Downloaded from UvA-DARE, the institutional repository of the University of Amsterdam (UvA)
http://dare.uva.nl/document/68577

File ID	68577
Filename	CHAPTER 6. COMPONENTS OF INDUCTIVE LEARNING

SOURCE (OR PART OF THE FOLLOWING SOURCE):

Type	Dissertation
Title	Exploring the limited effect of inductive discovery learning : computational models and model-based analyses
Author	D.H. van Rijn
Faculty	Faculty of Social and Behavioural Sciences
Year	2003
Pages	170

FULL BIBLIOGRAPHIC DETAILS:
http://dare.uva.nl/record/220416

Copyright

It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other then for strictly personal, individual use.

UvA-DARE is a service provided by the library of the University of Amsterdam (http://dare.uva.nl)

CHAPTER 6.

COMPONENTS OF INDUCTIVE LEARNING

Abstract

In this chapter, four different dimensions are presented along which inductive learning tasks can be compared. Also, we will discuss to what extend the conclusions from previous chapters of this thesis are applicable to inductive learning in general. Based on this discussion, three factors are presented that explain both successful and less successful results on inductive learning tasks. This chapter concludes with suggestions for further research and with a discussion of the implications of this research for the application of inductive learning in the curriculum of modern education.

In the introduction of this thesis, inductive learning was compared to scientific discovery, as is common practice (e.g., Klahr & Dunbar, 1988; Klahr, 2000; Kuhn et al., 1995; Schunn & Anderson, 1999). The assumed correspondence between scientific discovery and inductive learning implies a central role for hypothesis testing in inductive learning (c.f., SDDS, Klahr & Dunbar, 1988). Given this centrality of hypothesis testing, conducting experiments is not so much a task in itself, but a necessity for constructing, testing and refining hypotheses. The studies presented in this thesis, however, shed a different light on inductive learning.

In contrast to the notion of a central hypothesis, inductive learning in the tasks presented in this thesis appears to be mainly focused on the construction of experiments

to "test the effect of that variable", without initially having a carefully constructed hypothesis. When learners construct hypotheses, these are often based on a simple evaluation of their prior knowledge. Instead of criticizing this behavior as imperfect or (too) shallow, this concluding chapter, in line with the reasoning in previous chapters, presents a rationale for this behavior and explains it in terms of bounded rationality (Simon, 1957).

To generalize the conclusions of this thesis to the domain of inductive learning, it is necessary that the conclusions are based on a set of inductive learning tasks instead of being specific to one particular task. The set should reflect a range of inductive learning tasks that is as broad as possible. In the next section of this chapter, we present four different dimensions along which inductive learning tasks can be compared and argue that the tasks discussed in this thesis cover a broad range of the spectrum of inductive learning tasks. Then, we will discuss the tasks presented in this thesis focusing on the conclusions from previous chapters that are suitable for generalization to inductive learning in general. Based on these conclusions, three factors are presented that explain both successful and less successful results on inductive learning tasks. Finally, we will conclude with suggestions for further research and with a discussion of the implications of this research for the application of inductive learning in the curriculum of modern education.

Four Dimensions of Learning Tasks

Chapters 2 to 5 discussed inductive learning in different domains using different tasks. Dimensions of inductive learning tasks that determine the difficulty of the task include the extent to which learners need to perform experimentation, variable identification and conceptualization, and complexity. The tasks are compared on these dimensions to ground the assertion that they form an appropriate subset of the complete spectrum of inductive learning tasks.

Experimentation The tasks differ in the availability of the data on which conclusions have to be based. In the balance scale task, the problems presented to the learners are selected by the experimenter, the learners cannot construct experiments by themselves. This constrains how easily a learner can test constructed hypotheses, as the presented experiments might not be the experiments the learner needs to test a current hypothesis. On the other hand, this could also be

an advantage, as a carefully constructed sequence of experiments might guide the learner in the right direction for deriving new knowledge (c.f., Chapter 2). In both the Peter-task and the Optics tasks, learners have to construct experiments. Although this enables learners to actively test their hypotheses, it also burdens them with the task to construct correct experiments and the decision about when to stop creating experiments.

