DOCUMENT RESUME

ED 334 996 IR 015 163

AUTHOR Misanchuk, Earl R.; Schwier, Richard

TITLE Interactive Media Audit Trails: Approaches and Issues.

PUB DATE 91

NOTE 26p.; In: Proceedings of Selected Research Presentations at the Annual Convention of the Association for Educational Communications and Technology; see IR 015 132.

PUB TYPE Viewpoints (Opinion/Position Papers, Essays, etc.) (120) -- Speeches/Conference Papers (150)

EDRS PRICE MF01/PC02 Plus Postage.

DESCRIPTORS Data Collection; Feedback; Formative Evaluation; *Hypermedia; *Instructional Design; *Learner Controlled Instruction; *Multimedia Instruction; Qualitative Research; Research Methodology

IDENTIFIERS *Audit Trails

ABSTRACT

This discussion of uses for audit trails in instructional research begins by pointing out that interactive media provide the learner with the opportunity to shape the program, and consequently, the learning experience. The paper focuses on one of the new questions for instructional designers that have come with the advent of these technologies, i.e., the determination of the effects of taking different paths through instruction. The audit trail is described as comprising all the responses generated by a learner going through interactive or hypermediated instruction, and the nature of audit trails is discussed in the context of the different types of program structures in which they are generated: linear; branching; and multimedia and hypermedia structures. Three distinct purposes to which audit trails can be put are identified and discussed: (1) as data collection devices for formative evaluation in instructional design; (2) as tools for basic research into the instructional design of computer-based instruction and hypermedia; and (3) as a means of auditing usage of mediated presentations in a public forum. Ways in which the audit trail could be used in quantitative research are considered, e.g., in a raw data matrix, to present data associated with each decision point (node) in the treatment, to represent non-parametric data in a Petit-Point Pattern, or as an audit trail tree combining graphical representations with numerical accuracy. Qualitative approaches are also discussed, e.g., content or document analyses and case studies. It is suggested that inferential approaches to the audit trail would be used when comparisons are being made between groups of learners or treatments. Advantages and limitations of the audit trail approach are considered, and a discussion of the issues and challenges raised by this approach concludes the paper. Seven figures are appended. (13 references) (BBM)

Title:

Interactive Media Audit Trails: Approaches and Issues

Authors:

Earl R. Misanchuk
Richard Schwier

In any independent learning experience, from traditional print to computer assisted instruction, the learner strikes a path through the medium. In linear media, such as audiotaped instruction, that path is fixed and predetermined by the producer; learners share essentially similar instructional experiences. Instructional designers of such media have been able to assume a certain constancy to their instruction.

Interactive media, on the other hand, provide the learner with the opportunity to shape the program, and consequently the learning experience. Two learners may be able to traverse the same instructional materials, perhaps even achieve the same objectives, yet not share a common experience. At minimum, the degree of commonality in the experience is reduced over that experienced with the more linear media. By their very nature, interactive media are programs intentionally designed in segments, in which learner responses to structured opportunities such as menus or questions influence the sequence, size, and shape of the program (Schwier, 1987). The paths individuals strike through interactive media vary; we can assume that most learners will not share common learning experiences.

While obviously having a liberalizing (some would say humanizing) effect on instruction, the advent of interactive and hypermediated technologies brings with them a host of new questions for instructional designers, as well. Among these is the determination of the effects of taking different paths through instruction. To address such a question requires that there be some way of recording and analyzing the instructional path taken by each learner. We refer to that record of the path as the audit trail (after M. W. Petruk, personal communication, February 7, 1990).

It is a relatively simple matter to collect information about an individual's path using contemporary computer-based instruction authoring environments (e.g., *Authorware Professional* or *Course of Action*) or program development tools such as Macintosh's HyperCard. Learners' responses can be trapped and recorded at each decision point, or node. Of course, the descriptive data captured can provide very useful formative evaluation information. However, we speculate that there are additional purposes to which such information can be put, purposes that might produce generalizable rules about the construction of various paths through interactive and hypermediated instruction.

