Cognitive Task Analysis in Service of Intelligent Tutoring System Design: A Case Study in Statistics

Marsha C. Lovett

Center for Innovation in Learning, Carnegie Mellon University
Pittsburgh, PA 15213, USA
lovett@cmu.edu

Abstract. Cognitive task analysis involves identifying the components of a task that are required for adequate performance. It is thus an important step in ITS design because it circumscribes the curriculum to be taught and provides a decomposition of that curriculum into the knowledge and subskills students must learn. This paper describes several different kinds of cognitive task analysis and organizes them according to a taxonomy of theoretical/empirical ∞ prescriptive/descriptive approaches. Examples are drawn from the analysis of a particular statistical reasoning task. The discussion centers on how different approaches to task analysis provide different perspectives on the decomposition of a complex skill and compares these approaches to more traditional methods.

Introduction

One of the first steps of intelligent tutoring systems (ITS) design involves performing a cognitive task analysis of the curriculum to be taught. There are a variety of approaches to cognitive task analysis. This paper argues for a new taxonomy of those approaches. Theoretical approaches tend to emphasize the informational requirements of the task: what information is included in a problem description? how is this information represented? how must it be transformed to produce a solution? Empirical approaches, on the other hand, emphasize the processes people are engaged in as they perform the task. Another distinction is between prescriptive approaches that analyze a task in terms of how it *should* be performed and descriptive approaches that analyze a task in terms of how it *is* performed. One can combine these various features to get a variety of cognitive task analysis styles. As a structure for this paper, take the two distinctions (theoretical vs. empirical and prescriptive vs. descriptive) as orthogonal dimensions and cross them to create four distinct combinations. This paper will demonstrate how all four cognitive task analysis styles can be useful in service of ITS design.

The task domain to be analyzed is a topic in statistics called exploratory data analysis (EDA). EDA involves using statistical tools to describe, summarize, and draw conclusions about data. It is a good example for demonstrating cognitive task analysis for several reasons. First, EDA is a complex problem-solving skill that requires the solver to bring several different types of knowledge to bear on real-world problems (e.g., domain-specific and domain-general knowledge, higher-level reasoning and low-level perceptual processing, and both declarative and procedural knowledge). Thus, providing a cognitive analysis of this task will shed some light on the structure of what is sometimes considered to be more of an art than a science. Second, the need for EDA arises frequently across a wide variety of disciplines and in everyday life [2]. As such, it is commonly viewed as an important target of instruction [11]. Third, teaching EDA is a challenge [4, 17]. Many students at the undergraduate level, even after a statistics course, lack the necessary skills for selecting appropriate data-analytic tools [17], interpreting statistical graphs [13], and drawing conclusions from data [7]. Thus, identifying the key components of this complex skill could help improve instructional design and hence student learning.

The outline of this paper is as follows. The next section describes related work to set this research in context. The subsequent section presents a taxonomy of four types of cognitive task analysis and exemplifies how each applies to the domain of EDA. The final section discusses how these different task analyses contribute different perspectives to ITS design and relates them to more traditional methods.

2 Related Work

Although there has been a good deal of computer-assisted instruction in the domain of statistics, very little of this research specifically addresses the domain of EDA. There are two standard ways that computers are used in statistics instruction. The first is as a computational tool. Many introductory statistics classes include off-the-shelf statistical software packages to help students in analyzing real data sets, but these packages do not have an instructional component. The second standard way that computers are used in statistics instruction is as a simulation tool. These simulation programs are used as vehicles for students to view certain statistical phenomena in a discovery-world context. Examples include ExplorStat [9] and StatVisuals [10]. Although these simulations have an instructional component, they generally focus on isolated statistical concepts (e.g., standard deviation, confidence interval) and so do not deal with the larger skill of EDA.

The most notable exception to this rule is StatLady, an ITS for teaching probability and statistics [14]. This system has been shown to lead to substantial learning gains both in terms of pre- to post-test measures and when compared with other (non-StatLady) instructional conditions [15, 16]. The descriptive statistics module of StatLady is most closely related to EDA, so it will be discussed here. The StatLady Descriptive Statistics Modules covers a range of skills from sorting data to generating informative data displays. The Descriptive Statistics-1 component of StatLady was based on a task analysis that identified approximately 80 different curriculum elements. These elements are arranged into groups of elements that are closely related to each other and that tend to be taught in a unit. These curriculum elements also are categorized as representing three types of knowledge: procedural, symbolic, and conceptual. This short description shows that analyzing a task in the domain of statistics can lead to a highly complex, but also highly structured, set of task components.

