# Knowledge and the Simultaneous Conjoint Measurement of Activity, Agents, and Situations[1]

**Peter Pirolli**
Xerox Palo Alto Research Center and University of California, Berkeley
and
**Mark Wilson**
University of California, Berkeley

## Abstract

We outline a measurement theory developed by integrating ideas about knowledge level analysis, production system models of transfer, additive conjoint measurement, and Rasch models of measurement. Productions are assumed to represent situation-action elements of knowledge. The model views the performance of such a knowledge element as the combination of affordance properties associated with the element and ability properties associated with an individual. Under specified conditions, observed behavior can be used to separate and quantify variables measuring situation-action affordances and subject abilities. A specific version of this model is applied to data from four studies involving the CMU Lisp Tutor.

## Introduction

We seek to develop a theory of quantitative measurement that captures the intendedly rational behavior that results from bringing individuals and situations into contact. In this theory, knowledge, agents, goals, and situations are constitutively defined, but can be assigned independent quantitative measures under appropriate conditions. The theory is grounded in cognitive models of transfer and fundamental theories of measurement. We present an overview of this theory and a measurement model applied to a corpus from four studies using the CMU Lisp Tutor (Anderson, Conrad, & Corbett, 1989).

Outside of cognitive science, the pursuit of more direct quantitative measurement of theoretical constructs has a long history (Campbell, 1928) because it is such a strong marker of understanding and control of over the phenomena of interest (Michell, 1990). Many cognitive scientists are unfamiliar with models of fundamental measurement, such as *additive conjoint measurement*, and we need note, at the outset, that this means something much deeper in our scientific ontology and epistemology than just assigning numbers to observations. Such measurement implies strong hypotheses that essentially state that properties and relations in the world behave as recognizable mathematical structures that we call quantitative.

## Knowledge = Observer-ascribed Intention

The *knowledge level* was offered (Newell, 1982) as a new systems level, above the *symbol level*, to describe intelligent systems. Knowledge-level systems can be specified completely by an observer examining a system's interaction with the external world. A knowledge-level system consists of an *agent* behaving in an *environment*. The agent consists of a set of *actions*, a set of *perceptual devices*, a *goal*, and a body of *knowledge*. The operation of such systems is governed by the *principle of rationality*: if the agent knows that one of its actions will lead to a situation preferred according to its goal, then it will *intend* the action, which will then be taken if it is possible. Knowledge refers to *intentions* or *beliefs* and it is defined functionally as "whatever can be ascribed to an agent, such that its behavior can be computed according to the principle of rationality" (Newell, 1982, p. 105). The reason for specifying knowledge as a function that does not reside in any particular structure at the symbol level is that knowledge about the world cannot be captured in finite structure. The knowledge level is defined from the stance of an observer ascribing knowledge to a system based on observable external situations and actions.

## The Symbol Level and Knowledge Access

The symbol level is the medium of symbol structures that run in a physically realizable computational architecture. Relations among the knowledge and symbol levels can become complex. In general outline (Newell, 1982), a *representation* scheme is defined at the symbol level as a combination of data structures and processes specified in some architecture. Knowledge representation schemes determine the access functions to the knowledge available to the system in a given situation, as well as their cost structure. The principle of rationality is mechanistically realized by the total operation of the symbol level system.

Newell (1982, p.108) found it useful to think of knowledge as an infinite table representing a relation

[1]This work was supported by Office of Naval Research Grant N00014-91J-1523. We thank Karen Draney who performed the data analysis.

and containing elements in which situations and actions were conditionally associated according to goals. A compact way to think of the relationship between knowledge and symbol levels is provided by Rosenbloom and Aasman (1990). Consider the infinite table of knowledge mentioned above. At the knowledge level, the principle of rationality simply operates on the entire infinite table. At the symbol level, for any given environmental context, the access to that knowledge can be imagined as a degree-of-belief function over that table. The values of this degree-of-belief function determine the degree to which each knowledge element is plausible for that context, and this degree-of-belief value predicts the ease with which the knowledge element can be brought to bear within that context. Changes to this access function may occur through learning.

