Analogy and Transfer: Encoding the Problem at the Right Level of Abstraction

Emmanuel Sander (sander@univ-paris8.fr)
University Paris 8, Department of Psychology, 2 Rue de la Liberté
Saint-Denis Cedex 02, 93526 France

Jean-François Richard (richard@univ-paris8.fr)
University Paris 8, Department of Psychology, 2 Rue de la Liberté
Saint-Denis Cedex 02, 93526 France

Abstract

Analogy and transfer studies belong to distinct research areas regarding both the theoretical frameworks and the experimental paradigms, although they both contribute to the fundamental question of the generalization of learning. The aim of this paper is to show that those fields might be reanalyzed within a common framework, emphasizing the role of the abstraction level at which the problem is encoded, which might contribute to an unified theory of generalization of learning.

Analogy and transfer

Analogy and transfer of learning, which are two distinct areas in cognitive psychology as the lack of cross references between the two fields illustrates, are both concerned with the same basic questions: how can what has been learned in a specific situation be generalized to new situations and what conditions are necessary and/or sufficient to generate transfer from one situation to another? We shall attempt to reanalyze both phenomena within the same theoretical framework and show how this analysis contributes to the fundamental question of the generalization of learning.

Differences and similarities in the study of analogy and transfer.

The prevailing theory of transfer, proposed originally by Thorndike and Woodworth (1901), is transfer between the learning and the test situations relies on common elements. This theory has more recently been revived in the contexts of learning complex devices and programming languages. In these contexts, the common elements are production rules, which have both a stimulus side (their conditions) and a response side (the action triggered by the conditions). The idea that the degree of transfer between two tasks is proportional to the relative number of productions rules they have in common, each task being described by a set of production rules corresponding to the allowable procedures, has received substantive confirmation in text editing and computer programming (Anderson & Singley, 1993; Bovair, Kieras & Polson, 1990).

On the other hand, the analogy theories have been developed in the last two decades mainly through the works of Gentner (1983, 1989) and Holyoak (e.g. Holyoak & Thagard, 1989; Hummel & Holyoak, 1997, 2003) and their co-workers. While differing in many respects, both theories decompose the analogy process into several basic constituent processes (access, mapping, inference, evaluation, generalization) and distinguish several kinds of similarities (pragmatic, semantic, superficial, structural), which govern analogy process differently depending on the constituent at hand. Theories differ in the kind of similarities taken into account and in the way these similarities are involved within the constituent processes.

Several differences distinguish the experimental paradigms used between transfer and analogy studies. (i) Most of the problems in the analogy studies concern story problems that may be solved by pure reasoning and do not involve any physical action. The solution consists in the verbal statement of a procedure in order to attain the goal and not in the effective attainment of the goal (e.g. Ross, 1989). The reverse is true for transfer studies: solution is reached through physical actions from initial to final states of the problem (e.g. Kotovsky, Hayes & Simon, 1985). (ii) Due to the prominent use of verbal material in the analogy studies, the semantic content of the tasks is usually much richer and irrelevant features are more numerous than in the transfer studies; therefore the selection of features for mapping is more critical in analogy studies than in transfer studies. (iii) In analogy studies, the solution to the source problem is usually not worked out by the participant but provided by the instructions, as it is the case in traditional teaching situations, whereas in transfer studies, the solution is discovered by the participant; thus, what is transferred is more likely to be declarative knowledge in the analogy studies and procedural knowledge in transfer studies. (iv) In analogy studies, a problem, or several isomorphic problems, is (are) usually given only once, while in transfer studies, several trials may be given for the same task by varying the initial state and the goal. In this way it is possible to control the degree of prior training in transfer studies and to observe its effect on successive trials with the transfer problem.

In spite of these differences, many similarities may be emphasized in the way tasks are analyzed and in the results obtained. In both approaches, a sharp distinction is made between the elements that belong to the structure of the problem, and for this reason are relevant to the solution, and those that should be ignored. The theorists of analogy distinguish between the structural features that are relevant to
the task and the surface features that are not (Holyoak & Koh, 1987). In the transfer studies of problem solving, structure is defined through the notion of problem-space. This is an abstract structure that requires an embodiment, which results in irrelevant features being incorporated into the task (Kotovsky, Hayes & Simon, 1985).