The Audit Trail

An audit trail comprises all the responses generated by a learner going through interactive or hypermediated instruction. In general terms, audit trails contain words, phrases, sentences, and paragraphs that a learner types into a computer, as well as a record of the "multiple choice"-like responses made, either from the keyboard or via some other input device, such as a mouse or touch-sensitive screen. However, to keep this preliminary discussion relatively simple, we restrict our discussion here to the "multiple choice"-like responses. Hence, we conceive of the audit trail as a string of characters (numerals or letters) representing choices made by learners as they progress through choice points—or nodes—within the instruction. For example, suppose the first point at which a learner had to make a choice had three different paths the learner could follow, and let us further suppose that she chooses the second path. As she proceeds through the subsequent instruction, she encounters a second node, with two choices, and chooses the first of those. The first two characters of her audit trail would therefore be 21 (read "two-one", not "twenty-one"). If she then chooses the first of three paths at the next node, and the third of three at the following node, her complete audit trail up to this point would be 2113. A second learner might have chosen the third path at the first node, the first path at the second node, the first path at the third node, and the second path at the fourth node, to give an audit trail of 3112.

In creating an audit trail, the computer would simply record the numerical value assigned to each option presented at a decision point, in a vector-like arrangement. As the number of decision points in the treatment increases, so will the length of the vector; as the complexity of the path taken by an individual increases (e.g., exploring optional paths as opposed to forging straight ahead), so will the length of the vector.

Audit Trails and Program Structures

Audit trails differ in nature according to the program structures in which they are generated. The structures described below represent the simplest cases; in reality, they are often combined within programs.

Linear Structure

In the most basic structure—a linear one (Figure 1a)—all learners necessarily go through the same experiences. An audit trail from a linear structure is quite simple, if not trivial. It can only contain such information as how many tries were made before the desired response was attained, and/or how long it took to attain it.

Insert Figure 1 about here.

Branching Structures

In a more complex structure—a branching one—the audit trail is correspondingly more complex, as well. In this situation, the learner is presented with choice points, or nodes, at which different responses will occasion different alternative paths through the instruction. In some structures, the paths will re-converge at the same point at which they diverged (as when a feedback loop provides remediation or supplemental information, then sends the learner back to the original node for another try) (Figure 1b); in another structure—dubbed the learner controlled parallel path, or LCPP, structure—the various (linear or branching) paths will run parallel to one another, and convergence comes somewhere down the line (Figure 1c). What distinguishes the feedback loop path from the LCPP is where the learner ends up after passing through it. In the feedback loop, the learner exits the loop back at the node of departure; in the LCPP, the learner exits the loop at some point further down the instructional sequence.

In LCPP structures, the audit trail is more complex than in the linear structure insofar as the learner's choice must be recorded. If the parallel structures are linear, then once again only the number of unsuccessful attempts and the length of time taken to criterion can be recorded in conjunction with the choice made. If, however, the branches themselves contain branches, another level of complexity obtains, and the problem of dependency shows itself: The meaning of any character in the audit trail is dependent for meaning upon the character preceding it. That is, the character '1', the second character in the first individual's audit trail, does not mean the same thing as the character '1' which is the second character in the second individual's audit trail, because the two individuals chose different paths at the first node.

In the feedback structure, a different problem is created by repeat visits to the node. An individual may or may not pass through the same node in the program more than once. Stated differently, the learner may traverse portions of a path repeatedly. If a simple frequency count is made of visits to a node, it is difficult to discriminate, for example, between the case in which two learners traversed the same path segment once each, and one in which a single learner traverses the same path segment twice. We refer to this problem as the looping problem. The looping problem is context-dependent: Interactive treatments are typically designed to permit repeated visits to some portions of instruction but not others. A second implication of this phenomenon is that not all audit trails will be of the same length.

Multimedia and Hypermedia Structures

Multimedia and hypermedia structures represent the most complex case of all, and both the dependency and looping problems manifest themselves. These structures are difficult to describe because they are so variable, both in terms of the number and kind of choices available to the learner within each display, and in terms of where the learner may go next. It is therefore not possible to draw a single generalizable flow chart for hypermedia as it is for a feedback loop or an LCPP; hence the one in Figure 1d is merely one of many possible basic structures.

We suggest that there are at least two levels of hypermediated structures: those in which cycling is permitted through relatively structured paths; and those in which paths are almost completely unstructured. While Nelson's original description of hypermedia closely matches the latter situation (Rezabeck & Ragan, 1989) that which is commonly termed multimedia matches the former description of hypermedia. The primary difference between the two levels of hypermediated structures is one of how much control is exercised by the program author and how much is given to the learner. Indeed, although the various descriptions of structures given above closely parallels their historical emergence, it also describes a continuum of control. In linear structures, the instructional designer/programmer exercises almost total control of the path of the learner (see Figure 2); in Nelsonian hypermediated structures (the second kind), almost no control is exercised by the instructional designer/programmer (except through what path or feature options are made available to the learners).

Insert Figure 2 about here.