Variable Identification The variables that underly the simulation behavior in the tasks can be either relatively easy identified from the task or need to be actively induced by the learner. With respect to the identification of the variables, it is not only important that the learner identifies a particular representation in a simulation as an important variable, but also identifies its structure, i.e., its quantity space (see Chapter 5). In the Peter-task, both the variables and their quantity spaces are easily identified. Both the five variables and the quantity spaces of these variables are given by the interface. In the balance scale task, only two aspects of the balance scale are modified over different representations, indicating that these are useful variables for the inductive learning process. However, these variables are not identified as significant variables in either the instruction or the task-setup. Therefore, part of the difficulty of this task is to identify that these variables play a role in determining the behavior of the system. If these two variables are selected as interesting, the interface prescribes their quantity-spaces (i.e., the maximum number of weights that can be placed on a single peg, and the number of pegs for the maximum distance). As discussed in Chapter 2, the identification of variables is an important aspect in the explanation of children's behavior on the balance scale task. However, for the third task, Optics, both aspects are not straightforwardly derived. Although all potential variables are visible in the interface, the learner has to select the correct variables for his or her experiments and derivations *and* has to induce the quantity space to derive correct conclusions.

Conceptualization Although the identification of the relevant variables is an important, independent part of the inductive learning process, it is also a prerequisite for a correct conceptualization of the task. Conceptualization refers to finding the relevant set of concepts to describe the behavior of the domain under study, for example, the type of relations between the identified variables, but

also more generally how the learners think about the task.

In the Peter-task (and especially in the Peter-goes-shopping variant), the task-format and the associated instruction direct the learner toward an adequate conceptualization. Learners have no problem identifying that they have to discover the effect of the different levels of the shown variables on a pre-specified outcome. Even if the existence of an interaction effect is not known to the learners, the task-setup or domain appears to guide the learner toward a correct conceptualization of interactions (i.e., a conditional-effect).

However, in both the balance scale task and in Optics, the learner has not only to identify which variables are important, but also has to identify the type of relation that links different values of an independent variable to an outcome on a dependent variable. In the balance scale task, the learner has to come up with the idea to multiply the values for weights and distances (even if the often earlier thought of addition of weights and distances also proves relatively successful). In Optics, the learner has to induce that most of the effects are, qualitatively speaking, discontinuous. (That is, both the virtual focal point and the heart of the lens are points in quantity spaces in which the behavior in the domain shows a discontinuity.)

Complexity As in all domains, different inductive learning tasks have different levels of complexity. However, it is difficult to judge what defines complexity of a task as there are multiple, not necessary related, aspects that define it. Here we will not focus on task aspects in itself, but on how these task aspects are perceived by the learners.

In all three tasks presented here, a learner might come up with a set of relations that appears to give a correct or at least an adequate description of the domain, whereas in fact this description is only valid for a particular subset of the domain. Although a description that covers the complete domain is obviously the best description, the partial description might be correct for the subset of data the learner has seen. Take, for example, the balance scale task. If a learner uses the Addition Rule and does not encounter any items for which the Addition Rule renders the incorrect answer (e.g., the item with two weights on peg two on the one side, and four weights on peg one on the other side), the used Rule fits the data perfectly. Similar examples can be found in the two other

tasks, for example, the main effects without interactions in the Peter-task and experimenting without moving the lamp over the virtual focal point in Optics.

Obviously, one can find such a subset in almost all relatively complex tasks. However, complex tasks can differ in the ease with which learners are able to infer whether their current description fits the complete domain. An important concept in this respect is whether or not a task is ill-defined All three tasks presented in this thesis are ill-defined in that it is not clear what the outcomes of the discovery process should be, nor which variables should be incorporated in the discovery process. Therefore, learners need to use heuristics. For example, because of the ill-definedness of the variables, learners do not know what the complete experiment-space is. During experimenting, they need to assess whether they have covered a suffient area of this unknown experiment-space to warrant generalizing their conclusions. But not only does this influence the experiment-space, the ill-definedness also influences the hypothesis-space as learners do not know what type of relations need to be searched for.

Because of the differences in ill-definedness per task, the tasks also differs with respect to how salient the incompleteness (compared to a complete description) of a certain set of observed regularities is. For example, although the Peter-task has a more complex set of underlying regularities, these regularities have a simpler structure than the multiplication rule in the balance scale task. Related to this is the relative gain of constructing experiments; in some tasks it is easier to construct experiments that refute the current set of regularities than in others. For example, it is relatively easy to reject hypotheses in BigTrak (Klahr & Dunbar, 1988), as almost all incorrect hypotheses have a limited scope of accurate predictions. On the other hand, if a learner does not know about the virtual focus point, a lot of experiments can be constructed that support an incorrect hypothesis.