This distinction has been used in both fields to control and analyze the effects of the context of the task on the difficulty of the solution. Three cases have been considered: (i) the source and the target problems have similar structural features but different surface features, (ii) both problems are similar with regards to both types of features or (iii) they are similar with regards to surface but not to structural features. Some converging phenomena were observed:

In the absence of surface similarity, transfer is seldom observed between the source and the target problem unless the participant is informed of the relevance of the first problem. This has been shown for the tumor problem among others in analogy studies, with the military problem of the attack on a fortress as the source problem (Gick & Holyoak, 1980,1983) as well as for the MC (missionaries and cannibals) problem and its isomorphs, with the Jealous Husbands problem (Reed, Ernst & Banerji, 1974) as the learning problem, in transfer studies.

In contrast, positive spontaneous transfer is obtained when the source and the target problems are similar as regards both content and structure. For instance, in analogy studies, Holyoak and Koh (1987) have observed positive transfer on the tumor problem (to be treated with X-rays) with an analog problem involving the division and convergence of laser rays. In transfer studies, massive transfer has been observed by Luger and Bauer (1978) from the tea ceremony problem, first studied by Hayes and Simon (1974), to the tower of Hanoi problem (TOH). Reed, Ernst and Banerji (1974) have observed the same result with an isomorphic version of the MC problem.

In cases where both problems are similar as to surface features but structurally different, transfer is negative: the procedure learned first is transferred but is inappropriate to the new problem as it has been show in analogy studies (Novick, 1988; Ross, 1989) and in transfer studies (Zamani & Richard, 2000).

The importance of the interpretation of the situation: a factor neglected in both approaches

Despite the numerous experimental results in both fields that confirm the heuristic value of distinguishing structural and surface features, an intriguing question originates from the fact that the distinction between structural and surface features is made from the point of view of the expert, but has no meaning for a participant who tries to solve the problem. This distinction is useful in planning experiments but its relevancy to exploring the processes involved in the recognition of analogy is questionable. Experiments that show no transfer between situations very similar regarding physical cues but different regarding semantic interpretation of these cues illustrate this issue. Kotovsky and Fallside (1989) devised two isomorphic TOH problems using two displays with identical stimulus that could be interpreted in two different ways. A change in the apparent size was presented either as a real change of the object or as an apparent change due to a change in distance of the object. Three colored spheres were presented on a graphic display inside a square box at the intersection of the diagonals and the size of the sphere could be changed by positioning a joystick and pressing one of three keys. In the size condition, the box was presented as a frame and the keys were labeled small, medium and large. In the depth condition the box was presented as a tunnel along which the sphere could be moved and the keys were labeled far, middle and near corresponding respectively to the small, medium and large sphere. In that way the same physical changes were interpreted either as changes in the size of the sphere either as moving the sphere from front to rear or vice-versa. Four groups of participants were given size-size, size-distance, distance-size, and distance-distance interpretations for the training and the transfer problems, respectively. A marked transfer was observed when both problems pertained to the same representation, but there was no transfer when the representation was different in spite of the fact that the display was exactly the same. In another experiment in the same study, three situations were considered: Peg Move, which is the standard TOH situation except that the three disks are replaced by three balls that are stacked. In the Dish Move problem, the material is the same, but as the balls are not stacked, it is not immediately apparent which ball is moveable. In the Dish Change problem, the situation is physically similar but the interpretation is different. A ball is placed on each dish and behind each dish there is a reserve dish containing the two balls which are different in size from the ball placed in the front dish. It is possible to exchange the ball placed in the front dish with a ball from the reserve dish and in this case the result of an exchange is a change in size of the ball in the front dish. The constraints for exchanging balls are the same as the constraints for changing the size of a globe in a change problem: they depend on the size of the balls in the other locations, instead of on the size of the balls in the same location, as in move problems. Interestingly, there is positive transfer from a Peg Move problem to a Dish Move problem and vice-versa but not from these two problems to the Dish Change problem. That means that although change of location and change of size should be conceived as a single operator, since they define the same transitions in the problem graph, the similarity of these operators is not perceived. This is a case in which the physical similarity of the actions is blurred by a difference in the semantic interpretation of these actions.

