Philosophy, University of Waterloo, Waterloo, Ontario, Canada.

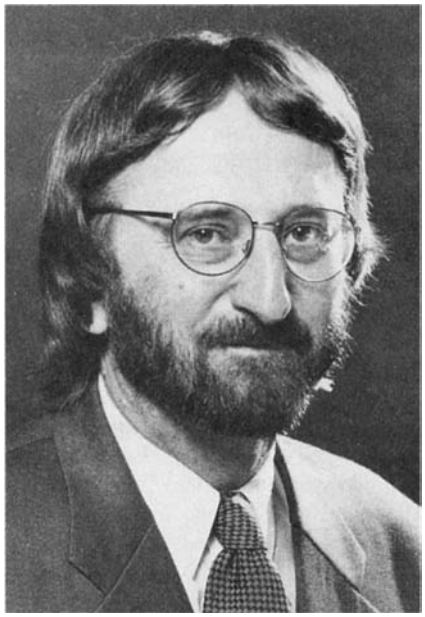
Preparation of this article was supported by National Science Foundation Grant SBR-95-11504 and by a grant from the Natural Sciences and Engineering Research Council of Canada.

The sections of this article reviewing mapping experiments and the Analogical Mapping by Constraint Satisfaction model are adapted from Holyoak and Thagard (1995), with permission of MIT Press.

Correspondence concerning this article should be addressed to Keith J. Holyoak, Department of Psychology, University of California, Los Angeles, Franz Hall, Los Angeles, CA 90095-1563. Electronic mail may be sent via Internet to holyoak@psych.ucla.edu.

[1] We thank Aaron Novick for producing this example and his mother, Laura Novick, for providing it to us along with an interpretation of the episode as an instance of analogical thinking (personal communication, June 14, 1993).

 0003-066X/97/$2.00
Vol. 52, No. 1, 35–44



Keith J. Holyoak

Similarity, Structure, and Purpose

Three broad classes of constraints form the basis of the multiconstraint theory. Each of these constraints can be illustrated with young Aaron's analogy. First, the analogy is guided to some extent by direct similarity of the elements involved. We noted that Aaron did not simply map his mother to herself (which would have maximized one local similarity between mapped objects). However, the analogy clearly depended on similarity of key relations between objects: Both the source and the target involved an injury sustained by a loved one. In general, similarity of concepts at any level of abstraction contributes to analogical thinking, particularly in the initial access step (e.g., Keane, 1986; Ross, 1989; Seifert, McKoon, Abelson, & Ratcliff, 1986).

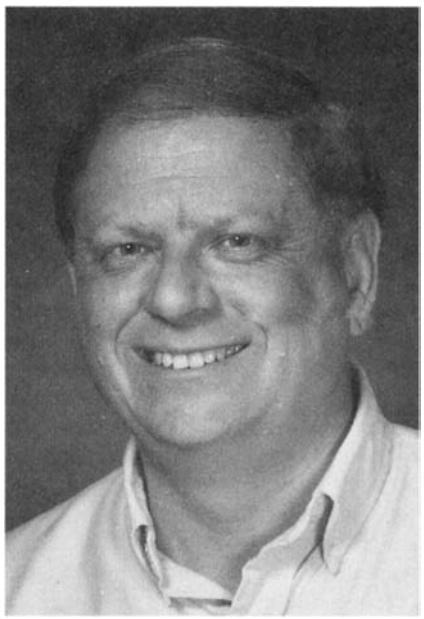
Second, the analogy is guided by a pressure to identify consistent structural parallels between the roles in the source and target domains (Gentner, 1983). The key structural constraint underlying analogical mapping and inference is a pressure to establish an *isomorphism*—a set of consistent, one-to-one correspondences—between the elements of the source and the target. Thus, once Aaron had decided to place the source and target "injuries" into correspondence (based on similarity of relations), structural consistency required that the person injured in the source (the child) be mapped onto the person injured in the target (the mother), because each was playing the same relational role. In this case, the constraint of maintaining consistent correspondences apparently dominated the rival similarity constraint, which by itself would have encouraged mapping the mother to herself. In the subsequent inference stage, consistency further required that it be the child in the target (now mapped to the mother in the source) who provided the soothing kiss.

Third, the constraint of purpose implies that analogical thinking is guided by the reasoner's goals—what the analogy is intended to achieve. Why did Aaron even consider the analogy with the kissing ritual? It appears that his mother's expression of pain gave rise to the goal of alleviating it; this goal in turn caused the child's attention to focus on those aspects of the target situation that were relevant to achieving a solution. Once his attention was biased so as to favor goal-relevant aspects of the situation, Aaron was led to access source analogs involving injuries rather than, for example, earlier instances of being dressed by his mother.

These three kinds of constraints—similarity, structure, and purpose—do not operate like rigid rules dictating the interpretation of analogies. Instead, they function more like the various pressures that guide an architect engaged in creative design, with some forces in convergence, others in opposition, and their constant interplay pressing toward some satisfying compromise that is internally coherent. When we describe computational models of analogy, we suggest how such local contradictions between constraints can be resolved by a process of constraint satisfaction. First, however, we briefly review some examples of experimental tests that reveal the operation of the constraints in the analogical thinking of college students.