## Measurement at the Knowledge Level

The basic assumption about formal representations of the knowledge level is that "to ascribe to an agent the [symbol] structure *S* is to ascribe whatever the observer can know from [symbol] structure *S*" (Newell, 1982, p. 112, italics in original). In any particular model we need not worry about the complete unbounded knowledge potential. Production systems could be used as the basis for the logic for the knowledge level analysis. Many studies (Pirolli, 1991; Polson, Bovair, & Kieras, 1987; Singley & Anderson, 1989) suggest a very simple and direct correspondence between formal production rules and elements of knowledge which results in an identical elements theory of transfer (Singley & Anderson, 1989). The following are some key properties of the Singley-Anderson formulation that are relevant to the knowledge level analysis:

- *Permanence property.*[2] Productions, once acquired, remain in the system.
- *Abstractness property.* The production rule conditions are patterns that variabilize over the space of possible situations. The production system maps observable situations and goals onto the conditions of productions. Each production therefore defines an equivalence class of situations. The production conditions may be regarded as specifications of situational invariants controlling cognition and behavior.
- *Independence properties* Each production rule can be acquired and can transfer independent of other productions.
- *Asymmetry property.* Each production specifies a state change in the interaction of agent and environment. The system can reason from conditions to actions but not vice versa.

To which we add the following restriction to constrain the system to knowledge-level analysis

- *Knowledge-level restriction.* Only productions (or specifiable compositions of production execution structures) that match identifiable external situations and yield identifiable external actions are relevant to the knowledge-level analysis.

We need to also define properties indicating the extent to which observed agents approximate knowledge-level systems. For the purposes of developing a broad measurement model, we seek to make the weakest possible assumptions about knowledge access. We will concentrate on knowledge access properties in relation to individual differences, skill acquisition, and practice:

- *Affordance properties.* Individual productions have properties that reflect the performance of a particular situation-action knowledge element. We may also refer to the opposite of affordance as *difficulty*.
- *Ability properties.* Individual agents differ in their ability to exhibit knowledge.

## Intelligent Tutoring Systems as Knowledge-ascribing Instruments

Concrete examples of knowledge level analyses are provided by *overlay* student models in intelligent tutoring systems. An overlay model is often specified in the form of a production system model that can solve problems in a particular domain such as medical diagnosis, geometry, or Lisp. The intelligent tutoring system (ITS), in addition to its pedagogical role, is a complex observational instrument. Based on its mechanical observation of different interface situations with a student (e.g., the computer screen state) the ITS uses its production system model to predict possible actions, and observes the actual student actions. The ITS records external observable situations and actions. The observed situation-action mappings are then matched against the mappings embodied in the production rules. The ITS basically has an internal table of productions that captures the elements of knowledge that are possible (both within and across students), and it ascribes these knowledge elements to the observed student agent when the student exhibits the appropriate behavior. This process is depicted in Figure 1: Behavior over time is matched to productions (P1, P2, and P3), and response measures (Ri) associated to a particular production are tabled by trials on that production (as for P1 in Figure 1). The ITS knows the mapping of situation to action implied by a production symbol structure *S*, and mechanically fulfills Newell's (1982) role of observer: "to ascribe to an agent the [symbol]

[2]This property is not explicitly stated by Singley and Anderson, but it is clearly part of their theory.

structure $S$ is to ascribe whatever the observer can know from [symbol] structure $S$" Thus, in practice, an ITS can be viewed as an automated knowledge-ascribing instrument that treats a student as a knowledge-level system.