2.1 The Context for the Examples

Because the different kinds of cognitive task analysis will be exemplified in the context of EDA, it is important to have a more specific idea of what is involved in this complex task. We take an EDA problem to include (1) a set of data, (2) some description of those data (e.g., what the different variables contain, how the data are represented) and (3) a question that requires one to draw conclusions from the data. For example, take the following problem statement:

> A weather modification experiment was conducted in south Florida to investigate whether "seeding" clouds with silver nitrate would increase the amount of rainfall. Clouds were randomly assigned to the treatment group (to be seeded) or to the control group (not to be seeded), and data were collected on the total rain volume falling from each cloud. A variable named *group* contains data on whether each cloud was seeded or not (1 = seeded, 2 = not seeded), and a variable named *rain* contains data on each cloud's rain volume. Does cloud seeding increase rainfall?

In this problem, (1) there are data available, (2) the data are described as coming in the form of two variables, *group* and *rain*, and (3) the question being posed is "Given these data, does cloud seeding appear to increase rainfall?" These three features compose the initial state of this problem. The goal state or "solution" to such a problem includes the following: (1) interpretation of appropriate statistical analyses performed on the data (e.g., interpretation of a graphical display of the distribution of rain volume for each of the two groups) and (2) discussion of how those results address the question that was posed (e.g., was the seeded group's rain volume larger, on average, than the not-seeded group's rain volume). Getting from the initial state to the goal state requires many steps—some mental steps (e.g., transforming the problem statement into a plan for action) and some physical steps (e.g., working with a statistical software package to produce the displays). In general, a cognitive task analysis is aimed at identifying the separate task components necessary to progress through all the necessary steps and reach a solution.

3 Different Types of CTA

Figure 1 presents the 2∞2 table of analysis styles to be explored in the following sections.

	Theoretical	Empirical
Prescriptive	Information-theoretic analysis of how the task should be performed	Analysis of how experts produce high-quality solutions
Descriptive	Complexity analysis to predict how people actually do the task (e.g., typical errors)	Analysis of how novices actually do the task (e.g., their misconceptions)

Fig. 1. Two-by-two table of dimensions along which cognitive task analyses can vary.

3.1 Theoretical/Prescriptive

One approach to deriving a theoretical/prescriptive cognitive task analysis is to consider the information present in the initial state and in the goal state and then to specify how this information must be transformed to progress from one state to the other. In the case of EDA, the initial state includes a data set, a description of the data, and a question. The problem solver must somehow take this information and produce interpretations of analyses performed on the data and conclusions regarding the original question (see Figure 2, top row). Given this, it is logically implied that an intermediate step in the task must involve producing an analysis (i.e., the analysis to be interpreted, see Figure 2 middle row). To progress with a theoretical analysis then, one continues to ask, for each intermediate step generated, what is required to go from the preceding state to that intermediate state or from that intermediate state to the next state.

In the EDA example of Figure 2, we can apply this technique to the middle row and ask how a solver would get from the initial state to a state that has an appropriate analysis of the data. In EDA there are many different kinds of analyses that are appropriate in different situations. One must choose an analysis based on the type of variables involved in the problem (e.g., are they categorical or quantitative). Therefore, we must add a preceding intermediate state specifying the proper identification of variables (Figure 3, bottom row). Repeating this process to identify intermediate states between each pair of adjacent states leads to the cognitive task analysis in Figure 3. This task analysis could be refined further still, but we stop at this level for current presentation purposes.

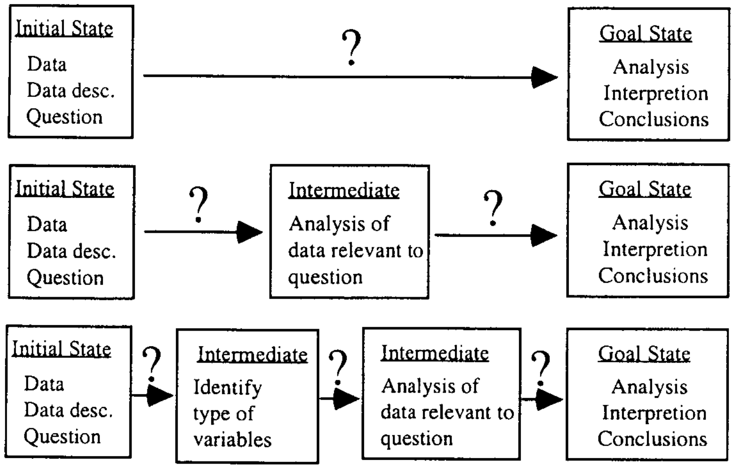

Fig. 2. Progressive revinement of a theoretical/prescriptive cognitive task analysis.