Ambiguity in Mapping: Comparing the Persian Gulf War With World War II

The analogy between the Persian Gulf War and World War II, which in 1991 figured prominently in debates about whether the United States should make a military response to Iraq's invasion of Kuwait, provides a striking historical example of the role of analogy in shaping public opinion. In addition, this analogy illustrates the interactions between the multiple constraints that guide construction of a mapping. During the first two days of the counterattack in January 1991, Spellman and Holyoak (1992) asked a group of undergraduates at the University of California, Los Angeles, a few questions to find out how they interpreted the analogy between the Persian Gulf situation and World War II. The two situations were by no means completely isomorphic; in fact, the analogy was messy and ambiguous. On the one hand, similarity at the object level favored mapping the United States of 1991 to the United States of World War II, simply because it was the same country, which would in turn support mapping President George Bush to President Franklin Delano Roosevelt. On the other hand, the United States did not go to war until it was bombed by Japan, well after Adolf Hitler had marched through much of Europe. Therefore, one might argue that the United States of 1991 mapped to Great Britain of World War II and that Bush mapped to Winston Churchill (because Bush, like Churchill, led his nation and Western allies in early opposition to aggression). However, other relational similarities supported mappings to the United States and Roosevelt; for example, the United States was the major supplier of arms and equipment for the allies, a role parallel



Paul Thagard
Photo by University of Waterloo Central Photographic.

to that played by the United States in the Persian Gulf situation. These conflicting pressures made the mappings ambiguous.

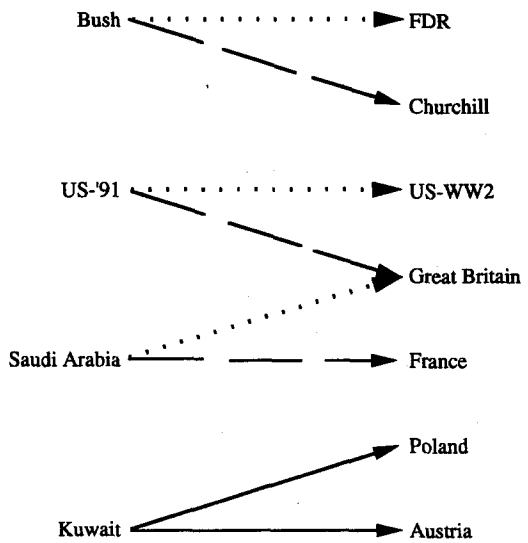
The pressure to maintain structural consistency—a central component of our multiconstraint theory—implies that people who mapped the United States to Great Britain should also have tended to map Bush to Churchill whereas those who mapped the United States to the United States should instead have mapped Bush to Roosevelt. Notice this is not a logical requirement—a person might think the United States should map to itself but Bush should map to Churchill because each was the dominant leader of the war effort. Indeed, nothing prevented people from giving both mappings as answers. However, the multiconstraint theory would predict that people should prefer one-to-one mappings—Bush to either Churchill or Roosevelt, but not to both—and mappings that maximize structural consistency by keeping leaders and their countries together.

At the same time, the multiconstraint theory allows the possibility of mappings that violate the one-to-one constraint when enough evidence favors multiple mappings. For example, several European nations (Austria, Czechoslovakia, and Poland) were targets of German aggression prior to the outbreak of World War II, and Kuwait might be mapped to more than one of them. Similarly, Saudi Arabia played a role somewhat similar to that played by Great Britain in World War II (a staging area for the counterattack and a target of missile attacks) and also somewhat similar to that played by France (under threat from Germany at the time Great Britain responded to the invasion of Poland).

The undergraduates were asked to suppose that Saddam Hussein was analogous to Hitler. Regardless of whether they thought the analogy was appropriate, they were then asked to write down the most natural match in the World War II situation for Iraq, the United States, Kuwait, Saudi Arabia, and President George Bush. For those students who gave evidence that they knew the basic facts about World War II, the majority produced mappings that fell into one of two patterns, as depicted in Figure 1. Those students who mapped the United States to itself also mapped Bush to Roosevelt; these same students also tended to map Saudi Arabia to Great Britain. Other students, in contrast, mapped the United States to Great Britain and Bush to Churchill, which in turn (by the one-to-one constraint) forced Saudi Arabia to map to some country other than Great Britain (usually France). The mapping for Kuwait (which did not depend on the choice of mappings for Bush, the United States, or Saudi Arabia) was usually to one or two of the early victims of Germany in World War II, usually Austria or Poland (or to a grouping such as "countries Hitler took over").

The analogy between the Persian Gulf situation and World War II thus generated a bistable mapping: People tended to provide mappings based on one of two coherent but mutually incompatible sets of correspondences. Spellman and Holyoak (1992) went on to perform a second study, using a different group of undergraduates, to show that people's preferred mappings could be changed by manipulating their knowledge of the source analog,

Figure 1
Bistable Mapping



Note. If Bush is FDR (Franklin Delano Roosevelt) then the US-'91 (United States during the Persian Gulf War) is the US-WW2 (United States during World War II) and Saudi Arabia is Great Britain; if Bush is Churchill then the US-'91 is Great Britain and Saudi Arabia is France. (Dotted lines indicate mappings that result when Bush is mapped to FDR; large dashed lines indicate mappings that result when Bush is mapped to Churchill. Solid lines represent mappings that are constant regardless of the mapping for Bush.) From "If Saddam Is Hitler Then Who Is George Bush? Analogical Mapping Between Systems of Social Roles," by B. A. Spellman and K. J. Holyoak, 1992, *Journal of Personality and Social Psychology, 62,* p. 917. Copyright 1992 by the American Psychological Association. Reprinted with permission.

