Using Interactive Tutors to Overcome Cognitive Learning Hurdles in Quantitative Reasoning

Charles L. Cappell
Northern Illinois University

April, 2001

The project reported in this paper was funded by the National Science Foundation, DUE-9752619.
Abstract

This paper focuses on several of the major conceptual hurdles encountered when trying to understand sampling. Portions of the Sampling Tutorial are demonstrated that address each of these hurdles. The tutor is designed as one component, the component emphasizing conceptual learning, of a course on Quantitative Research Methods. Major conceptual hurdles encountered when learning conceptual material for a module on sampling are outlined and related to more general hurdles in obtaining quantitative reasoning skills. A strategy for addressing these hurdles learners encounter is described, and learning architectures built into the Sampling Tutorial are demonstrated.
What are the Cognitive Hurdles in Quantitative Reasoning and Analysis (QRA)?

The discipline of mathematics has identified four major cognitive transitions that lead to "mathematical reasoning":

1. The transition from arithmetic to algebra: culminating in the ability to think symbolically and to manipulate symbols to solve problems.
2. The transition between algebra and geometry: culminating in the ability to think logically. Logical proof replaces calculation.
3. The transition from technique to theory: culminating in the ability to create proofs. Thinking shifts from a focus on solving a problem to creation of proofs of a more general nature.
4. The creative transition culminating in the ability to pose new problems needing solution. (AAC, 1991: Ch. 5: Mathematics).

Jones et al. (1995) used the Delphi method to survey a panel of experts that identified a set performance skills comprising an independent dimension of statistical critical thinking. In their broad analysis of critical thinking abilities for liberal arts majors, a fifth factor emerged from their analysis, which they labeled: Statistical Argumentation Skills (presented below with their factor loadings in parenthesis):

"Employ graphs, diagrams, hierarchical trees, matrices, and models (.77).
Apply appropriate statistical inference techniques (.69).
Assess statistical information used as evidence in an argument (.65).
Use multiple strategies in solving problems (.55).
Determine if a conclusion is based on an adequate/representative sample (.54).
Determine how data might confirm or challenge a conclusion (.53).
Assess how well an argument anticipates and responds to objections (.48).
Determine and judge the strength of a causal reasoning argument (.41)."
(Jones et al., 1995)

The parallel learning transitions (and associated hurdles) found in the quantitative/statistical thinking used in social science can be stated in equally general goal-oriented terms: What are the components of applied quantitative/statistical reasoning we want students to learn? I propose that contemporary treatment of undergraduate social science research methods should integrate the more general set of liberal arts critical thinking skills and QRA; to wit:

1. Critical Thinking Skills, the ability to think logically and to assess the quality of evidence. Identify logical fallacies and well-posed questions amenable to scientific inquiry. Assess the natural process, the sampling process, the measurement process that generate the data.
2. Formal, Symbolic Thinking Skills. Translate text and ideas into symbols, read algebra, create and manipulate symbolic models, concrete to formal reasoning.
3. Conceptual/Variable/Hypothetical Language (CVHL). Think in terms of hypothetical implications, scientific method, causality, use concepts of counterfactuals, sample space.
5. Statistical Modeling Skills. Create and evaluate statistical models representing a current condition or process. Apply useful statistical summary measures to descriptions and explanations: correlation and regression coefficients, conditional probabilities, odds, summary measures of association.
6. **Inferential Reasoning.** Think inferentially, to recognize the limits of generalization. The basic gamut of inferential statistics, sampling distributions, confidence intervals, p-values, hypothesis testing. (see also: AAC, Ch. 5, 1991; Browne, N., and Keeley, S., 2001; Chase, 2000; Jones, 1995; Mallows, 1998).

The foundation concepts are embedded mostly in probability theory and application. The overall goal of the lesson is to produce a conceptual map of basic sampling theory and procedures: centering around the concept of a random process and sample space. The cognitive QRA goals are to introduce the mathematical and logical thinking skills that accompany this basic understanding at a formal – operational level; that is a level of understanding where the learner can use algebraic notation and formulas; understand the more general and theoretical conceptual space in which QRA operates; and thereby, transfer what is learned in one local contextual domain to a wider range of problem solving situations.