The same conclusions may be derived from Bassok, Wu and Olseth (1995) study who oppose what they call the object-mapping hypothesis to the interpreted structure hypothesis. In a study by Ross (1989), participants first learned to solve a permutation problem in which cars were randomly assigned to mechanics and then had to apply the equation in a transfer problem where scientists were randomly assigned to
computers. Although scientists and cars have similar roles (the randomly assigned set), participants erroneously instantiated the unknown value in the equation by matching scientists with mechanics and computers with cars. According to the object-mapping hypothesis, participants remember that one variable was associated with animate objects and the other with inanimate ones and therefore placed the scientists in the role of mechanics, because both are animate, and computers in the role of cars, because both are inanimate. Bassok, Wu and Olseth (1995) suggest that it is not the attribute similarity, animate versus inanimate, that is mainly responsible for the erroneous matching, but the semantic relation \(x \text{ gets } y\): Scientists get computers in the same way as mechanics get cars. Because the Ross (1989) results are compatible with both hypotheses, an experiment was devised using three sets of elements: carts, caddies and golfers, the same variable, caddies, being present both in the training and the transfer problems. All the participants initially learned to solve the permutation problem in which caddies are randomly assigned to golfers. In the transfer problem, for some participants, caddies were randomly assigned to carts, while for the remaining participants, carts were randomly assigned to caddies. According to the interpreted structure hypothesis, in the training problem participants infer that golfers get caddies to carry equipment and will place carts in the role that caddies had in the first problem since they consider that caddies get carts to carry equipment. The results are consistent with the hypothesis: performance is high when carts are assigned to caddies and very poor when caddies are assigned to carts.

These converging results support the view that the properties of objects in a situation are not considered in isolation but rather in relation to the other properties with which they are linked through prior knowledge related to them. These relationships between properties are as salient as the isolated properties detected through perception (Richard & Tijus, 1998; Sander & Richard, 1998). As such, the measurement of similarity is more complex than a measurement based on the properties that are present in the source and the target: it should include the interpretations that are inferred from the relationships between the properties as they are encountered in everyday life. The question “how do people recognize the analogy between a given situation and another situation?” becomes one of “what mechanisms make two interpretations compatible?” and “how do people change their interpretations of situations?” It means that if the initial encoding of a situation is not adequate, people must not only eliminate the non-relevant features but also change their representation. This process, called re-representation, is addressed in the following paragraph.

The generalization of learning

Re-representation

Until recently, the observation of failures to spontaneously retrieve a relevant source once few surface features are shared did not initiate studies aiming at identifying how such a failure may be overcome. Participants were led to change their representation of the task by means of explicit hints that the learning problem could be useful; The re-representation process was cut short. Initiating work on this point, Gentner and Medina (1998) propose that comparison promotes learning by inviting a re-representation of the situation. A recent body of research addresses this question explicitly and shows that the process of encoding, which had been somehow put aside so far in analogy studies, turns out to be actually crucial: Gentner, Loewenstein and Thomson (2004), as well as Kurtz and Lowenstein (2004), showed that the comparison of two partially understood situations or two unsolved problems could be beneficial, since it influences encoding and favors re-representation. Our claim is that, when not provoked by an experimental setting, the process of re-representation may be due to a spontaneous search and that this process is basic in situations where a source is not provided and is actively searched for. In our study of learning text editing (Sander & Richard, 1997), we have shown that, during a first phase, the typewriter is a source of analogy. The text editor is first categorized as a typewriter, as this is the known domain which shares the greatest number of salient features. The general goal is the same: to type a text, and objects are shared: a keyboard and a surface on which what is typed appears. This analogy is limiting and it is crucial to provide a mechanism explaining how this can be overcome, that is, how another source, other than the first one selected, can be used if the first analogy turns out to be limiting. We have shown that progress in learning was guided by analogies with sources at a higher level of abstraction, defined along the class inclusion relation. We considered two categories more abstract than typewriting, ordered by an abstraction relation, namely writing in general (typewriting is a specific way of writing in general, as handwriting is another specific way); and manipulating objects (we manipulate the components of a text when we write it, when we correct it, when we duplicate and move parts of the text from one place to another). We first identified the knowledge concerning each of these categories by placing the participants in the relevant context (for instance manipulating tokens for the context of manipulating objects) and asking them to solve tasks isomorphic to the ones that can be solved on a text editor (a task of moving a string of contiguous colored tokens was isomorphic to a task of moving a word with a text-editor). Doing this with all of the objects and all of the goals allowed us to identify the knowledge about the hypothesized sources and to compare the learning that was actually observed with the successive use of these sources. Once knowledge about typewriting turns out to be inadequate, tasks were solved by using knowledge about writing in general, or, if the writing level turns out to be inadequate, knowledge about manipulating objects. Participants progressively discovered the properties of the text editor in this order. Thus, the analogy with typewriting was only the first step of learning: the entire learning process turned out to be guided by analogy through recategorization at increasingly higher levels of abstraction.
Interpreting at the relevant abstract level