World War II. Because many undergraduates were lacking in knowledge about the major participants and events in World War II, it proved possible to "guide" them to one or the other mapping pattern by having them first read a slightly biased summary of events in World War II. The various summaries were all historically "correct," in the sense of providing only information taken directly from history books, but each contained somewhat different information and emphasized different points. Some versions emphasized the personal role of Churchill and the national role of Great Britain; other versions placed greater emphasis on what Roosevelt and the United States did to further the war effort. After reading one of these summaries of World War II, the undergraduates were asked the same mapping questions as those used in the previous study. The same bistable mapping patterns emerged as before, but this time the summaries influenced which of the two coherent patterns of responses students tended to give. People who read a "Churchill" version tended to map Bush to Churchill and the United States to Great Britain, whereas those who read a "Roosevelt" version tended to map Bush to Roosevelt and the United States to the United States. Even summaries that had been written to support a crossed mapping (e.g., making Churchill the most important leader but the United States the most important country) tended instead to produce one of the two patterns in which the mapping kept the leader and his country together. It appears that even when an analogy is messy and ambiguous, the constraints on analogical coherence produce predictable interpretations of how the source and the target fit together.

Processing Goals and the Resolution of Ambiguity

The research we have described so far demonstrates the impact of both structure and similarity on mapping. What about the purpose of the analogy? It is generally accepted that people seek and use analogies to achieve their goals. However, it has been less clear whether the purpose can actually change the mappings that people generate, rather than just the initial selection of a source or the later adaptation of a solution. One way to investigate this issue is to have people draw analogies between situations for which the mapping is ambiguous and then see if their goals alter people's preferred mappings. Spellman and Holyoak (1996, Experiment 3) performed an experiment of this sort in which college students were asked to map the characters in two soap-opera plots. The students were told to pretend that they were writers of a successful new soap opera and that they were in court trying to prove that writers from another soap opera had stolen their ideas. Each soap opera involved the entanglements of multiple characters. In the first soap opera, which was set at a university, an ex-alcoholic professor named Peter was in love with his research assistant, Mary, and had cheated his brother out of his inheritance. These characters were connected by three types of relations: professional (Peter was Mary's boss), romantic (Peter was in love with Mary), and inheritance (Peter cheated his brother). The second soap opera was set in a city and involved two fairly distinct sets of characters. The "lawyer set" included Nancy, an ex-addict entertainment lawyer, and John, a young lawyer working at Nancy's law firm who had often filled in for her. The "doctor set" included David, a prominent physician who had become an alcoholic, and Lisa, an intern who was now treating most of David's patients. Nancy and David were half-siblings, and John and Lisa were cousins. Both pairs had aging relatives ready to leave them money in a will. In one version of the story, Nancy and Lisa (the women) cheated David and John (the men), respectively, out of their shares of the inheritance; in the other version of the story, the men cheated the women out of their shares of the inheritance. From this description, the object mappings were ambiguous; for example, if the women were the cheaters, then Peter would seem to map equally to Nancy and Lisa.

To manipulate the purpose of using the analogy, the students were told that the judge in the plagiarism trial wanted them to predict what would happen in the next episode of the city soap opera. If they could figure out who would do what to whom in the episode, this would be solid proof that the writers of the city soap opera had plagiarized ideas from the university soap opera. Half of the students were told that the crucial episode was "just like" what had happened in an episode of the university soap opera in which Peter had tried to steal credit for Mary's ideas. For these students, the professional relations between the characters were most important for the plot development. The other half of the students were asked to predict an episode "just like" one in which Peter had tried to seduce Mary, in which case the romantic relations were critical. The inheritance relations did not play any direct role in either of the two episodes. After they had written extensions of the plot, all of the students were asked to select the best match for each character in the university soap opera from among the characters in the city soap opera. Thus, the experiment measured people's preferred mappings both indirectly by which characters were used to extend the plot and directly by the mapping task.

So, which characters in the city soap opera correspond to Peter, the professor, and Mary, his assistant? Without taking the goal into account, the mapping is actually ambiguous in four ways, as schematized in Table 1. The basic ambiguity is that Peter is somebody's boss, as are Nancy and David, and he pursues someone, as do John and Lisa. But consider how the mapping would be expected to shift if people placed greater weight on the relations that were most pragmatically central in extending the plot to predict the crucial new episode. If the episode hinged on the professional relations, then Peter would seem more like Nancy or David (the bosses) than like John and Lisa (the underlings). Suppose that one were dealing with the version of the city soap opera in which the women cheated the men out of their inheritance. If people placed at least some weight on the incidental inheritance relations, the mapping of Peter to Nancy would be preferred over the mapping to David

Table 1
Optimal Mappings for the Source Characters Based on Pragmatic Manipulation and Gender of Cheater in the Inheritance Relation

Professional plot extension				Romantic plot extension		
Gender of cheater					Gender of cheater	
Male	Female	Role	Source characters	Role	Male	Female
David	Nancy	Boss	Peter	Pursuer	John	Lisa
Lisa	John	Underling	Mary	Pursued	Nancy	David

Note. From "Pragmatics in Analogical Mapping," by B. A. Spellman and K. J. Holyoak, 1996, *Cognitive Psychology, 31*, p. 328. Copyright 1996 by Academic Press. Adapted with permission.

(because Nancy, like Peter, cheated someone out of an inheritance). To be consistent with this mapping for Peter, Mary would then be mapped to John. (Similarity of the characters' gender was controlled in the experiment by counterbalancing and therefore is ignored in our discussion.)