As the balance scale task is relatively well defined; it is known to the learners what to predict and the variables of interest are easily identified (by adult learners), the main difficulty is finding the right method for combining the weights and distances. The Peter-tasks are one step more complex, in that in these tasks the learners have to detect which variables are actually influencing the outcome. Moreover, the right description can only be found when a learner constructs

the right set of experiments; there is nothing in the task that alerts the learner to this set. The most complex of the three tasks discussed is the the Optics task. In this task, the variables are identified, but not the type of knowledge that learners are expected to gain nor the quantity spaces (see Chapter 5) of the variables. And, similar to the Peter-task, learners can come up with a partial description that accurately fits a large set of observed experiments while not being hinted by the conducted experiments that the current description is inaccurate.

Table 7.1: Categorization of the inductive learning tasks presented in this thesis.

	Balance scale task	Peter task	Optics
Experiments			
Self-directed	.	✓	✓
Externally generated	✓	.	.
Variables (QS)			
Self-induced	✓	.	✓
Visible in interface	✓	✓	.
Conceptualization			
Self-induced	✓	.	✓
Provided by task setup	.	✓	.
Complexity	low	intermediate	high

Table 7.1 gives an overview of the position of the tasks on the four dimensions.

Three Tasks

Even though the three tasks discussed in this thesis cover the complete spectrum of the above presented dimensions, the inductive learning behavior in these tasks shows a relatively large overlap.

Balance Scale Task

In the chapter on the balance scale task, a computational model is presented that explains how children's behavior on this task develops from using a simple guessing rule to using the correct rule which involves calculating the torque. A central feature of the explanation of development is prior knowledge about solving forced choice tasks. If the model is presented with a forced choice task, it has access to knowledge

that states to search for a difference between the alternatives of the forced choice. In the context of this task: Search for differences between the right and left arm of the balance scale. If the model finds a difference, this difference is used in solving the problems and new knowledge is constructed that relates those problems to that difference. Initially the model uses only one type of difference (i.e., the difference in the number weights between the left and right side of the balance scale). This causes a relatively large number of erroneous predictions about the movement of the balance scale. However, as searching for differences is resource intensive, the model does not immediately search for a new difference if an erroneous prediction is made. Only if the number of erroneous predictions becomes too large compared to the effort that it takes to find a new type of difference, the model will invest in a search for these new differences. This architecturally implemented mechanism leads to a model that performs according to satisficing principles: Do not search for new explanations as long as the current knowledge is not too often falsified.

PETER-TASK

In Chapter 3, the first chapter on the Peter-task, computational models are presented that focus on conducting experiments instead of focusing on constructing complex hypotheses as is put forward by theories like SDDS (e.g., Klahr & Dunbar, 1988). Nevertheless, without the emphasis on hypothesis construction or testing, these models do capture the main behavioral patterns. This illustrates that instead of a complex search for the correct type of hypothesis, inductive learning performance in a simpler task like the Peter-task is constrained by the learner's ability of constructing correct experiments and deriving knowledge from these experiments. But what guides the construction of the experiments? In the Peter-task, the main determinant is the computer-interface. Most learners proceed by starting with the top-most variable working their way downward. However, another determinant is prior knowledge. Learners also construct experiments based on an evaluation of their prior knowledge. If they assume a particular effect to be associated with a level of a variable, they will explicitly test this level. Moreover, if an effect is discovered that is not in line with their prior knowledge, this is often reason for a more thorough investigation of the effects underlying that particular level or variable. In the strongest form, this leads to behavior that limits inductive learning to what is covered by the experiments that are "dictated" by the interface, supplemented with experi-

ments constructed on the basis of prior knowledge. If the behavior satisfied these two constraints, learners often see no reason to continue learning although aware of the not conducted experiments. This can be seen as an illustration of the satisficing principle: Although there is maybe more to be known about the task, if there are no more loose ends, the inductive learning process is considered finished.

Chapter 4 takes this one step further and proposes a measure of consistency as an alternative measure for the quality of the inductive learning process. Based on a comparison of two Peter-experiments, it was shown that the score on the new consistency measure is more stable over domains than the previously used measures. These more stable scores are a consequence of the focusing on the *process of knowledge derivation* instead of (implicitly) on the completeness of the experiment-space coverage. This way, the varying amounts of prior knowledge for different domains does not influence the consistency score as readily as it does the completeness and comprehension scores.

Optics

Apart from the methodological issues discussed in Chapter 5 on the Optics task, the conclusions from Chapter 5 resemble those from Chapter 3 and 4. Learners, even in the more complex Optics environment, are not focused on creating and testing hypotheses, but are mainly focused on constructing experiments. An analysis of learners' think aloud protocols showed that they do learn, however, that the learned knowledge does not necessarily overlap with knowledge tested in the post-test.