Although few studies looked for successive analogies, many other results of the literature may be reinterpreted as indicating that interpreting at the relevant level of abstraction is crucial in analogy making. Sander and Richard (1997, 1998) provide converging analyses for several analogical problem solving situations (the candle problem, Duncker, 1945; the tumor problem Gick & Holyoak, 1980, 1983; the genie problem, Holyoak, Junn & Billman, 1984; Archimedes problem, Goswami, 1992). For instance, for the well-known candle problem (Duncker, 1945), in which participants, provided with tacks, matches and boxes, had to fix a candle on a wall, the required solution was to tack the box onto the wall and to use it for supporting the candle; one major difficulty was to consider the box not only as a container but also as a platform and it has been shown that the task was easier if the box was empty than if it was provided full with tacks (Adamson, 1952). Finding the solution with tacks in the box requires two levels of abstraction, and is thus more difficult than finding it with empty boxes, which requires one level of abstraction and is still not as obvious as would be using the box as a container. In Archimedes' analogy between a crown and his own body, encoding both entities at the same abstract level implied neglecting specific properties such as, for the crown, symbol of kingship, made of precious metal, etc.; as well, the human body is a living body and has to be considered as a lifeless body.

In the situations just considered, the spontaneous retrieval of potential sources presupposes that these sources are very familiar and overlearned: such a degree of learning is not attainable in experimental tasks on transfer. Moreover, the success of this retrieval depends on the existence of a relevant source in existing knowledge. This is probably not the case in difficult transfer problems, such as in the TOH situation, where the transfer of learning is from the dimension of location to the dimension of size. These dimensions have very different properties in our familiar environment as Richard, Clement and Tijus (2002) have shown. For instance, several objects may have the same size, but not the same location: if they are on the same peg, one is necessarily above or below the other, so that additional cues are involved which do not exist for size. Given these differences, it cannot be expected that a spontaneous search for abstraction will succeed. Actually, the type of conception that is necessary to map a change of size onto a change in location problem is very abstract, since it is necessary to exchange the role of the dimensions. Clément and Richard (1997) have shown that the difference in difficulty between the TOH isomorphs with a move operator and those with a change of size operator is due to the point of view induced by the situation, which is adopted in encoding the action. A change (of location, size...) may be considered from two different points of view: it may be perceived as a mere change of state, without considering the process which has taken place between the initial and the final state or it may be seen as a process of transition, as a path leading from the initial to the final state, with all of its intermediate steps. According to the first interpretation, the change is an all or nothing process; according to the second one, it is a graded process. The second interpretation is more specific since it concerns the way the change is made, the first one is more abstract, since it posits only that a change has happened. In TOH isomorphs, the adequate interpretation happens to be the perception of a change as an all or nothing process: it allows interpreting the second rule of the instructions as a condition for leaving a state and the third rule as a condition for entering a new one. Adult participants faced with a standard TOH problem interpret a move as a pure change of location, because of the way it occurs on a computer screen following a click: the object appears in one place, then it appears in another, and no intermediate state appears. When faced with a change of size problem, they assume there are intermediate stages: they perceive the change of size as a graded process, as a kind of growth process and do not break down the action of changing the size into two sub-actions: removing a given size from an object and endowing it with a new one.