Now consider the situation from the point of view of someone who had to predict the plot focusing on the romantic relations. Peter would now map best to either John or Lisa, who shared the role of pursuer. Of these two possibilities, the mapping to Lisa would be preferred if people were sensitive to the inheritance relations as well as the romantic relations. Consistency would then tend to make Mary map to David.

By seeing how the students actually mapped Peter and Mary as a pair, it can be determined whether their mappings were sensitive to the students' purpose in using the analogy. Those students who gave greater emphasis to the type of relation that was pragmatically important for extending the plot (either the professional or the romantic relations) would give one of the two mappings consistent with the goal-relevant relations. In addition, those students who also gave at least some weight to the inheritance relations—even though these were not relevant to the plot extension—would select a mapping in which Peter mapped to a cheater.

The top panel of Figure 2 displays the results for the plot-extension task. The majority of the students developed a sensible plot extension in which Peter and Mary mapped consistently to one of the two character pairs that matched on the important type of relation. Of these two possibilities, there was a weak preference for the pair that also matched on the incidental inheritance relations. The goal clearly had a strong inference on people's choice of characters.

The bottom panel of Figure 2 displays the results for the explicit mapping task. This task, unlike the plot-extension task, did not actually require people to focus on the type of relation needed to write the new episode. Nonetheless, people may have continued to give greater weight to whichever type of relation had been relevant to their goal. And, in fact, people preferred to map Peter and Mary on the basis of the goal-relevant relation rather than the opposing relation, although this preference was weaker than it was in the plot-extension task. In addition, people tended to prefer a mapping that was consistent with the inheritance relations.

Notice that in both the plot-extension and explicit mapping tasks, the majority of the students mapped Peter and Mary to some consistent pair of characters (i.e., two people who interacted with each other) rather than splitting the mapping in some way (the "other" responses). Spellman and Holyoak's (1996) experiment thus shows that people are sensitive to all three of the basic constraints we have been talking about—structure (making consistent mappings for the pair of characters), similarity (mapping professional relations to professional relations and romantic relations to romantic relations), and purpose (resolving ambiguous mappings on the basis of whichever type of relation is most relevant to the person's goal in using the analogy).

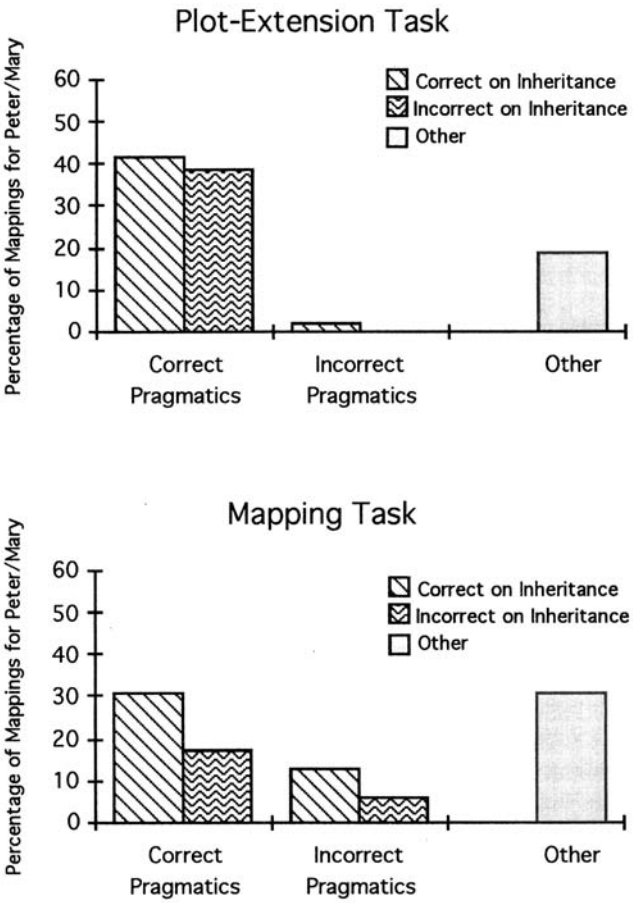
Analogical Thinking in Everyday Life

People's sensitivity to analogical constraints of similarity, structure, and purpose is vividly exhibited outside the laboratory. Conventional wisdom has it that people at the beginning of their careers—a young professional in training, such as a medical student, an articling lawyer, or a graduate student working on a doctoral dissertation—can benefit substantially from having role models. Consider Jane, an intelligent, hard-working student who aspires to be a clinical psychologist. She may encounter an older, successful clinical psychologist on whom she can model, at least partly, her own choices concerning her career and personal life. Using a role model in this way is a kind of analogical thinking, in which Jane's own career is the target problem and the role model's career becomes a potential source of insight.[2]

We are not aware of any empirical research that has addressed the question of how people choose their role

[2] We are grateful to Ziva Kunda for ideas about role models as analogies.