This is mainly caused by learners not knowing what they are looking for, both in terms of the final theory they are supposed to find (e.g., how the dependent variable is structured), nor in terms of which (independent) variables to use in explaining the behavior of the system. Again, learner's performance depend on their background knowledge (e.g., levels and variables to test) and a satisficing principle as testing all the combinations of levels and variables is pragmatically impossible.

Task General Observations

Regardless the differences in the tasks described above, a number of observations can be made in each of these tasks, indicating that these observations are relatively task-independent:

1. Discovery skills used by learners in inductive learning tasks are often less complex than envisioned. That is, instead of extensive hypothesis-driven behavior, the behavior shown by learners in the studies in this thesis can to a large extent be described by a more simple algorithm: (1) think about something that can have an effect, (2) construct an experiment for the different situations in which that effect occurs, and (3) induce the existence of that effect based on the outcomes of these experiments. (See Chapter 3, and the rationale of the model presented in Chapter 2.)

2. Which variables are tested, and what levels are chosen by the learner for that variables depends on the task properties and prior knowledge, often taking the form of assumptions about the effect of variables. Moreover, the outcomes of experiments are interpreted by the learners in terms of this prior knowledge. If the newly discovered effects are unexpected, the learner is likely to engage in further examination of that variable. (See Chapter 2, 3 & 4.)

3. If no distinct stop criterion is given and the learner has no way to know whether the discovered effects are the ones searched for, a satisfying principle determines when to stop. That is, when new experiments are unlikely to uncover new results, the costs of conducting these experiments becomes higher than the profit associated with knowing these new results. In those situations, it becomes a rational decision to stop experimenting. (See Chapter 4 & 5.)

Based on the above, inductive learning in the tasks described in this thesis can be described as a process utilizing a relative straightforward experiment-construction strategy, which shows a pronounced influence of prior knowledge, and constrained by a stop-criterion based on bounded rationality.

Three Factored Explanation for Inductive Learning Behavior

The above discussion focused on the overlap between the three tasks discussed in this thesis along four dimensions. These dimension can also be generalized into three more general factors that are related to the learners' behavior. These factors

explain learners' inductive learning behavior at a global level, and also explain why learners' performance is often below the expected levels.

Simplicity Inductive learners appear to strive for an as simple as possible explanation for the observed behavior. Their behavior is probably best described by the principles of bounded rationality (Simon, 1957). Instead of performing an exhaustive search of the task's experiment-space, learners weigh the possible gain of conducting more experiments against the associated costs. Even in relative simple and well-structured tasks, the amount of effort associated with conducting all experiments prevents learners from conducting all experiments, making the learners resort to satisficing. Presumably, the perceived possible increase in knowledge does not countervail the costs of conducting more experiments. Another explanation for this effect is related to the law of diminishing returns. Initially, each correctly constructed experiment enables a learner in the Peter-task to derive a new main effect simply by comparing the previous experiment with the just constructed experiment. However, after conducting all the experiments associated with the main effects, a learner has to compare three already conducted experiments with a newly created experiment to derive a first order interaction. Moreover, if a learner assumes that simpler effects occur more often than more complex effects, searching for complex effects is more expensive because on average more experiments have to be constructed and compared per discovered effect. Therefore, the more a learner discovered, the smaller the chances are of discovering new information given the same investment.

Another aspect related to simplicity is that learners have a preference for simple hypotheses. If a simple hypothesis appears to predict the behavior of the task relatively well, learners are likely to stick to that hypothesis. Although this can be seen as an example of confirmation bias (Klayman & Ha, 1987) or as a preference for parsimony, this can also be accounted for on the basis of bounded rationality. As complex hypotheses tend to be both more complex to construct *and* to test, keeping an hypothesis as simple as possible decreases the chance to get entangled in complex hypothesis testing.

Guiding knowledge Apart from the aim for simplicity, learners' behavior is also influenced by the application of prior knowledge. The most straightforward example of this notion is the importance of the availability of the weight and

distance concepts in the balance scale model presented in Chapter 2. A similar effect is present in models of Peter-task behavior, presented in Chapter 3. Besides testing the five variables presented in the interface, the learners only test for effects that seem probable or plausible on the basis of their prior knowledge about the domain. Given the overall complexity of the Optics task it is harder to pinpoint the exact contribution of prior knowledge. However, even in the relative short protocols presented in Chapter 5, both positive and negative effects of prior knowledge are clearly visible.