Note that this style of cognitive task analysis is similar to means-ends analysis [12] in that one repeatedly evaluates the difference between two problem states and establishes an intermediary to link them, eventually creating a complete solution path from initial state to goal state. Thus, the resultant analysis emphasizes the goal structure of the task more than the methods for achieving those goals. Thus, this technique is most appropriate for domains where planning is very important to good problem solving. This is true of EDA, especially in complex problems, where several planning steps must be taken before any analysis is initiated. Also, any situation where the ITS design involved an emphasis on goals or planning would be well informed by this style of task analysis. For a detailed enumeration of the procedures involved in performing a task, however, this approach should probably be combined with one or more of the other approaches described below.

3.2 Theoretical/Descriptive

The difference between a theoretical/*prescriptive* and a theoretical/*descriptive* task analysis involves a consideration of how people will actually perform the task, not just how they should perform the task correctly. This can involve an analysis of where errors are likely to occur or what parts of the task will be difficult to learn. One can apply cognitive psychological theories to make these predictions. For example, tasks that involve high working-memory demands are predicted to produce more errors [3].

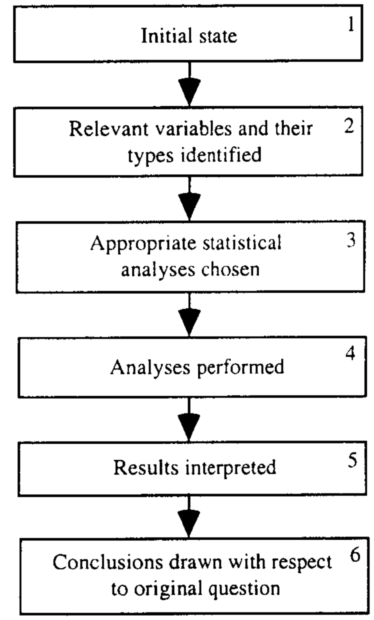

Fig. 3. Resulting cognitive task analysis.

One can apply this type of task analysis at more detailed level as well. Previous research suggests that steps that are not overt (e.g., "mental" planning steps) will be more difficult for students to learn than steps that have coresponding observable actions [8]. This suggests that the first three steps depicted in Figure 3 will be difficult to learn because students will tend not to see these steps modeled often and will tend not to be corrected when they make mistakes on these steps. It may spotlight particular areas where special attention should be directed during ITS design. This approach to task analysis is especially useful when data on student performance and errors are not already available.

3.3 Empirical/Prescriptive

The next type of cognitive task analysis involves an empirical/prescriptive approach. Here, experts can be asked to solve a series of problems *while* they are providing concurrent talk-aloud protocols [6]. Such protocols provide step-by-step information about what features of the problem the experts are attending to, what intermediate results they are producing, and what goals they are posting during the solution of a problem. The experts are ideally those who will produce high-quality (i.e., accurate and complete) solutions. This could range from a professional in the target domain to a "good" student, where "good" is defined in terms of some performance criterion.

1 So, it looks like what the problem is asking for is a comparison between breeder A and Breeder B puppies and specifically it says that Breeder A wants to compare the heights of the Great Dane puppies.
2 So actually the information on weight and litter is not necessarily going to be useful. ...
2 So, we've got 2 variables heights , well actually 1 var which is height.
2 And that would be the response variable and the explanatory variable is breeder.
3 So, one of the first things I'd like to do is do a boxplot to compare the heights for breeder A vs Breeder B. ...
4 <S selects boxplot in statistics package; display appears>
5 Ok there we go. Ok so now what do I see? Well, I'm looking at boxplots of heights of these Great Danes. Um. for the two breeders.
5 Immediately what I see is that the central part of the distribution for breeder A's puppies seems to be shifted lower than, relative to the central part of the distribution for breeder B's puppies.
5 And being specific, the third quartile for the distribution of heights of breeder A is actually even lower than the median for breeder B, the distribution of puppies from breeder B.
5**!** So that means that 75% of the pups were shorter than the median height pup for breeder B, that's for group A.
5 There's, I would say the IQR in the two groups is about the same.
5 The overall range is bigger in breeder A. Breeder A also has a low end outlier,
5**!** a little dwarf Great Dane.