We caution that the figure ought not to be interpreted too literally. We do not have any reason to believe, for example, that *all* LCPP structures grant half the control to the designer and half to the learner; nor can we state unequivocally that feedback loops are located exactly half-way between linear and LCPP structures. In other words, the horizontal scale should not be interpreted as being anything more than ordinal, and we are not certain that the horizontal axes representing the degree of designer control and the degree of learner control should be exactly the same length.

The audit trails of all structures described above suffer from the problem of conceptual distance. This is a classical measurement problem which surfaces in the interpretation of data from interactive treatments. Given different treatments, selections at nodes may represent nominal, ordinal or even integer data points, and each type of data imposes restrictions on how data can be analyzed. For example, one menu in a program offers the viewer a choice of "river memories," "river bridges," "river travel," or "river science." In this case "river memories" is represented as 1 and "river science" as 4. Because the data are nominal, the numerical representation is misleading; the conceptual distance implied between the two choices is in fact no greater than the conceptual distance between any two other choices from the menu. The problem is not isolated to numerical data. With tree diagrams and other graphic approaches, conceptual distance is implied as branches diverge on a diagram. In actual fact, however, choosing "A" at the seventh level of the farthest branch on the right side of a diagram is not necessarily different conceptually from choosing "A" at the fourth level of the left branch. Nevertheless, a casual observer can be seduced into thinking that spatial or numerical distance in data indicate conceptual distance as well.

Purposes of Audit Trails

To date, we have identified three distinct purposes to which audit trails can be put: as data-collection devices for formative evaluation in instructional design; as tools for basic research into the instructional design of CBI and hypermedia; and as a means of auditing usage of mediated presentations in a public forum.

Formative Evaluation in Instructional Design

The primary purpose for which audit trails have historically been used is for formative evaluation of instructional materials. It is useful to be able to determine which paths are perceived as attractive or significant by learners, and to learn where and how they make errors. If particular paths or segments of instruction receive less traffic than others, this may indicate a need for revision. In some cases, no traffic along particular paths may allow the instructional designer to eliminate those options from the system, perhaps improving the efficiency of the program or liberating space for other options.

The purpose of formative evaluation is always to optimize the performance of the product. In linear structures, learner variables are reduced in importance—they cannot be accommodated by the design. There is, of course, the option of developing parallel linear treatments for different subgroups of learners, but this is not usually a reasonable option for developers.

In branching structures, the possibility exists for either learners or designers (or both) to factor learner variables into the path decisions. The number of learner options can be increased. Efficiency is still still a goal, of course, but a greater possibility now exists for changes to be made to the instruction than in the case of linear structures.

Formative evaluation becomes less significant when used in the context of multimedia/hypermedia structures, however. Since designer influence is reduced, and learner influence is increased, efficiency ceases to exist as a construct against which to judge the performance of the materials: What does it mean to attempt to optimize the learner's path through Nelsonian hypermedia, where the learner is the only arbiter of the "correct" path? In multimedia/hypermedia, the purpose is not to optimize the treatment, but rather to open up the number of possibilities available to the learner. Formative evaluation concerns will likely be limited to cosmetic issues such as "ease of navigation" and "meaningful transitions" among elements of instruction learners encounter.

Formative evaluation is essentially a tool for instructional designers, and as structures are used which minimize the influence of instructional designers, the value of audit trails for formative evaluation purposes becomes increasingly unclear.

Basic Research in Instructional Design

Audit trails can also be used to track learner performance in research settings. Individuals, and indeed groups of individuals, can approach instruction differently, and this has traditionally been of theoretical interest. For example, consider individual differences or cognitive style constructs, such as locus of control. How might internalizers and externalizers differ in their approaches to highly organized interactive treatments? Would they react differently to, and take different paths through, very linear treatments and hypermedia treatments? One learner may select the shortest path available; another may select every available remedial segment in the same treatment. Resultant paths would be very different from one another, but, short of actually watching both individuals progress through the materials, how can these differences be expressed? As interest grows in the effectiveness of learner control of instruction (e.g., see Higginbotham-Wheat, 1990; López, 1990; Ross, Morrison, & O'Dell, 1990; Steinberg, 1977), and especially as it broadens into learner control of interactive and hypermediated instruction, these kinds of questions will command increasing interest.

Research on linear structures, as on many branching structures, is necessarily quantitative in nature. As we progress into multimedia/hypermedia structures, however, the research mode takes on a decidedly qualitative bent. On the surface, this may sound like a curious statement, but consider the following. There are relatively few questions one can ask about use of linear media beyond achievement/efficiency, performance and interactions with designs and learner variables. Meaning is imposed on the designs studied, usually by the producer or designer. For example, we could examine the effectiveness of a particular cueing strategy on different types of learners in linear media. We can ask questions like "did the cueing strategies help one group more than another?" But the treatment is fixed, so we are largely restricted to quasiexperimental designs unless parallel treatments are developed for comparison and control.