Thus, in all of these tasks, learners' behavior is actively guided by their knowledge about a domain. Not only in terms of their procedural knowledge that describes how to conduct correct experiments, but also in terms of declarative knowledge associated with the domain of the task.

An example of this is the difference in performance in the two tasks presented in Chapter 3. In the Peter-bikes-to-school task, learners are *not* told about the possibility of interactions between variables. Only two of the fifteen learners correctly discover and report the interaction during the post-test. In the Peter-goes-shopping task, the instruction contains information about the existence of an interaction, without making explicit how this interaction should be tested. Five of the fifteen learners in this task correctly report the interaction. Moreover, when these learners are given the Peter-bikes-to-school task directly after being tested in the Peter-goes-shopping task, thus after their concept of interaction has been "primed", 9 out of 15 discover the interaction in this task whereas without being primed only 2 learners discover the interaction (Schoutsen, 1999).

More formally, given the aim for simplicity, the knowledge a learner has about the task is used in a process that adheres to the satisficing principle and limits the effective hypothesis and experiment search space to the regions of the total space in which the learner *expects* an effect (cf. the learner search space, Van Joolingen & De Jong, 1997).

Salient Discrepancies The "guiding knowledge" not only has a direct effect on the inductive learning task in that it determines which effects are researched, but also plays a role in the evaluation of discovered effects. In both the Peter-tasks and in Optics, two types of reactions have been observed after encountering a discrepancy between prior knowledge or hypotheses and data. In most of

the situations, learners just acknowledged the newfound effect, and revised their earlier beliefs. However, in some situations, learners were struck by the discrepancy and engaged in further research, trying to figure out what caused this discrepancy. This elaboration caused by the salient discrepancy often led to improved knowledge about the domain (e.g., in the Peter-task, discovering the interaction or in the Optics-task discovering the effect of the virtual focus point). This way, prior knowledge about possible effects does not only guide inductive learning by guiding the focus on what is tested, but the saliency of the discrepancy also determines whether the learner discovers less shallow effects in the domain under study.

Note that the notion of saliency also plays an important role in the simulations in Chapter 2. By means of increasing the saliency of a yet unused feature, the learner becomes aware of this feature and incorporates it into the discovery process. Although saliency plays a somewhat different role in this task compared to the Peter and Optics tasks, in all three tasks the saliency of features or concepts is an important predictor of whether or not that feature or concept is included in the final knowledge of the task.

CONCLUSIONS

The analyses presented in this thesis show that the claimed generality of the SDDS theory (Klahr & Dunbar, 1988) is probably overstated (see for a similar argument Johnson & Krems, 2001). In simple inductive learning tasks, as often encountered in normal life, hypothesis construction does not play the all-important role as sketched in the SDDS theory. Instead, learners seem to "simply" test for the effect of variables. This tendency toward *simplicity* also plays a more general role in inductive learning. Learners tend to keep their representations of the task as simple as possible. The level of this simplicity is determined by what is deemed necessary to test by the interface combined with *guiding knowledge* that states which variables and what type of relations are tested. If during the initiated tests *salient discrepancies* are found between the experimental outcomes and the assumed effect, elaborate on the found results. Because of this, these discrepancies can guide the learner toward a more complex set of hypothesis than initially constructed on the basis of the simplicity and guiding knowledge factors.

IMPLICATIONS

As was shown in all three tasks, learners have a difficult time discovering complex relations if they do not have guiding knowledge. That is, without explicit help, learners do not discover the multiplication rule in the balance scale task, seldom discover the interactions in the Peter-task, and do not easily discover the important landmarks in the Optics task. This clearly illustrates that the emphasis on "self-discovery" in modern curricula of secondary schools can only lead to satisfying results if the students are actively guided in their discovery process and if the tasks and domains to which discovery learning is applied are carefully chosen. First, students need to have relevant prior knowledge, so that they know what to look for. Second, they have to be aware that it is not necessary to cover the complete experiment-space. However, they also have be aware that this only holds if they select their experiments from all parts of the experiment space (e.g., using heuristics like "test the extreme values" or "when you think that you're ready, do some more random experiments to see if your predictions hold"). And third, given the important role of discrepancies in discovering the more complex relations in all three tasks, students need to be aware of the importance of discrepancies between their own hypotheses and assumptions and the discovered effects. Only if these three conditions are met, discovery learning might be a useful *addition* to more traditional educational methods.