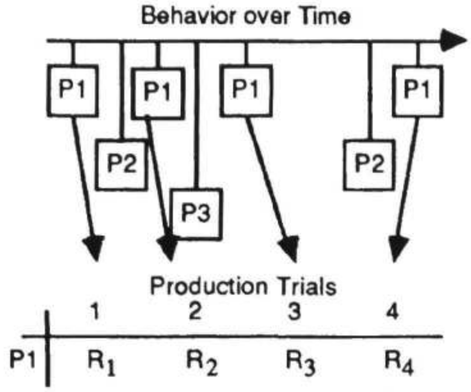


**Figure 1. Behavior over time is matched to productions and trial.**

## Theories of Fundamental Measurement

Fundamental measurement concerns the quantification of observed relations or properties (Campbell, 1928; Luce & Tukey, 1964). The marks of quantity are established by ordinal relations and additive structure among the variables of interest. In the case of extensive properties, such as length, specific ordinal relations are clearly manifest in comparisons of objects of different lengths and additivity is manifest in the manner in which lengths can be concatenated to produce new lengths. More abstractly, the marks of quantity can be established by determining if the observed relations or properties conform to algebraic structures that satisfy specific axiomatic conditions (Michell, 1990).

Both abilities and affordances are reflected in the same performance, so our knowledge-level assumptions provide *constitutive definitions* of the situation-action affordance, individual ability, and the measurable performance. Unlike extensive properties, such as length or mass, constitutively defined variables raise the issue of separating out and establishing the quantities associated with each of the constitutively defined concepts. This too can be achieved, as in the theory of additive conjoint measurement (Luce & Tukey, 1964), which establishes the axioms that must be met to establish the appropriate order and algebraic structure to quantify constitutively defined (conjointly measured) variables. The general idea to separate out the measures for constitutively defined variables is familiar to anyone who has used the additive factors logic of experimental design. If variables defined over the two classes of entities and the resultant response variable can be simultaneously scaled so that an ordinal additive (noninteractive) structure results, then one can separate the scales associated with the variables. Imagine that an ITS uses an overlay model to dichotomously score actions made by a person $i$ and in a situation class corresponding to the conditions specified by production $j$, such that $X_{ij} = 1$ scores that "subject $i$ acted as if they had knowledge $k_j$ appropriate for situation $j$" otherwise $X_{ij} = 0$.[3] Assume that we simply associate a parameter $z$ with each individual to indicate level of ability and each situation-action knowledge element will be associated with a parameter $d$ indicating its difficulty (the inverse of the degree to which it affords understanding). One might assume a simple stochastic model in which the probability of a person $i$ exhibiting the appropriate knowledge in situation $j$ is some function of $z_i/d_j$. The following is a simple model with the necessary properties for illustration. Let

$$\Pr(X_{ij} = 1 | z_i, d_j) = \frac{z_i}{(z_i + d_j)} \quad (1)$$

and the complement is

$$\Pr(X_{ij} = 0 | z_i, d_j) = \frac{d_j}{(z_i + d_j)} \quad (2)$$

Examination of this formulation shows that a person of zero ability, $z_i = 0$, has a zero probability of exhibiting knowledge in all situations, whereas a person of infinite ability, $z_i = \infty$, has a probability of one of exhibiting knowledge in all situations. A situation-action element of zero difficulty, $d_j = 0$, has a probability of one of being successfully performed across individuals of all ability, whereas an element of infinite difficulty, $d_j = \infty$, has a zero probability of being successfully evoked in anyone. If the difficulty of the situation is equal to the ability of the individual, $z_i = d_j$, then there is a .5 probability of it being successfully performed. The estimation of these parameters becomes realizable through the observation of the relative proportions of correct and incorrect solutions across individuals and problems. This is because the relative proportion of correct to incorrect solutions is predicted by the combination of Equations 1 and 2 to be

$$\frac{\Pr(X_{ij} = 1)}{\Pr(X_{ij} = 0)} = \frac{z_i}{d_j} \quad (3)$$

We can look at a particular situation $j$ and use as its parameter the relative proportion of individuals scoring as successfully having the relevant knowledge. Alternatively we can look at a particular

[3]This example is a modification of one discussed by Rasch (1960).

individual and examine their relative proportion of success to estimate that person's parameter. A particular situation (or individual) can be chosen as an arbitrary zero point of reference, and another situation (or individual) as the unit. Although the example here examines probability correct, the analysis can be easily mapped on to performance time.