Figure 2
Resolving Ambiguous Mappings



Note. Top: Percentages of participants in the plot-extension task of Experiment 3 of Spellman and Holyoak's (1996) study who mapped characters in accord with the goal-relevant relation (either professional or romantic) and in accord with the incidental inheritance relation. Other = participants who did not write an analogous plot extension or whose analogous plot extension included two sets of characters. Bottom: Percentages of participants in the mapping task of Experiment 3 of Spellman and Holyoak's (1996) study who mapped characters in accord with the goal-relevant relation (either professional or romantic) and in accord with the incidental inheritance relation. Other = participants who did not write an analogous plot extension, included two sets of characters in the plot extension, or did not map to a congruous Peter/Mary pair. From "Pragmatics in Analogical Mapping," by B. A. Spellman and K. J. Holyoak, 1996, *Cognitive Psychology, 31*, p. 332. Copyright 1996 by Academic Press. Reprinted with permission.

models and apply them to inform their lives. From the perspective of the multiconstraint theory of analogy, choosing a role model is a kind of analog selection, and applying the role model is a kind of analogical mapping. For example, Jane may select her role model by remembering an established clinical psychologist who is similar to her in many respects such as race, gender, and personality. We conjecture that, like analogical retrieval, role-model selection is primarily dominated by such salient similarities, although subtler aspects of relational structure (e.g., how the older psychologist has been connected with other people and institutions) and purpose (e.g., Jane's career and personal goals) may also affect her selection of a role model. Once she has identified a role model, however, Jane's analogical thinking will be much more affected by structure and purpose than by more superficial similarities. To make decisions in her own life (e.g., whether to emphasize therapy or research, whether to marry another psychologist, or whether to have a baby in graduate school), Jane may be able to take into account the positive and negative results of similar decisions in her role model's life. Jane's thinking might implicitly proceed along the following lines: "I don't know whether I should put a lot of energy into my PhD research or get more clinical experience. My role model, Alice, is similar to me in that she is a woman who is interested in both therapy and research. She completed a fine PhD thesis that yielded several publications, which helped her get a good placement that started her off on a very successful career as a clinical psychologist. So maybe I should also see research as furthering my career aspirations." At this point in Jane's thinking, the fact that Alice is also a woman will be important to the extent that the structure of her life maps onto Jane's and suggests to Jane how she might accomplish her goals. For example, Alice's being a woman may turn out to affect structure and purpose if Jane's situation involves gender-related impediments to accomplishing career goals.

If Jane becomes a practicing clinical psychologist, she may find herself noticing frequent use of analogies by both her clients and herself. According to Meichenbaum (1994), victims of posttraumatic stress disorder frequently use analogies and metaphors to describe their own situations. Here are some of the metaphors used by people recovering from severe psychological trauma:

- I am a time bomb ticking, ready to explode.
- I feel like I am caught up in a tornado.
- I am a rabbit stuck in the glare of headlights who can't move.
- My life is like a rerun of a movie that won't stop.
- I feel like I'm in a cave and can't get out.
- Home is like a pressure cooker.
- I am a robot with no feelings. (pp. 112–113)

All of these examples involve analogical mappings from a familiar situation to the situation of the patient. (For a discussion of the relationship between metaphor and analogy, see Holyoak & Thagard, 1995, chap. 9.)

Meichenbaum (1994) described how changes in metaphors used by clients can mark improvements in their conditions. Recovering trauma victims replace metaphors such as those in the list in the previous paragraph by metaphors such as the following:

- One door closes and another opens.
- I want to be the author of my own stories.
- Get back in the driver's seat.
- Put a new coin in my juke box and play a new tune.
- I want to move out of whirlpools and into still waters. (p. 115)

The use of these hopeful metaphors for clients' problems suggests that people can map themselves to persons or

situations in ways that suggest solutions to personal problems. Recognition of patients' changing metaphors can therefore be a useful part of Jane's clinical practice. There is, as yet, no experimental evidence that metaphor change plays a causal role in patients' improvements, but clinical observations of patients suggest that metaphor change is an integral part of healing, not just a reflection of emotional states before and after treatment.

Jane is likely to find herself not only noticing her patients' analogies but also using them herself. One of the major predictors of therapeutic success is *empathy*, the extent to which Jane is able to understand her clients' emotions by putting herself in the clients' shoes and getting a sense of how she would feel if she were in the clients' situations (Dawes, 1994). Barnes and Thagard (in press) argued that empathy is essentially a process of analogical mapping, in which the empathizer is able to produce a structured comparison that results in transfer, not just of verbal information but also of an emotional state. For example, if Jane's client is a rape victim, she can imagine how she would feel if she had experienced the same trauma and then map this emotion back to the client in order to understand more deeply the ongoing distress. Empathic therapy may also involve analog retrieval or construction, when therapists have to work hard to find situations in their own lives that are semantically and structurally similar to those of the clients.

The use of empathy in therapy involves an analogy between the client and the therapist, but the therapist may also find useful analogies between the client and someone else, perhaps a previous client with a similar problem. Such a mapping may help the therapist better understand the current client and also may be used to provide a role model that the client can use as a source analog to suggest steps toward recovery. Barker (1985) provided additional examples of the use of analogies and metaphors in psychotherapy.

In addition to her life as a therapist, Jane's career as a researcher may benefit from analogical thinking. When she designs her experiments, she may look carefully at experiments already done in the area in which she has an interest. An experiment conducted previously may provide a source analog to suggest, in part, how she should structure her own experiments. (See Dunbar, in press, for analyses of the use of analogy in working laboratories within the field of molecular biology.) Which experiments Jane selects to guide her own design, and how she maps them to produce her own experiment, should depend primarily on structure and purpose, although other similarities may get carried over as well.

Theoretical ideas in psychology and other fields often arise by analogy from related fields. In current cognitive science, the analogy between thinking and computation is the major source of theories of mind (Thagard, 1996). Conceiving the mind as a rule-based computer, a neural-networklike computer, or a chaotic computer makes possible precise specification of mental mechanisms that may explain people's psychological capacities. Current theories of analogical thinking have been heavily influenced by such computational analogies.