Fig. 4. Expert protocol excerpt

Figure 4 presents some excerpts from a statistics instructor solving an EDA problem that asked whether puppies bred by two different breeders tend to differ in height. To the left of each line in the protocol is a number indicating the step number (from the task analysis in Figure 3) that the expert's verbalization appears to represent. One can see by a quick scan of the column of numbers in Figure 4 that this problem solver is following the theoretical/prescriptive task analysis quite closely. He makes one statement corresponding to step 1 from Figure 3, then three statements corresponding to step 2, etc. This analysis can be useful to confirm a previously generated theoretical/prescriptive task analysis.

Also noted in Figure 4 are two cases where the expert makes some particularly insightful comments during interpretation of his analyses. These lines are noted with a bold exclamation mark (**!**). In the first, the expert makes a specific inference about the proportion of data in one group that would be lower than the median for the other group. This step demonstrates not only that he can "read off" the values on a graphical display, but that he can use these values flexibly and appropriately. The final step shown in Figure 4 shows not only that this expert is willing to make a pun (not yet part of the EDA curriculum!) but that he takes into account the context of the problem in making his interpretations. This is important for the final step of drawing conclusions even though it was not present in the earlier theoretical/prescriptive analysis. Thus, seeing such features of expert performance can suggest revisions and additions to a theoretical/prescriptive analysis that might not have included them. In particular, it may revise the ITS design to emphasize the use of contextual information in interpreting statistical displays.

This theoretical/descriptive analysis style is also useful on its own. For example, it may be the preferred approach when the ITS designer does not have a lot of domain knowledge and hence would have difficulty in applying a theoretical/prescriptive analysis based on first principles. In addition, the theoretical/descriptive analysis focuses more on procedures required for good performance because it focuses on the expert's problem-solving processes. Therefore a domain that involves learning a lot of procedural knowledge would be well analyzed under this approach.

3.4 Empirical/Descriptive

Finally, empirical/descriptive cognitive task analysis can be performed by collecting talk-aloud protocols from students who are currently learning the target domain or from people who may perform the target task but have not received formal training. Since EDA is covered in many introductory statistics classes and yet students often exit these classes with misconceptions and incorrect strategies, we collected data from students whose problem-solving protocols might reveal such areas of difficulty.

Because the technique for collecting the data was similar to that in the empirical/prescriptive analysis above, we will just summarize the results of this analysis. Several areas of weakness were revealed as students often did incomplete or incorrect exploratory data analyses. By studying their talk-aloud protocols, the following difficulties were identified:

- Skipping of steps 2 and 3 (from Figure 3) such that students began carrying out statistical analyses without having considered the variable and analysis types
- Selecting incorrect statistical analyses (step 3) or just guessing at analyses
- Incompletely interpreting the statistical analyses generated (e.g., forgetting to describe the spread of a distribution and only describing its mean value)
- Failing to address the results of the statistical analyses to the question (step 7)

The first two of these areas of difficulty are consistent with the predictions of the theoretical/descriptive analysis, i.e., that planning steps would be difficult for students to learn because they are not overtly taken. The other areas of weakness help to highlight problem areas that may be representative of student misconceptions. For example, previous research has shown that students learning statistics often fail to understand the importance of variability. This could be related to some students failing to mention the spread of distributions in their interpretations of graphical displays.

4 Discussion

Each of the cognitive task analysis styles described above provided important results about EDA to inform the design of ITS for this domain. Here we summarize some of the strengths and weaknesses of these approaches, highlighting that different styles may be more effective for different domains or at different stages of ITS design. Theoretical/ prescriptive analysis helps identify the goals that must be achieved for high-quality performance of a task. This can lead to the specification of the knowledge components an ITS needs to cover. It is especially useful in providing an abstract picture of the goal structure of a task; this may be helpful in the early phases of ITS design. However, it does not emphasize the methods for achieving task goals, so heavily procedural domains may require further task analysis of another style. The theoretical/descriptive approach similarly derives its results from theory-based predictions but deals more with how actual performance may depart from the prescriptive model. This is useful in highlighting areas where special attention to student errors and difficulties in learning is warranted. When data on typical student difficulties and misconceptions are not available, this approach can be very useful. The empirical approaches differ from the theoretical in that they are based on studies of people performing the target task. This provides a "reality check" on the theoretical analysis at the cost of additional data collection and coding time. Empirical/prescriptive cognitive task anlaysis focuses on studying people who will produce high-quality solutions, whereas empirical/descriptive cognitive task analysis focuses on "typical" performance. These analysis styles are also useful on their own (not just as validation methods) because they (a) emphasize the processes involved in problems solving and (b) offer the potential for very fine-grained analyses. Thus, empirical analyses are appropriate for domains requiring a good deal of procedural knowledge of for later stages of ITS design where detailed refinements of the task analysis are helpful.