Note that we can transform the metric of $z_i$ and $d_j$ above to a logarithmic metric and achieve an additive structure. Let $\theta_i = \log(z_i)$ and $\xi_j = \log(d_j)$, so that Equations 1, 2 and 3 become:

$$\Pr(X_{ij} = 1 | \theta_i, \xi_j) = \frac{\exp(\theta_i - \xi_j)}{1 + \exp(\theta_i - \xi_j)}, \quad (4)$$

$$\Pr(X_{ij} = 0 | \theta_i, \xi_j) = \frac{1}{1 + \exp(\theta_i - \xi_j)}, \quad (5)$$

and

$$\frac{\Pr(X_{ij} = 1 | \theta_i, \xi_j)}{\Pr(X_{ij} = 0 | \theta_i, \xi_j)} = \exp(\theta_i - \xi_j) \quad (6)$$

This is a form of the logistic or Rasch model that is more commonly used, and we will use in what follows. The values of $\theta$ and $\xi$ are reported as logits.

## Rasch Models

Rasch (1960) developed a measurement theory with similarities to additive conjoint measurement but based on assumptions of stochastic processes. We use a Rasch framework in our measurement model. Rasch (1960) established that sufficient statistics could be obtained quite simply for the parameters in his models. Estimators based on such statistics fulfill the dual role of establishing the empirical conditions under which a model applies, and provides the underpinnings for statistical estimation and inference. *Rasch models* satisfy an important property of *specific objectivity*. Specific objectivity means that a response measure is a conjoint measure of two entities (such as an agent and situation) whose measures can be separated and quantified, similar to the manner discussed above. Rasch stated that specific objectivity holds when

> the result of any comparison of two objects [persons] ... *is independent of everything else within the frame of reference other than the two objects* [persons] *which are to be compared* (Rasch, 1977, p.77, italics in original).

That is, the parameter describing the person must be inferentially separable from the parameters describing the situation-action knowledge element. This must hold, in a dual fashion, for comparisons of situation-action knowledge elements.

It can be shown that Rasch models satisfy the conditions of additive conjoint measurement. Because of space limitations, we leave that proof to a formcoming paper. It is important to observe that this proof is shown to hold for the *determinate* Rasch model (Rasch, 1977. p.6). It remains to be shown whether in any actual situation the Rasch model holds, and evidence for all such real situations will always be empirical rather than theoretical (i.e., the conclusion that "the Rasch model holds" or otherwise, will for all such situations be based on weight of empirical evidence, rather than be true or false according to some theory). Numerous procedures for testing fit to the Rasch model have been given in the literature. This is paralleled by the search for "indirect evidence" (Michell, 1990, pp. 78-84). That is, the evidence that an additive conjoint structure holds in a given (sufficiently large and interesting) situation can only be based on a weight of empirical evidence, just as for the Rasch model.

## Application of the Model

Programming is perhaps the most studied cognitive skill. Production system models of this domain reveal remarkable regularities in performance, learning, and practice. We begin by noting that Anderson et al. (1989) report factor analyses of the Subject by Production matrices of error rates. Their findings suggest that subjects and productions have independent additive effects (under appropriate transformation) on error rates. This provides some indirect evidence that the general Rasch approach requiring separable parameters for agents (people) and knowledge elements is appropriate.