Computational Models of Analogical Coherence

Our development of the multiconstraint theory has heavily depended on computational models designed to simulate aspects of human analogical thinking. ARCS (Analog Retrieval by Constraint Satisfaction; Thagard, Holyoak, Nelson, & Gochfeld, 1990) and ACME (Analogical Mapping by Constraint Satisfaction; Holyoak & Thagard, 1989) address the steps of access and mapping, respectively. ACME has also been extended to make inferences based on the mappings it computes (Holyoak, Novick, & Melz, 1994). As their names imply, these systems, in essence, attempt to find the optimal fit to the constraints postulated by the multiconstraint theory. The models attempt to make use of the strengths of both the symbolic and connectionist approaches to modeling cognition (Barnden, 1994), combining symbolic representations of explicit knowledge with connectionist processing.

The structures in a parallel constraint-satisfaction model consist of elements and various kinds of constraints among them. One can classify constraints as being either internal or external: Internal constraints involve relations only among the elements, whereas external constraints come from outside the system of elements. In addition, constraints can be either positive or negative, depending on whether they imply that two elements are compatible or incompatible.

We illustrate the general approach by focusing on the ACME model of analogical mapping, which specifies how the constraints of similarity, structure, and purpose can be jointly optimized to yield a coherent set of correspondences between a source and a target (see Holyoak & Thagard, 1989, for a full description). Consider a simplified version of the Persian Gulf analogy that includes only the information that Saddam Hussein was president of Iraq, which invaded Kuwait, and Hitler was führer of Germany, which occupied Austria:

Target	Source
president of (Saddam, Iraq)	führer of (Hitler, Germany)
invade (Iraq, Kuwait)	occupy (Germany, Austria).

Considered in isolation, the mappings for these fragments are obvious, but in the context of more realistic representations of people's knowledge about the two wars, the computational difficulty of the task is apparent. The ACME model shows how multiple constraints make mapping possible.

If we focus on structure, we can constrain the mapping problem considerably by mapping predicates to predicates and objects to objects, so that the correspondence *invade* ↔ *Hitler* would never be considered. The elements in our constraint-satisfaction theory of analogical mapping include only hypotheses that relate analog components of similar types: predicate–predicate hypotheses such as *invade* ↔ *occupy* and *invade* ↔ *führer of* and object–object hypotheses such as *Saddam* ↔ *Hit-*

ler and *Saddam* ↔ *Germany.* We can also ignore hypotheses that involve objects that never fill corresponding slots, such as *Saddam* ↔ *Austria.*

Among the hypotheses worth considering, two additional kinds of structural constraints can be applied: the positive constraint of structural consistency and the negative constraint of one-to-one mapping. For example, structural consistency requires that the hypothesis *invade* ↔ *occupy* should encourage and be encouraged by the mappings *Iraq* ↔ *Germany* and *Kuwait* ↔ *Austria.* Similarly, one-to-one mapping requires that the hypothesis *Iraq* ↔ *Germany* should discourage and be discouraged by *Iraq* ↔ *Hitler.* In ACME, structural consistency and one-to-one mappings are both "soft" constraints, encouraging mappings but not insisting on them, whereas ruling out mappings between objects and predicates is a hard, inviolable constraint.

Both similarity and purpose are treated as external constraints on mapping. We want to favor mappings that involve semantically similar components such as *invade* and *occupy,* not ones involving elements as different as *invade* and *führer of.* Again, this is a soft constraint, as we want the system to be able to discover correspondences between elements that were not previously seen as related to each other. Similarly, the purpose favors mapping hypotheses that fit with the goals of the analogist: If the point of the analogy is to show that Saddam is evil like Hitler, then the mapping hypothesis *Saddam* ↔ *Hitler* would be encouraged by the soft constraint that mappings should serve the purpose of the analogy.

Now we can move from the constraint theory to the computational model, in which (a) elements are represented by units; (b) positive and negative constraints are represented by excitatory and inhibitory links, respectively; (c) external constraints are represented by links to special units; and (d) parallel constraint satisfaction is achieved by algorithms for updating activations of the units on the basis of their links to other units. In the simple example above, 11 units are needed to represent all the mapping hypotheses. (For simplicity, we ignore hypotheses about mappings between propositons.) These units will be interconnected by excitatory and inhibitory links to represent the positive constraint of structural consistency and the negative constraint of one-to-one mapping. To implement the external constraints, we need two special units, one for semantic similarity and the other for purpose. A special unit will be linked with each unit that represents a mapping hypothesis that satisfies the constraints of either semantic similarity or relevance to purpose (or both). For example, ACME produces a link from the special "similarity" unit to the unit representing the *president of* ↔ *führer of* correspondence but not the *president of* ↔ *occupy* correspondence. Of course, satisfying a constraint can be a matter of degree. For example, the concept of being a president is somewhat similar to that of being a führer but perhaps less so than to that of, say, being a prime minister. The magnitude of the positive or negative weight on each link reflects the degree to which the corresponding constraint is satisfied or violated.

Figure 3 depicts the network created by ACME when it is given as input the source and the target represented above. Once this network is created, a simple "relaxation" algorithm updates the activation of each unit in parallel to determine which mapping hypotheses should be accepted (see Rumelhart, Hinton, & McClelland, 1986). All units start with activations near zero, except for the special units for semantic similarity and purpose, which start with and retain full activation of one. These units start to activate the units with which they are linked; then, activation spreads throughout the system, fostered by excitatory links and suppressed by inhibitory links.