**Development of an Interactive Tutor for QRA: Sampling**

The interactive tutor developed for sampling is designed as one of three pedagogical components that organize an introductory level Quantitative Methods course in the social sciences that I teach. The tutor carries the burden of the “conceptual learning” component. We are fortunate to have a dedicated computer lab for QRA training where the tutor can be run; it can also be distributed via CDROM and the internet. The lab also is the site of the second pedagogical component: hands on active learning using real (or simulated) data that gives the student the technical performance skills to implement the material being covered. To accompany the section of sampling, the students are given a sample frame and are taught how to use EXCEL to draw a random sample. The third component, classroom, still retains a component of the traditional lecture structure, but is used to pose a problem at the outset of each major lesson: in practical terms – the student’s next assignment, but also in general terms: giving a brief overview of the types of problems the material about to be introduced will help solve. This classroom component also helps to contextualize the material, and to motivate the students. So keep in mind, the strategy and example drawn from the tutor on sampling discussed below is only one component of the overall pedagogical strategy directed at the learner.

**Conceptual Hurdles**

In terms of declarative knowledge, or content knowledge, the hurdles are simply transferring the information presented into long-term knowledge. Conceptual maps have to be learned as well that link the concepts.

The problem of developing a formal-operational understanding of the material from the inchoate empirical representations, or metaphors, is the second major hurdle. This problem is evident when learners are not able to formally define knowledge or operations and transfer theme to new settings. In some of the evaluation quizzes I’ve given after a lesson on sampling; I’ve asked students to define in their own words, yet as technically as precise as they can, what is meant by a random sample. A typical response from a learner grounded in the empirical metaphor phase or thinking is: “It’s like drawing names out of a hat”. The learner has not articulated the more formal mechanism that defines this procedure as random.

**Sampling Tutor Learning Objectives**

The substantive goals for the learners are to:

1. Understand the 50 major concepts used in sampling.
2. Understand the different designs used well enough to recognize them in the literature; to know their strengths and weaknesses; and to implement the simpler forms.
The cognitive learning goals are to:
1. To lay a foundation of the concepts for random process and sample space from which inference can be learned.
2. Move the learner from concrete mastery to formal-operational thought regarding these concepts.

**Strategy**

**Contextual Framework for Motivation.**
The general idea is that learning occurs when the student is interested in a substantive story or problem. The tutor uses a variety of real world issues to attract some interest. (Shank, 1995).

Outline key concepts.
In the sampling tutor, around 50 key concepts were identified as a basic set of concepts students needed to master in order to understand how sampling is used in the social sciences and to prepare the groundwork for understanding inference. (Shau and Mattern, 1997).

**Concept Grounding.**
All views of cognitive development stress the grounding of the most complex cognitive operation in some inchoate experience that generated an initial “empirical metaphor”. This approximates what Piaget and Inhelder referred to as the “concrete” level thinking. More recently, some of the beginning foundations of mathematical thinking are viewed as having seeds in these empirical metaphors (Lackoff and Nunez, 2000).

**Conceptual Priority.**
Nearly every sentence uttered when teaching abstract QRA is ladden with presuppositions. The concepts need to be prioritized. The first time a concept is used, the designer of the lesson needs to ask: what concepts are prerequisites for the learner to integrate this new concept into their working knowledge base? This obviously can lead to “…it’s turtles..all the way down”, so at some point the lesson planner draws a boundary and declares all prior concepts as prerequisites to this lesson.

**Conceptual Maps.**
Priority is one form of conceptual map, or local “schema” that links concepts into a meaningful whole (Winn and Snyder, 1996; Schau and Mattern, 1996). A conceptual map is a mental representation that links the concepts into a network, at the deepest level, a neural network, that places concepts in such a way that recall is triggered (Gopnick and Meltzoff, 1998). At a higher level of abstraction, the concept takes part of its meaning, often the most significant part, from its relation to other concepts in the local concept map. New material has to be integrated into conceptual maps for it to be available in the future.

**The Strategy Implemented for a Segment of Sampling Tutor**
The cognitive learning goal of QRA is to develop “formal-operational” level thinking when applied to the modeling of the variable quality of social life and events. In the outline below, the links between concepts is made and the progression from concrete – empirical conceptualizations to formal – operational representations is discussed. Along with the discussion are snapshots of the portion of the tutorial designed to facilitate such learning. This illustration consists of just one “learning thread” through the lesson, albeit a major one.