Clearly people vary in their abilities, and we will assign each person $i$ a parameter $\theta_i$. Performance varies across productions rules, and we will assign each production $j$ a parameter $\delta_j$ to indicate its difficulty. Particularly important to us is the observation of practice effects for individual productions. Performance time and errors improve as a sublinear function of practice. Error rates improve as exponential or power functions of trials of practice (the specific form may depend on a number of other factors, Pirolli, 1991). We will find it useful to think of each trial of practice as an increment that improves performance. The same increment can be applied to model each trial, if the practice function is transformed into a space in which it is linear. If we let $\alpha$ be the increment, then on trial $t$ there will be $(t - 1)\alpha$ incremenets received by a production. We will denote this sum by $\alpha_{t-1}$, so for trials 1, 2, ...$n$ the difficulty $\delta_j$ of production $j$ will be improved by $\alpha_0 = 0$, $\alpha_1 = \alpha$, $\alpha_2 = 2\alpha, \ldots, \alpha_{n-1} = (n - 1)\alpha$.

Another relevant finding for us is that various kinds of treatments can have specific impact on specific productions. For instance, instructional examples

| | Production Trials | | | | |
|---|---|---|---|---|---|
| | 1 | 2 | 3 | ••• | n |
| $P_1$ | $\theta_1+\delta_1+\tau_0+\alpha_0$ | $\theta_1+\delta_1+\tau_0+\alpha_1$ | $\theta_1+\delta_1+\tau_0+\alpha_2$ | ••• | $\theta_1+\delta_1+\tau_0+\alpha_{n-1}$ |
| $P_2$ | $\theta_1+\delta_2+\tau_0+\alpha_0$ | $\theta_1+\delta_2+\tau_0+\alpha_1$ | $\theta_1+\delta_2+\tau_0+\alpha_2$ | ••• | $\theta_1+\delta_2+\tau_0+\alpha_{n-1}$ |
| ⋮ | | | | | |
| $P_k$ | $\theta_1+\delta_k+\tau_0+\alpha_0$ | $\theta_1+\delta_k+\tau_0+\alpha_1$ | $\theta_1+\delta_k+\tau_0+\alpha_2$ | ••• | $\theta_1+\delta_k+\tau_0+\alpha_{n-1}$ |
| | ↓ Different subject (θ) and different treatment (τ) on $P_2$ ↓ | | | | |
| $P_1$ | $\theta_2+\delta_1+\tau_0+\alpha_0$ | $\theta_2+\delta_1+\tau_0+\alpha_1$ | $\theta_2+\delta_1+\tau_0+\alpha_2$ | ••• | $\theta_2+\delta_1+\tau_0+\alpha_{n-1}$ |
| $P_2$ | $\theta_2+\delta_2+\tau_1+\alpha_0$ | $\theta_2+\delta_2+\tau_1+\alpha_1$ | $\theta_2+\delta_2+\tau_1+\alpha_2$ | ••• | $\theta_2+\delta_2+\tau_1+\alpha_{n-1}$ |
| ⋮ | | | | | |
| $P_k$ | $\theta_2+\delta_k+\tau_0+\alpha_0$ | $\theta_2+\delta_k+\tau_0+\alpha_1$ | $\theta_2+\delta_k+\tau_0+\alpha_2$ | ••• | $\theta_2+\delta_k+\tau_0+\alpha_{n-1}$ |

**Figure 2. A matrix depicting parametric variations with people, practice, and treatments for a set of knowledge elements or productions, $P_i$.**

improve learning to program, but more importantly an example can be analyzed to determine the specific productions that will be improved (Pirolli, 1991). Learners may employ different learning strategies when processing instructional materials, and these have specific effects on production performance (Pirolli & Recker, in press). Each such specific treatment $k$ will be associated with a parameter $\tau_k$ that alters the difficulty of performing a specific production.

We can elaborate the model in Equations 4 to 6 by expanding $\xi$ to be a linear combination of the relevant production parameter $\delta$, treatment parameter $\tau$, and practice increment $\alpha_{t-1}$, or $\xi = \delta + \tau + \alpha_{t-1}$. Figure 2 illustrates a hypothetical analysis of performance on productions $P_1, P_2, ..., P_k$, over trials 1, 2, ... $n$ for different subjects under different treatments. In the top table of Figure 2, performance is attributed to a person parameter ($\theta_1$) and production parameters $\delta_1$, $\delta_2$, ..., $\delta_k$. Performance is improved by practice effects $\alpha_0 = 0$, $\alpha_1 = \alpha$,..., $\alpha_{n-1} = (n-1)\alpha$. We also assign a treatment parameter $\tau_0$ to this base treatment. The bottom table of Figure 2 shows how we would indicate a different subject ($\theta_2$) who received some treatment ($\tau_1$) that specifically affects performance of production $P_2$ (perhaps an example program).