Audit trails in hypermedia fit the increasingly popular paradigm of naturalistic observation in that they are collected unobtrusively, in a natural setting. The instructional developer, given this orientation, is charged with designing a rich context within which learning can occur. Rather than being concerned about the direction and substance each learner encounters in instruction, the instructional developer is more concerned about the landscape of the instruction—the contours, breadth and depth of the terrain, and the ease with which learners can manoeuvre through the materials. The learner is viewed as part of an instructional ecosystem, simultaneously shaping and being shaped by the instruction encountered.

In their purest forms, natural interactive media are not based on instructional preconceptions, objectives or hypotheses, nor are they constructed to conform to the characteristics or needs of defined groups of learners. Rather, the learner enters the instruction unburdened, and the patterns of learning are allowed to emerge from the paths learners (and eventually groups of learners) construct. This is a potentially exciting orientation to instruction, and seems to be at odds with several of the assumptions underlying systematic instructional development (defined objectives, congruence among elements of instruction, reliable evaluation). At the same time, the instructional developer is still developing a system; this system is perhaps somewhat more organic, but its construction will still follow guidelines, conventions and rules.

One of the fundamental philosophical differences between quantitative and qualitative modes of inquiry is that a quantitative orientation emphasizes the existence of an externally definable reality. The researcher is trying to understand or reveal that reality (the absolute rules by which it operates, and how rules can be generalized). A qualitative orientation presumes the existence of multiple realities which arise from associated contexts. Given that reality is malleable and dependent upon the context in which it is observed, the researcher searches for meaning within a particular context and does not attempt to generalize meaning. Thus, meaning arises from a context, rather than by imposing meaning on a context to see if the general rules work. Since there are a huge number of potential contexts (probably finite, but who will bother counting?) multiple meanings may percolate from multiple realities.

Usage Audits for Unstructured or Public Environments

Sometimes multimedia packages are produced for use by a rather vaguely-defined audience (e.g., all visitors to a tourist site). Producers of such packages can only speculate about what content or paths in them will be most of interest and in demand, and observation of viewer/users is the only means of validating the producers' initial estimates.

A third use of audit trails therefore is to determine which paths of existing interactive media packages are of most interest to certain classes of viewers/users. The use of audit trails forms an unobtrusive way of effectively peering over the shoulders of groups of users to determine how they are traversing the interactive media package. This approach is similar to the classic unobtrusive measure of determining the amount of wear on floor tiles in front of various museum displays.

From such usage audit data, decisions might be made about optimal layouts for future treatments on videodiscs. By appropriate clustering of information, seek times may be reduced. If additional (unused) information exists on the videodisc, usage data might indicate re-vamping of the presentation is in order.

Anomalies is usage data might lead to testable research hypotheses. For example, based on observations, one might be prompted to ask the question "Do older users of a program persist in one set of activities while younger users sample from a range of activities?" As another example, in a point-of-purchase application, are people likely to make more purchases from a "motion video" display of products or from a "static video" display of products in a video catalog?

Other potential inferences include what similar subject-matter might be well received by future viewers/users (essentially needs assessment or marketing research.) If there has been heavy use of a certain class of information, it might be a clue that additional related information might also be popular.

This use of audit trails combines some of the characteristics of both formative evaluation and basic research—something of a hybrid of the earlier discussion. Certainly the goal can be to improve the effectiveness of a treatment. For instance, a point-of-purchase display might emphasize a particular type of product and de-emphasize others because of the design of selection screens. Audit trail data may suggest a need to redesign some of the selection screens. On the other hand, basic questions may arise too. For example, audit trail data may suggest that younger users consistently follow the shortest path available through certain portions of an interactive video presentation, while older users tend to linger in the same zones of the presentation. While this may ultimately have implications for formative evaluation, the data may also give rise to interesting questions of why specific design elements interact with the age of users.

In most cases, audit trail data for unstructured or public environments will include relatively simple census data. The primary interest of most users is "who is using this, and for what purpose?" The audit trail data offer information about who was exposed to which portions of a treatment, tracking user preferences rather than user performance.

Quantitative Description of the Audit Trail

In our initial search for a meaningful way to represent the audit trail, we investigated and considered several formats. The list we generated, below, is not exhaustive, but merely a point of departure. Some of the representations appear to be more useful and durable than others; some we considered briefly and discarded for various reasons outlined below.