In order to attain the relevant degree of generalization, it is probably necessary to relearn, in the new situation, what has been learned in the previous one, possibly with fewer trials. This hypothesis is supported by the discovery of the importance of solution failures experienced in the training problem (Gick & McGarry, 1992) with the problem of the mutilated checkerboard as a transfer task (Wickelgren, 1974). In this task, two squares have been cut off at opposite diagonal corners of the board, so that they are of the same color. The question is thus: is it possible to cover the remaining 62 squares with 31 dominoes? The response is obviously “no” if one thinks of the color of the dominoes. The problem actually is difficult, because participants do try to cover the checkerboard. As a training problem participants are given an isomorphic version of the checkerboard problem which is either easier to solve or more difficult depending on the semantic cues which are provided. Only when the more difficult problem was given in the learning phase, was transfer significantly higher than for a control group. These results indicate that presenting a difficult training problem that tends to induce solution failures analogous to those produced in the target problem facilitates spontaneous transfer. The conclusion is that solution failures experienced in the source problem are instrumental in recognizing its similarity to the target problem, because similar solution failures are experienced in the target problem. Richard and Zamani’s (2003) study of the Passalong test supports this hypothesis about the role of impasses. The authors demonstrate that impasses provide information that is useful for discovering which of the properties of a situation are relevant to the problem and to encode the situation at the relevant level of abstraction. We would like to observe that the conditions necessary to produce deep semantic changes such as those just considered are not present in the analogy paradigm since solution to the source problem is usually provided. This is probably the reason why few transfer is observed when surface similarities and hints are absent. We suggest that benefits resulting from analogical encoding
(Gentner & al., 2004; Kurtz & Lowenstein, 2004) may be interpreted within that framework: comparing two partially understood situations or two unsolved problems favor encoding at the relevant abstraction level in the same way that impasses favor recoding the problems in the case of transfer studies.

Chen (2002) has also been concerned with the role of abstraction in transfer. It is informative to situate our view relatively to his. Chen (2002) argues that another kind of similarity, namely procedural similarity, may be manipulated independently of structural similarity and surface similarity and consequently is distinct. In a series of experiments where the problem is how to measure the weight of an elephant with only small objects available, he has shown that the amount of transfer depends on the degree of abstraction shared by the target and the source problems. The problems may have in common the principle (equivalence of weights), the method (compression device or balance device) as well as the specific procedure implementing the method (for instance, measuring the equivalence by the resistance of a spring or the resistance of water); they may share the principle and the method but not the procedure, or they may share the principle only. The comparison of problem-solving performance shows that spontaneous transfer is effective only when the procedure is common to both problems, and hints have either no effect or very limited ones. Chen’s claim is that the difference between conditions lies in the implementation of the principle: when the target and the source problems are similar at an abstract level only, the procedural details are missing, therefore the procedure is difficult to work out; Since the hints have little effect, the difficulty is not in accessing the source but in implementing the procedure.

We shall illustrate our position by showing how it makes possible to reinterpret the experimental situation and data reported in this study. The three levels of similarity compared by Chen (2002) correspond to different descriptions of the objects and the descriptions required to recognize the analogy are more and more abstract as we move from the higher to the lower level of similarity. When target and source problems share a similar procedure, the coding of events does not need to be changed to recognize the analogy: a boat floats like a vessel in water; it sinks when it is too heavy, like a vessel. The only difficulty consists in mapping the objects. When the method is similar but not the procedure, there are important differences in the way actions are conceived and named. Let us consider the compression method. The effect of putting a weight on the holder is perceived as compressing the spring; the effect of putting a weight in a boat is that the boat sinks. In both cases the result is a difference in height, but although the difference in height seems physically the same, it is semantically different: the spring changes in height, the boat changes in depth. The height is measured from the level of the ground to an upper level; the depth is measured from the level of water to a level under water. In order to consider the spring device and the boat device as similar, weight has to be thought as acting against a resistance: two weights are equal if they compensate the same resistance. The ground to cover is still bigger when the source and target problems differ in methods. In the balance device, the weights are compared directly, while in the case of the compression device they are compared indirectly by means of a resistance. In order to consider the balance device as similar to the spring or boat devices, the weights have to be viewed as forces. According to this point of view, it is not surprising that participants do not recognize analogy at an abstract level: the difficulty does not lie in the implementation of the procedure, it is due to the fact that the difficult concepts necessary to draw the analogy do not exist or, if they do, they have to be retrieved. In other words, the differences reported by Chen are not procedural but conceptual differences.

Conclusion

We have shown that:

(i) Neither physical nor surface similarities are directly responsible for analogy recognition and transfer of learning. In fact, the way the object properties are coded and the way the actions and events are interpreted is the major factor in the process.

(ii) Constraints of the task (or the structural relations) require a definite type of coding of the properties of the objects and of their relations for the analogy to be recognized: this is particularly manifest in TOH isomorphs and permutation problems.

(iii) Recognition of analogy is a matter of recoding at the appropriate level of abstraction, which depends on the relations between the source and target problems. Actually it may happen that situations are not perceived as similar because they are coded at a too specific level. If they were coded at a more general level, the same situations would be seen as similar.

Acknowledgments

Special thanks to the four anonymous reviewers for their insightful comments on a first draft of the paper.

References