We have developed and applied such a model to a corpus of data from several studies involving the CMU Lisp Tutor. The formal expression of the measurement model is

$$f_i(x;\mathbf{A},\mathbf{B},\xi|\theta) = \frac{\exp(\mathbf{b}'_{ix}\theta + \mathbf{a}'_{ix}\xi)}{\sum_{u=1}^{K_i}\exp(\mathbf{b}'_{iu}\theta + \mathbf{a}'_{iu}\xi)} \quad (7)$$

where $\mathbf{A}$ is a design matrix describing how the data relate to productions, trials of practice, and treatment, and whose rows are $\mathbf{a}_{ix}$, $\xi$ is a vector of parameters that describe the production, trial, and treatment parameters, $\mathbf{B}$ is a score matrix describing how the data relate to the subjects and whose rows are $\mathbf{b}_{ix}$, and $\theta$ is a parameter that describes each subject. A detailed description of this model and the marginal maximum likelihood algorithm used to estimate its parameters is given in Adams and Wilson (1992). The model is applied to particular circumstances by specification of the $\mathbf{A}$ and $\mathbf{B}$ matrices. Pirolli and Wilson (1992) also provide some simple illustrative examples.

The data come from four experiments (Bielaczyc, Pirolli, & Brown, 1991; Pirolli & Recker, in press; Recker & Pirolli, 1992) investigating people learning to program recursive functions. Although conditions vary across experiments, all used the same programming problems (though possibly in different sequences) and involved the same production system models. For our first tests of the measurement model, we selected 22 productions for analysis across the 76 participants in the studies. These productions represented new knowledge learned in the recursion lesson or elements that could be induced from the instructional examples given to subjects. One set of subjects ($N$ = 32) saw an example illustrating recursion on numbers, another set of subjects ($N$ = 44) saw an example illustrating recursion on lists. Some productions that could be induced from the numeric recursion example but not the list recursion example we called *number productions* (4). Another set of productions could be induced from the list recursion example but not the numeric recursion example was called *list productions* (5). Some productions could be induced from both examples (1) and some could be induced from neither example (12).

Application of the model in Equation 7 to these data shows that the production difficulties ($\delta$) vary from -1.98 logits to 1.57 logits, with the number productions having a mean difficulty of .41 logits, the list productions -.47 logits, and the productions available from neither example had a mean difficulty of -.02 logits. The mean of abilities ($\theta$) for the subjects seeing the number recursion example was -1.07 logits and for those seeing the list recursion example was -1.26 logits. Seeing a number recursion example had a treatment effect ($\tau$) of -.36 logits on the number production difficulties and seeing the list recursion example treatment effect was -..22 logits. Each trial of practice ($\alpha$) had an effect of -.10 logits. Figure 3 presents the predicted probability of error (in log-linear coordinates) based on these parameter estimates. Examples improve the error estimates on the productions they affect, and this effect is two to four times the practice increment on a logit scale (see also Pirolli, 1991).