Raw Data Matrix

The most basic way of representing audit trail data is to simply record the responses of each learner (as a vector), one above the other (see Figure 3). Matrices like these have the advantage of being easily constructed and relatively easily interpreted for each individual. The interpretation, however, can only be relatively limited, and context-bound. For crude formative evaluation purposes, the data are useful. An instructional designer can see which choices are attracting individual learners, and speculate about design decisions. But there is a serious limitation with this type of approach: raw data, by definition, aren't summarized and therefore conclusions based on the group are difficult to derive.

Insert Figure 3 about here.

Nodal Frequencies and Proportions

Another approach we investigated was to present data associated with each decision point (node) in the treatment. Data can be presented in at least two forms, as raw data (Figure 4) or as proportions (Figure 5).

Insert Figure 4 about here.

Insert Figure 5 about here.

Raw nodal frequencies, like raw data matrices, are easy to create. They are perhaps easier to interpret, since the data are now summarized. Magnitudes of differences are obvious, and at any node, comparisons have high precision and are intuitively satisfying. At the same time, relationships across nodes, or among variables are difficult to interpret. For example, how should four choices of "A" at one node be compared with 103 choices of "A" at another node, if they occupy different locations in the treatment? Perhaps only eight individuals encountered the first node, whereas several hundred encountered the second. The looping problems, described earlier in the paper, surface here to cause difficulties in interpretation. If an individual loops through a node several times, frequency data become distorted. Either the same choice is made several times, thereby inflating the frequencies at that node, or several different choices are made, thereby levelling the data at that decision point.

Nodal data can also be presented proportionally. Again, this type of format is easy to create and precision is retained at a high level. Proportions allow easier comparisons across nodes or among variables at different positions in the treatment. Of course some calculation is necessary, and the user must struggle with the question of what to use as a denominator. For example, is the denominator consistently the total number of learners encountering the treatment, or is the denominator the total number of learners who pass a particular decision point? Perhaps obviously, the denominator of any proportion will be determined by the comparisons the user chooses to make—yet another type of context dependence. As with raw nodal data, proportional data are also sensitive to looping problems. An individual looping several times through a particular node can inflate its proportion of the total. In addition, as one descends deeper into he data matrix, smaller raw numbers represent elevated proportions (see Figure 5). While proportional representation is useful for compressing large numbers, statistically bloating small numbers appears unnecessary and counterproductive.

When considering nodal data (either frequencies or proportions), individual differences are lost in the compression of data. Any design decisions based on these data are limited to conclusions about the group as a homogeneous entity, and we sacrifice any more subtle interpretations. Furthermore, data spread across several tables, each representing a single node (or, alternatively one large table showing multiple cross-breaks), are difficult to assimilate. Patterns that exist within them are difficult to detect.

Petit-Point Pattern

This early approach to the portrayal of the audit trail was suggested by some of John Tukey's work in representing non-parametric data in what he called a "stem and leaf" form (Hartwig & Dearing, 1979). It is a graphical approach, combining the intuitive appeal of a histogram with a character-based notational system symbolizing the choices. For example, in Figure 6, the X's represent the first choice, O's represent the second choice, and H's represent the third choice. A dot (•) is used to indicate that the learner proceeded past the node without making a choice, a situation made possible by an unfortunate bug in the program used to collect these sample data.

Insert Figure 6 about here.

Although virtually any symbol could be used to represent individual choices, but we chose characters that seem to occupy the same amount of space, so that inter-character differences would not influence the overall appearance, and perhaps the interpretation, of the display. The resulting pattern of characters is in some ways reminiscent of a pattern for petit-point embroidery.

To set up a display of this kind, data must be progressively or sequentially sorted decision point by decision point. That is, subsequent columns of data in a matrix must be sorted within the categories formed by the sorted data comprising the first column.

Advantages of this method included the ease of generating the display by using a search and replace function on a word processor to substitute characters, and a line sorting feature to assist in constructing the display. Although the display accurately shows the proportions of choices made at each node in a graphic and intuitive way, the dependency problem is very evidently in play: In the fifth column, for example, there are 11 distinct groups of O's; each group has a different meaning, depending on where it is located vertically. That is, although the O's all indicate that the second of the choices available was the one that was chosen, the first two O's and the third O represent choices on different content, due to the fact that different routes brought the learners to he node.

In one version of this kind of display, we also tried to use color to denote different choices, but found that it offered little advantage in interpretability.