List of Key concepts:
Random process
Sample space
Events and their Orderings (Permutations)
Factorial Operator
Size of Sample Space (Counting Formulas)
Random Process is the fundamental concept. This concept is introduced early and is grounded in "empirical metaphors": "drawing numbers out of a hat", "the rolling of a die".

The concept of a random process is linked to the concept of a sample space, that multiple realizations, or outcomes of such a process are possible. Any one observation, or generation of an event, is just one realization from the set of all possible events. A couple of graphical images are used to visually enhance the concept. Simulations demonstrate this empirically.

Orderings are introduced as particular types of sequences of events from a sample space. Again the concept is grounded in an empirical example, and a simulation allows the students to play with the orderings of a common set of life experiences could occur.

A formal representation, a counting formula (factorial operator), is introduced as a tool for quantifying the size of the sample space of permutations. A “rationale” accompanies the introduction, again using empirical metaphors, to give the learner a fuller understanding of how the formal operation relates to the earlier empirical representation of a sample space.

Other counting formulas for combinations are presented that build upon the first for permutations, again with their logical and mathematical rationale to help the learner make the transition from concrete representations to abstract, formal-operational understanding.

A flowchart of this thread running through the Sampling Tutor looks as follows:
At the conclusion of each lesson, an animation plays that summarizes the key concepts and their relations. It is this visual representation of the conceptual schema that we want to transfer to the learner’s cognitive structure.
Illustrations from the Sampling Tutor

The introduction and navigation menu use the same conceptual map diagram embedded in the pedagogical objectives.
The mapping of the concepts and the lesson’s learning objectives introduce each lesson.
Random sampling introduced with empirical metaphor

Recall that in an earlier lesson, we defined a study population to be the 25 largest cities in the U.S. If we want to draw a random sample from this population, we could create some mechanical procedure that would give each city the same probability of being selected.

With 25 cities in the population, applying the example procedure means that the probability of selection for any one city being chosen is 1 in 25 (1/25).

The intuitive procedure of putting all the names of the cities into a hat and selecting a sample from it works just fine...for small populations...if we have a hat.

This simple metaphor of drawing names or numbers out of a hat captures what is meant by a random sample.

A random sample is a subset of population elements selected according to objectives rules that make sure only the laws of probability govern which element is selected.

Computers can do the sampling work for us, even when the populations and samples are large.
Random process introduced by real time simulation using consistent context
Sample space introduced with formal and concrete visual representations.
A more abstract representation using Venn diagrams is then presented to reinforce the first empirical metaphor.
A concrete illustration is then presented that employs the idea of a sample space, here we use the idea of 3 Chicago Bulls players (from 5 players) forming the basis of the Triangle Offence.
Order and events introduced with visual representations in familiar context: common life events.

**Fundamental Sampling Concepts**

**Events and their Different Orders**

While we are primarily interested in you learning the logic behind counting formulas relevant to sampling theory, the study of the order in which events occur, or could occur, can be substantively interesting.

For example, consider the following 5 common events:

- Buy a dog.
- Graduate from high school.
- Have a baby.
- Get married.
- Get a full-time job.

Substantive applications of this idea can be found in the social science.
Simulation used to reinforce the formal concept of sample space, here for ordered events.

One can see by extension, that the final number of \textit{ordered arrangements} in this example is
$$5 \times 4 \times 3 \times 2 \times 1 \times 1 = 120$$

If we think and write symbolically, we can use the letter \(k\) to stand for the number of events, and the sequence becomes in symbolic notation:

$$k (k-1) (k-2) (k-3) (k-4) \cdots (k-(k-1))$$

To create one possible arrangement of the 120 orders click on one item in each of the following lists.

A. Graduate from high school
B. Get a full-time job
C. Buy a dog
D. Get married
E. Have a baby
Mathematical notation introduced, then the derivation is presented. Tutor moves from concrete representations to more formal – operational schemes.

**Fundamental Sampling Concepts**

We use a shorthand notation for expanding a number like this:

\[ k(k-1)(k-2)(k-3)(k-4) \ldots (k-(k-1)) \]

**This is simple mathematical shorthand.** The \( ! \) symbol in statistics is called the factorial operator: think of it as a mathematical operation that is performed on any integer, \( k \).