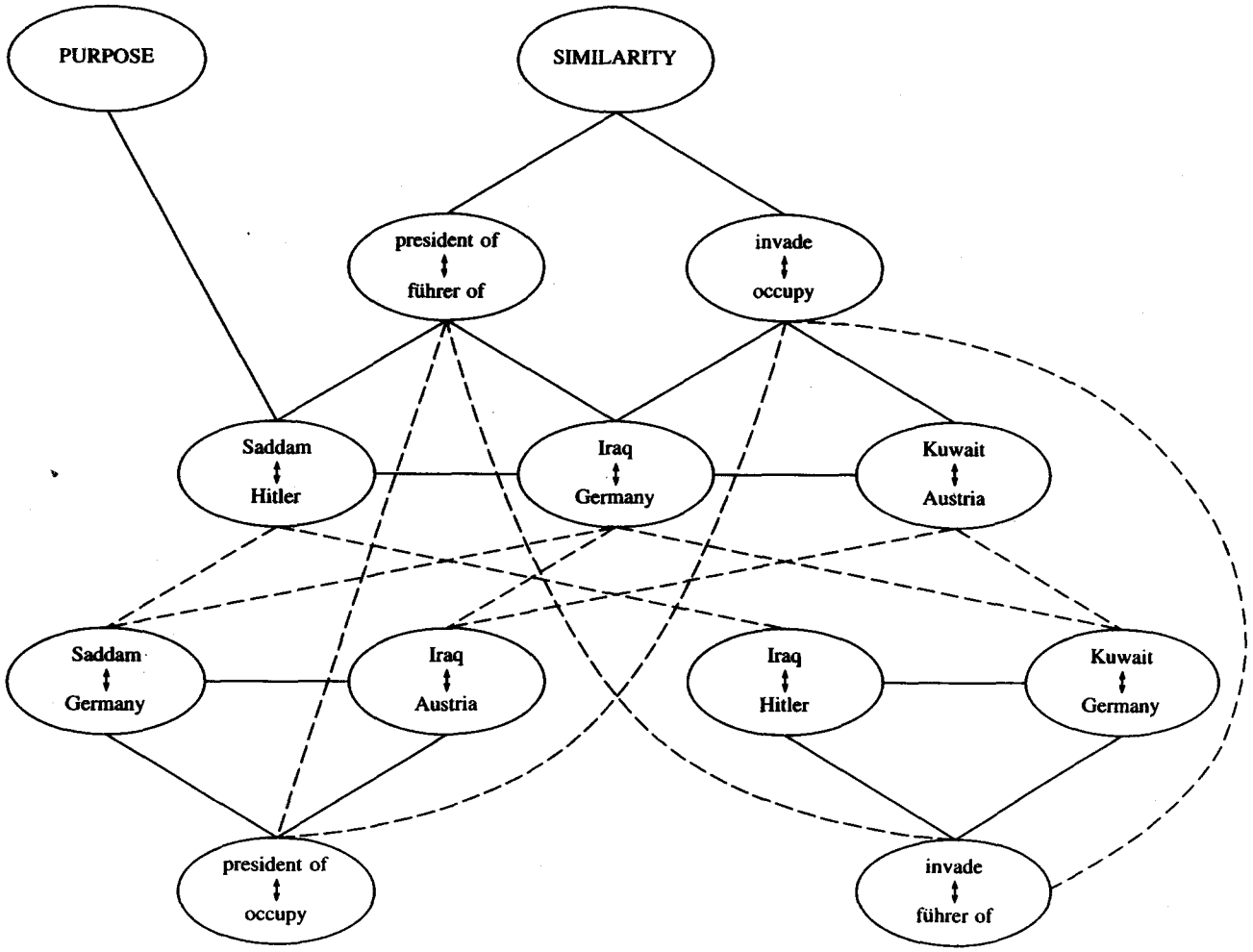
ACME is capable of exploiting higher order relations (i.e., relations that take propositions as arguments; see Gentner, 1983) to provide much deeper mappings. An enhanced representation of the Persian Gulf target and the World War II source might include the information that Saddam Hussein's being president of Iraq was a cause of Iraq invading Kuwait, just as Hitler's being führer of Germany was a *cause* of Germany occupying Austria. ACME would then map the two cause relations together and create additional mapping hypotheses, putting entire propositions into correspondence with each other.

We have tested ACME on dozens of examples and have showed how ACME can closely mimic human mapping behavior in a variety of psychological experiments. For example, ACME can find human-like mappings between the Persian Gulf and World War II analogs when propositions that capture an elaborate summary of each are given (Spellman & Holyoak, 1992). Like humans, ACME is sensitive to the "Necker-cube" quality of this ambiguous analogy, settling into one of two sets of coherent but mutually exclusive correspondences: President Bush and the United States tend to be mapped to Roosevelt and the United States, or to Churchill and Great Britain, but not to a mixed combination of a leader and a country, such as Churchill and the United States. Also like people, ACME occasionally tolerates a one-to-many mapping, such as that between Kuwait and Austria/Poland. Similarly, ACME is able to simulate the manner in which the processing goal guides the resolution of ambiguous mappings (Spellman & Holyoak, 1996). In general, the model seems to capture the human ability to find coherent relationships between complex and imperfectly understood situations on the basis of the interplay between the constraints of structure, similarity, and purpose.