This approach was eventually set aside because it didn't seem to do a great deal to describe what was happening. It was deemed to be moderately useful as a formative evaluation tool, but didn't appear to have sufficient power to make it useful for research purposes.

Audit Trail Tree

The audit trail tree (Figure 7) was the next approach we attempted. It combined both the graphical representation and the numerical accuracy of the Petit-Point Pattern approach, but, in addition, presented the data in a more intuitively powerful way.

Insert Figure 7 about here.

The audit trail tree is drawn so that the thickness of the line depicts the number of learners who chose the path represented. Of course, if large numbers of learners are involved, the line width could be scaled. Too, numbers (either frequencies or proportions) could be attached to each node to provide greater detail. The visual representation appeared to be useful and somewhat easier to interpret than the Petit-Point Pattern method. It was clear, it was graphical, it was intuitive, and it was grounded in reality. Comparisons were easy to make; flow could be read into the diagram as learners progressed from the beginning to the end of instruction. Although the drawing process is not difficult to do manually, it is somewhat tedious, and automation of the process on a graphic-interface microcomputer should be reasonably straightforward.

On the other hand, unless numbers were attached (as suggested above), the precision of the display was fairly low (i.e., it is difficult sometimes to tell the difference between 3 learners and 4, or between 11 and 13). The problem of conceptual distance remained. and perhaps was magnified by the ease with which other dimensions of the data were made manifest. And, of course, the problem of how to represent the loops remained.

Qualitative Description of the Audit Trail

The descriptive approaches described above are all quantitative. Descriptive analysis may also use qualitative methods productively. The goal of these descriptive approaches would be to examine the audit trail information within its context and not generalize beyond the context examined.

Content or Document Analysis

In one sense, an instructional product can be treated as a document, and a thorough examination of that document and related documentary sources can provide useful data for analysis. The usual purpose of document analysis is to explain the "status of some phenomenon at a particular time or its development over a period of time" (Best & Kahn, 1989). Any documents related to the development of an instructional product could be used in this type of analysis, and might include such things as client/developer contracts, outlines, storyboards, usage data, production schedules, other related instructional products, drafts of material, formative evaluations and the like.

One of the key cautions when using documentary sources as data is their trustworthiness. Documents , often as easily as individuals, may be inaccurate or even lack authenticity. For example, a formal contract between an instructional developer and a client might outline the parameters of a project in great detail, whereas a more accurate description of the intentions of the instructional designer might be scribbled on a cocktail napkin. Because the contract has legal and political implications, it may include cautious language which clouds the actual intentions of the contractual parties. The researcher must be careful to establish the trustworthiness of all data examined.

Some examples of audit trail studies which could employ this approach might include: to analyze the instructional design preferences of a producer of interactive media; to evaluate prejudice or bias in instructional products; or to reveal the underlying structures or levels of difficulty inherent in a particular product.

Case Study

Often used to longitudinally study development of social phenomena and change, case study methods can probably be extended to instructional products or processes. The "case" under study would not be a specific product or process, but rather an exemplar of, or prototype for, a category of products or processes. For example, a researcher conducting a case study would not be interested in a CD ROM treatment on architectural design as an entity. Rather, in a case study, the researcher might be interested in the development process and instructional design employed as an example of other products which might fall into a similar category.

Data are typically gathered from a number of sources, including direct observation, interviews, formal instruments and inventories or recorded data. The emphasis in conducting a case study is depth of analysis, not broad generalization. Generalizations would only be drawn from a series of cases which reveal consistent observations.

A number research questions in interactive media might employ such an approach. For example, a researcher interested in how people navigate through an unstructured program might choose one such program as an exemplar, and examine it in detail. Interviews with the instructional designer might reveal some of the options, limitations and assumptions inherent in the design of this particular product. Individuals using the product could be observed to see which of the navigational options are used, and which are not. Users might then be interviewed about their decisions, perhaps revealing some design reasons why certain paths were chosen or ignored. Of course, some of the quantitative approaches to describing audit trails might be useful to guide interviews, as might other formal inventories and tests. A thorough examination of this "case" from the multiple perspectives above, could reveal a great deal about the navigational preferences of individuals. Of course, the results would only be generalizable within the bounds of the characteristics of this treatment as an exemplar.

Inferential Approach to the Audit Trail

Inferential approaches would be used when comparisons are being made between groups of learners (e.g., differing on cognitive styles) or treatments (e.g., using instructors of the same gender as the learners and using instructors of differing genders).