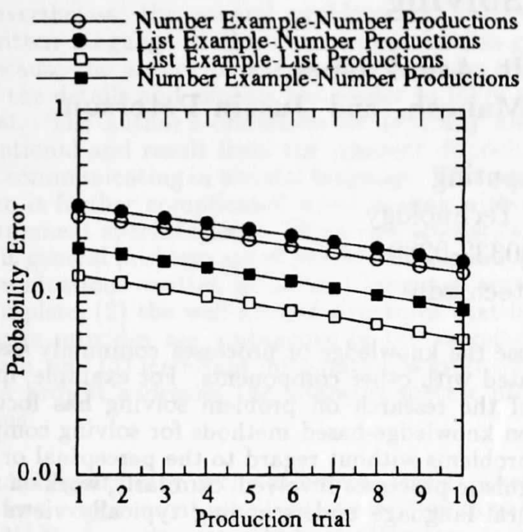


**Figure 3. Estimated learning curves for the CMU Lisp Tutor studies (log-linear scale).**

## General Discussion

Our measurement model integrates assumptions about the observation of knowledge, thé elementary nature of knowledge and its transfer and practice properties, and theories of fundamental measurement. The measurement model estimates the affordance properties of individual elements of knowledge and the ability properties of individual subjects on scales that are fundamentally quantitative (have order and additivity). Unlike analysis of variance, regression, or similar techniques, the model estimates are not of samples or populations, but the specific entities of interest. In essence the model treats knowledge as the combined additive effect of situation-action affordances, individuals, and their joint histories in terms of practice or treatments. We have treated these parameters as scalar properties, but it is possible to extend the model so that abilities and affordances are treated as multifaceted structures (i.e., matrices, see Adams & Wilson, 1992).

## References

Adams, R. J., & Wilson, M. 1992. *A random coefficients multinomial logit: Generalizing Rasch models*. Paper presented at the American Education Research Association, San Fransisco, CA.

Anderson, J. R., Conrad, F. G., & Corbett, A. T. 1989. Skill acquisition and the LISP tutor. *Cognitive Science* 13:467-505.

Bielaczyc, K., Pirolli, P., & Brown, A. 1991. The effects of training in explanation strategies on the acquisition of programming. Technical Report, CSM-3. School of Education, University of California.

Campbell, N. R. 1928. *An account of the principles of measurement and calculation*. London: Longmans, Green, and Co.

Luce, R. D., & Tukey, J. W. 1964. Simultaneous conjoint measurement: A new type of fundamental measurement. *Journal of Mathematical Psychology* 1:1-27.

Michell, J. 1990. *An introduction to the logic of psychological measurement*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Newell, A. 1982. The knowledge level. *Artificial Intelligence*18:87-127.

Pirolli, P. 1991. Effects of examples and their explanations in a lesson on recursion: A production system analysis. *Cognition and Instruction* 8:207-259.

Pirolli, P., & Recker, M. in press. Learning strategies and transfer in the domain of programming. *Cognition and Instruction.*

Pirolli, P., & Wilson, M. 1992. Measuring learning strategies and understanding: A research framework. In C. Frasson, G. Gauthier, & G. I. McCalla (Ed.), *Intelligent Tutoring Systems* (pp. 539-558). Berlin: Springer-Verlag.

Polson, P. G., Bovair, S., & Kieras, D. 1987. Transfer between text editors. In J. M. Carroll, & P. Tanner (Ed.), In*Proceddings of CHI '87 Human Factors in Computing Systems and Graphics Interface Conference.*

Rasch, G. 1960. *Probabilistic models for some intelligence and attainment tests* . Copenhagen: Danish Institute for Educational Research.

Rasch, G. 1977. On specific objectivity: An attempt at formalizing the request for generality and validity of scientific statements. *Danish Yearbook of Philosophy* 14:58-94.

Recker, M., & Pirolli, P. 1992. Student strategies for learning programming from a computational environment. In C. Frasson, G. Gauthier, & G. I. McCalla (Ed.), *Intelligent Tutoring Systems* (pp. 382-394). Berlin: Springer-Verlag.

Rosenbloom, P. S., & Aasman, J. 1990. Knowledge level and inductive uses of chunking (EBL). In (Ed.), In *Proceedings of the Eighth National Conference on Artificial Intelligence.*

Singley, M. K., & Anderson, J. R. 1989. *Transfer of cognitive skill*. Cambridge, MA: Harvard University Press.