Future Directions

Despite ACME's successful simulation of many analogies, we do not believe that it or other current computational models provide the final word on analogical thinking. Human use of analogies and metaphors still far surpasses existing computational models in semantic richness and flexibility of application. One promising

Figure 3
Structure of the Network Created by ACME for the Simplified Saddam Hussein Example



Note. Solid lines indicate excitatory links, and dashed lines indicate inhibitory links. Units representing mappings between whole propositions are not shown. ACME = Analogical Mapping by Constraint Satisfaction. From *Mental Leaps: Analogy in Creative Thought* (p. 250), by K. J. Holyoak and P. Thagard, 1995, Cambridge, MA: MIT Press. Copyright 1995 by MIT Press. Reprinted with permission.

direction for future progress in computational understanding of analogy involves the use of distributed representations of meaning. ACME is a localist connectionist model in which each neuron-like unit represents a mapping hypothesis linking pairs of predicates and objects. A distributed representation, in contrast, codes predicates, objects, and propositions by complexes of units, just as concepts seem to be distributed over a large set of neurons in the brain. We are currently exploring different ways of introducing distributed representations into our analogy models and finding that they do indeed enable greater flexibility than ACME affords (Eliasmith & Thagard, 1996; Hummel & Holyoak, in press). Models based on distributed representations can capture more subtle interactions among the constraints on analogical thinking. In addition, distributed representations make it much easier to understand the connection between analogical thinking and learning of abstract schemas. These newer models are nonetheless instantiations of the multiconstraint theory of analogy, as they perform analogical mapping using the three basic constraints of similarity, structure, and purpose. We anticipate that continued development of more sophisticated computational models will lead to a deeper understanding of the analogical mind.

REFERENCES

Barker, P. (1985). *Using metaphors in psychotherapy.* New York: Brunner/Mazel.

Barnden, J. A. (1994). On the connectionist implementation of analogy and working memory matching. In J. A. Barnden & K. J. Holyoak (Eds.), *Advances in connectionist and neural computation theory:*

Vol. 3. Analogy, metaphor, and reminding (pp. 327–374). Norwood, NJ: Ablex.

Barnes, A., & Thagard, P. (in press). Empathy and analogy. *Dialogue.*

Dawes, R. (1994). *House of cards: Psychology and psychotherapy built on myth.* New York: Free Press.

Dunbar, K. (in press). How scientists think: On-line creativity and conceptual change in science. In T. B. Ward, S. M. Smith, & J. Vaid (Eds.), *Conceptual structures and processes: Emergence, discovery, and change.* Washington, DC: American Psychological Association.

Eliasmith, C., & Thagard, P. (1996). *Integrating structure and meaning: A distributed model of analogical mapping.* Unpublished manuscript, University of Waterloo, Waterloo, Ontario, Canada.

Gentner, D. (1977). If a tree had a knee, where would it be? Children's performance on simple spatial metaphors. *Papers and Reports on Child Language Development, 13,* 157–164.

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

Gillan, D. J., Premack, D., & Woodruff, G. (1981). Reasoning in the chimpanzee: I. Analogical reasoning. *Journal of Experimental Psychology: Animal Behavior Processes, 7,* 1–17.

Goswami, U., & Brown, A. (1989). Melting chocolate and melting snowmen: Analogical reasoning and causal relations. *Cognition, 35,* 69–95.

Holyoak, K. J., Junn, E. N., & Billman, D. (1984). Development of analogical problem-solving skill. *Child Development, 55,* 2042–2055.

Holyoak, K. J., Novick, L. R., & Melz, E. R. (1994). Component processes in analogical transfer: Mapping, pattern completion, and adaptation. In K. J. Holyoak & J. A. Barnden (Eds.), *Advances in connectionist and neural computation theory: Vol. 2. Analogical connections* (pp. 113–180). Norwood, NJ: Ablex.

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., & Holyoak, K. J. (in press). Distributed representations of structure: A theory of analogical access and mapping. *Psychological Review.*

Inagaki, K., & Hatano, G. (1987). Young children's spontaneous personification as analogy. *Child Development, 58,* 1013–1020.

Keane, M. T. (1986). On retrieving analogues when solving problems. *Quarterly Journal of Experimental Psychology, 39A,* 29–41.

Meichenbaum, D. (1994). *A clinical handbook/practical therapist manual for assessing and treating adults with post-traumatic stress disorder (PTSD).* Waterloo, Ontario, Canada: Institute Press.

Ross, B. (1989). Distinguishing types of superficial similarities: Different effects on the access and use of earlier problems. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 15,* 456–468.

Rumelhart, D. E., Hinton, G. E., & McClelland, J. L. (1986). A general framework for parallel distributed processing. In D. E. Rumelhart, J. L. McClelland, and the PDP Research Group (Eds.), *Parallel distributed processing: Vol. 1. Foundations* (pp. 45–76). Cambridge, MA: MIT Press.

Seifert, C. M., McKoon, G., Abelson, R. P., & Ratcliff, R. (1986). Memory connections between thematically similar episodes. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 12,* 220–231.

Spellman, B. A., & Holyoak, K. J. (1992). If Saddam is Hitler then who is George Bush? Analogical mapping between systems of social roles. *Journal of Personality and Social Psychology, 62,* 913–933.

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

Thagard, P. (1996). *Mind: Introduction to cognitive science.* Cambridge, MA: MIT Press.

Thagard, P., Holyoak, K. J., Nelson, G., & Gochfeld, D. (1990). Analog retrieval by constraint satisfaction. *Artificial Intelligence, 46,* 259–310.