One inferential approach to the audit trail we considered and discarded involves the use of multiple regression. This traditional approach to path analysis did not appear to be appropriate for answering the types of questions we are addressing here. It is a method of analyzing linear relationships among sets of variables, and assumes that a *causal order* among the variables is known and that the relationships among the variables are *causally closed* (Duncan, 1966). Even though we didn't give this approach much consideration, we mention it because of a possible confusion of terms: We are attempting to analyze paths through instruction, in a way that bears no relationship to path analysis, as the term is used in a statistical sense.

Furthermore, in an inferential approach to analyzing choices made in hypermediated and interactive instruction, one cannot make the assumption of normal distribution that underlies parametric statistics; hence a focus on the non-parametrics is essential.

A productive approach to analyzing a class of problems such as that under discussion would be to collect data on choices made by a large group of people, and regard that distribution as the *usual distribution* (in fact, it would be an expected distribution that is "normal" for the particular content and treatment being investigated, but since the term normal distribution has a technical connotation, we must make a distinction). Given this expected distribution, one could then subject certain individuals to treatments of varying kinds, and compare the audit trails of those subjects to the audit trails generated by the "usual" population, on a decision by decision basis (and keeping in mind the dependency problem). That is, the comparisons could only be made for single decision nodes at a time (which could be a limitation).

A statistic such as the χ^2 one-sample test, a test of goodness of fit (Siegel, 1956), would appear to be an appropriate tool for determining the statistical significance of observed deviations from "usuality". Another likely candidate would be the Kolmogorov-Smirnov test (Marascuilo & McSweeney, 1977; Siegel, 1956). Indeed, since Siegel states that "...the Kolmogorov-Smirnov test may in all cases be more powerful than...the χ^2 test" (1956, p. 51), it would seem to be the test of choice.

Advantages and Limitations of the Audit Trail Approach

Sophisticated authoring languages and systems permit an instructional developer to collect a wide range of data very easily. For example "Authorware Professional" for the Macintosh has more than 100 resident system variables and functions which can be inserted into an instructional program. It requires little sophistication to use, and resultant data can be written to files for later analysis. This is a boon for instructional developers who have specific parameters of the instruction they wish to analyze. For example, if efficiency of instruction is of interest, it is a simple matter to track the amount of time individuals devote to various instructional components, measure achievement, and from these two variables calculate an efficiency index (achievement/time). Where the instruction is stable, and relatively predictable, this offers significant opportunities to the instructional developer and researcher alike.

Of course, collecting massive amounts of data is one thing; making sense of the data is quite another. When one is conducting traditional empirical research, questions and hypotheses determine which data are important. Similarly, focussed formative evaluations which externally define the parameters for criticism and potential revision will impose limitations on the data collected. However, when the instruction is approached assumption-free, how are data excluded as insignificant? Systematically excluding data may create an impoverished description of interaction, or perhaps worse, may introduce bias into the system under investigation. It appears that the glut of data is necessary, but how can it be tamed?

In addition, unbridled data collection can result in a theoretical wasteland for individuals who are curious or eclectic (read snoopy, sloppy and unable to discard anything). Data snooping can be raised to absurd heights, and we have visions of researchers and developers paralyzed beneath mounds of imponderable data.

"The richness of nonlinear representation carries a risk of potential intellectual indigestion, a loss of goal directedness and cognitive entropy. The availability of multiple types of representations in a hypermedia system presents a cognitive overload about which little is known." (Dede, 1990, p. 20)

It is important to realize that excessive attention to detail can be dysfunctional. Often the most useful data are those which provide generalizations and trends. Extremely fine-grained information may seem like a good idea, but the most useful observations may yet be drawn from seemingly crude, but well-derived, data. While trying several of the analytical tools described in this paper, we were impressed with how quickly data grew beyond our capability to represent it. Clearly, the more complex instructional designs and questions about data become, the more we are driven into underdeveloped strategies for analysis.

Analyzing complex audit trail data meaningfully involves making inferences about two things: whether a chosen path has integrity; and why a particular path was chosen from among alternatives. First of all, we look for integrity. When an individual or group of individuals follow a particular path, we must decide whether the series of choices comprising the audit trail has any external meaning. It is possible that a series of choices may be relatively random. Therefore, in an attempt to ascribe meaning to emerging audit trails, it is possible to draw spurious conclusions. Independent observations by researchers or developers should provide a measure of reliability to conclusions.