### A: Graduate from high school
### B: Get a full-time job
### C: Buy a dog
### D: Get married
### E: Have a baby

### A: Graduate from high school
### B: Get a full-time job
### C: Buy a dog
### D: Get married
### E: Have a baby

And so on, up to 120 different patterns for these 5 events.
Four major sampling selection designs are presented with accompanying counting formulas, each worked out on the same contextual problem of sampling 5 cities from the population of 25.

### Fundamental Sampling Concepts

#### Sampling with Replacement

<table>
<thead>
<tr>
<th>Order Matters (Permutations)</th>
<th>Order Does Not Matter (Combinations)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$N^k$</td>
<td>$\binom{N+k-1}{k} = \frac{(N+k-1)!}{k!(N-1)!}$</td>
</tr>
</tbody>
</table>

**N=Population Size.**

**k=Sample Size.**

In some texts, $n$, instead of $k$, is used to represent sample size.

#### Sampling without Replacement

| $\frac{N^k}{(N-k)!}$ | $\binom{N}{k} = \frac{N!}{k!(N-k)!}$ |

**Remember, in the social sciences, the most common sampling design is sampling without replacement, where the order does not matter.**

![Click on the notepad to take notes.](image-url)
Summary

This paper presented a brief overview of the architecture used in the sampling tutor to induce a fuller understanding of sampling theory and methodology at the formal – operational level of thinking. The lessons progressed from concrete empirical metaphors and representations to more formal, symbolic representations and operations. Simulations were integrated to reinforce the learning.

Other features of the tutor not covered in this brief presentation also are designed to induce the student to move from concrete thinking to the formal – operational level. Quizzes at the end of each lesson, and an overall “recitation” give the learner feedback on conceptual reasoning errors. In the recitation, the learner is given feedback on two dimensions of understanding: content (declarative – more concrete knowledge) and operational (relational and operational ability), rather than in substantive areas, in order to reinforce the metacognitive objectives of the tutor and to increase the learner’s awareness of the distinction between concrete and formal understanding.

Subjective evaluations of the tutor by learners have been overwhelming positive (Cappell and Haapoja, 2000). Some analysis of objective learning outcomes are underway that will assess not only content and performance ability, but also any change in the attitude learners in the social sciences may have toward their ability to learn quantitative reasoning and analysis skills after interacting with the tutor.
References

Bradstreet, Thomas.

Bravo, Maria.

Broers, Nick, J., and Berger, Martijn P., F.

Cappell, Charles, and Haapoja, Thomas.

Chance, Beth.

Cox, D.R.

Cuzzort, R.P., and Vrettos, J.S.

Dahaene, Stanislas

Earley, Mark.

Finney, Sara.

Garfield, Joan.

Geary, D. C.

Gelman, Rachel, and C. Randy Gallistel.

Gopnik, A., Meltzoff, A.N., and Kuhl, P.K.

Holcomb, John, P., and Ruffer, Rochelle L.

Jonasen, David. Ed.

Kahneman, D., Slovic, P., and Tversky, A.

Keeler, Carolyn, and Steinhorst, R. Kirk.

Koslowksi, Barbara

Kvam, Paul

Lakoff, George, and Nunez, Rafael.

Marasinghe, Mervyn G., Meeker, W.Q., Cook, Dianne, Shin, Tae-Shin.

Mathews, Colin

Minium, Edward W., R.C. Clarke, and T. Coladarc.

Nasser, Fadia M.

Newton, H. Joseph, and Harvill, Jane.

Parr, William, C., and Smith, Marlene.

Phillips, J.L.

Piaget, J. and Indelder, B.

Resnick, L. B.

Rogers, Bruce.

Rumsey, Deborah.

Seaman, Michael.

Schau, Candace, and Mattern, Nancy.
1997 "Use of map techniques in teaching applied statistics courses.” The American Statistician 51:171-175.

Velleman, Paul.
1997 ACTIV STAT (software). Reading, MA: Addison Wesley Interactive
Wachsmuth, Ipke.  

Winn, William, and Snyder, Daniel.  

Wisenbaker, Joseph, M., and Douzens, Cordelia.  

Wynn, Karen.  