It is also important to note the relationship between these analysis methods and those described in other work. Durkin, for example, contrasts interview and case-study methods of task analysis [5]. Interview methods are quite different from those covered in this paper; they primarily rely on people's introspections about the knowledge required for good performance in a domain. Case study methods, on the other hand, relate to the empirical approaches described here. In particular, observational case studies performed with experts correspond to the empirical/prescriptive approach, and observational case studies performed with end users or novices correspond to the empirical/descriptive approach. Retrospective case studies, which involve someone's retrospective report on a previously solved problem, are in some sense intermediate between interview and observational case study in that they require some amount of actual problem solving (the empirical side) and some introspection (a recollection of why/how particular solution steps were taken). Like those described in this paper, these traditional methods are best suited to different situations. The general rule, however, is that *some* type of task analysis (or possibly several types) will aid ITS design because the better we understand the problem, the better we can solve it.

References

1. Anderson, J. R., Corbett, A. T., Koedinger, K. R., & Pelletier, R. (1995). Cognitive tutors: Lessons learned. *The Journal of the Learning Sciences, 4*(2), 167-207.
2. Becker, B. J. (1996). A look at the literature (and other resources) on teaching statistics. *Journal of Educational and Behavioral Statistics, 21*(1), 71-90.
3. Byrne, M., & Bovair, S. (1997). A working memory model of a common procedural error. *Cognitive Science, 21*, 31-61.
4. Cohen, S., Smith, G., Chechile, R. A., Burns, G. & Tsai, F. (1996). Identifying impediments to learning probability and statistics from an assessment of instructional software. *Journal of Educational and Behavioral Statistics, 21*, 35-54.
5. Durkin, J. (1994). Knowledge acquisition. In J. Durkin (Ed.) Expert systems: Design and development. (pp. 518-599). New York: Macmillan.
6. Ericcson, K. A., & Simon, H. A. (1993). *Protocol Analysis* (Rev.). Cambridge, MA: MIT.
7. Garfield, J. & delMas, R. (1991). Students' conceptions of probability. In *Proceedings of the Third International Conference on Teaching Statistics*, D. Vere-Jones (Ed.), Volume 1. pp. 340-349. Voorburg, The Netherlands: International Statistical Institute.
8. Koedinger, K. R., & Anderson, J. R. (1993). Reifying implicit planning in geometry: Guidelines for model-based intelligent tutoring system design. In Lajoie, S., & Derry, S. (Eds.) *Computers as Cognitive Tools*. Hillsdale, NJ: Erlbaum.
9. Lang, J., Coyne, G., & Wackerly, D. (1993). ExplorStat. Computer Program.
10. Loomis, J., & Boynton, G. (1992). Stat Visuals. Univ. of Calif. Computer program.
11. National Council of Teachers of Mathematics (1989). *Curriculum and Evaluation Standards for School Mathematics*. Reston, VA.
12. Newell, A., & Simon, H. A. (1972). *Human Problem Solving*. NJ: Prentice Hall.
13. Pittenger, D. J. (1995). Teaching students about graphs.*Teach. of Psych. 22*(2), 125-128.
14. Shute, V. J., & Gluck, K. A. (1994). Stat Lady Descriptive Statistics Tutor: Data Organization and Plotting Module. Brooks AFB, TX: Armstrong Laboratory.
15. Shute, V. J. (1995). SMART: Student Modeling Approach for Responsive Tutoring. *User Modeling and User-Adapted Interaction, 5*, 1-44.
16. Shute, V.J., Gawlick-Grendell, L. A., Young, R.K., and Burnham, C.A. (1995). An Experiential System for Learning Probability: Stat Lady Description and Evaluation. *Instructional Science (4)*, 1-22.
17. Ware, M. E., & Chastain, J. D. (1991). Developing selection skills in introductory statistics. *Teaching of Psychology, 18*(4), 219-222.