Determining why a particular path was chosen from among alternatives is hazardous work, yet at the heart of conducting formative evaluation or basic research. Certainly post-hoc interviews and independent observation by experts can provide a measure of enlightenment; simply asking, "why did you make these choices?" can confirm suspicions. Nevertheless, individuals may choose similar paths for quite different reasons, suggesting a lack of coherence to paths, when two or more coherent cognitive paths may actually occupy the same physical geography. Of course, there is the danger that individuals confronted with the question will work to impose meaning on their own set of responses; self analysis is as natural as it is often flawed. Another potential pitfall when interviewing individuals, particularly learners, is that they have little or no understanding of the paths *not* taken. A choice of direction allows an audit trail to emerge for one group, but does not permit extensive comparison with competing audit trails for the same group. For example, an individual might say "I started here because the instruction included my name in the question, and then always chose the first option in a list after that" but can not comment on various paths which streamed away from the initial options using personal pronouns or impersonal references. When searching for meaning in audit trails, we are engaged in conceptual exploration, and it appears to us that this requires collaboration (almost conspiracy) among designers, SMEs and learners to be successful.

Issues and Challenges

One of the most pressing needs we see in audit trails research is the development of procedures for collecting and analyzing data. It is quite possible that existing procedures can be adapted to inform this area of study, but as new knowledge structures emerge, we predict that analytical tools will continue to fail to keep pace with instructional designs. Our ability to make sense of increasingly sophisticated webs of instruction depends on the development of extremely robust analytical tools and strategies. Other areas of study, such as geography or oceanography, may provide insight into ways to chart and analyze complex (and yet seemingly graphic) forms of data.

At the same time, sophisticated analytical approaches will open the door to many questions heretofore untouched. Recent work in knowledge-based management systems (KBMS) suggests research and design inquiry into multiple knowledge representation schemes and applications, improving and studying the effects of context-free information and context-sensitive advice, and the effects of immersing learners in "virtual reality" information environments (after Dede, 1990). Research in information saturated environments will allow researchers to revisit questions of learning styles and learner motivation from a fresh perspective. In a very real sense, qualitative studies will permit new questions to arise from the experiences of users within increasingly rich instructional environments. We can not afford to be arrogant at this stage of development, and suggest the range of research questions which may be important in multi-media environments. Our ability to enrich and structure information, and our continued progress in creating natural interfaces between knowledge structures and learners, far outstrip our current ability to analyze data, or even ask intelligent questions of data. We are facing an opportunity for real and meaningful exploration, and we should approach the task unfettered by some of our most fondly held assumptions (e.g.,. we can systematically manipulate instruction to adequately address the needs of learners). It is a time for listening, not necessarily answering.

References

Best, J. W., & Kahn, J. W. (1989). *Research in education* (6th ed.). Englewood Cliffs, NJ: Prentice-Hall.

Dede, C. (1990). Knowledge processing: Navigating the sea of information. *Authorware*, *3*(1), 17-22.

Duncan, O.D. (1966). Path analysis: Sociological examples. *American Journal of Sociology*, 72, 1-16.

Hartwig, F., & Dearing, B. E. (1979). *Exploratory data analysis*. Sage University Paper series on Quantitative Applications in the Social Sciences, series no. 07-016. Beverly Hills, CA: Sage Publications.

Higginbotham-Wheat, N. (1990, February). *Learner control: When does it work?* Paper presented at the Annual Meeting of the Association of Educational Communications and Technology, Anaheim, CA.

López, C. L. (1990, February). *Personalizing math word problems*. Paper presented at the Annual Meeting of the Association of Educational Communications and Technology, Anaheim, CA.

Marascuilo, L. A., & McSweeney, M. (1977). *Nonparametric and distribution-free methods for the social sciences*. Monterey, CA: Brooks/Cole Publishing.

Rezabeck, R. H., & Ragan, T. J. (1989, February). *Elaborated resources: An instructional design strategy for hypermedia*. Paper presented at the Annual Meeting of the Association of Educational Communications and Technology, Dallas, TX.

Ross, S., Morrison, G., & O'Dell, J. (1990, February). *Uses and effects of learner control of context and instructional support in computer-based instruction*. Paper presented at the Annual Meeting of the Association of Educational Communications and Technology, Anaheim, CA.

Ryan, T. A. (1972). Multiple comparisons in psychological research. In R. E. Kirk (Ed.), *Statistical issues: A reader for the behavioral sciences* (pp. 291–306). Monterey, CA: Brooks/Cole Publishing.

Schwier, R. A. (1987). *Interactive video*. Englewood Cliffs, NJ: Educational Technology Publications.

Siegel, S. (1956). *Nonparametric statistics for the behavioral sciences*. New York: McGraw-Hill.

Steinberg, E. R. (1977). Review of student control in computer-assisted instruction. *Journal of Computer-Based Instruction*, *3*(3), 84–90.

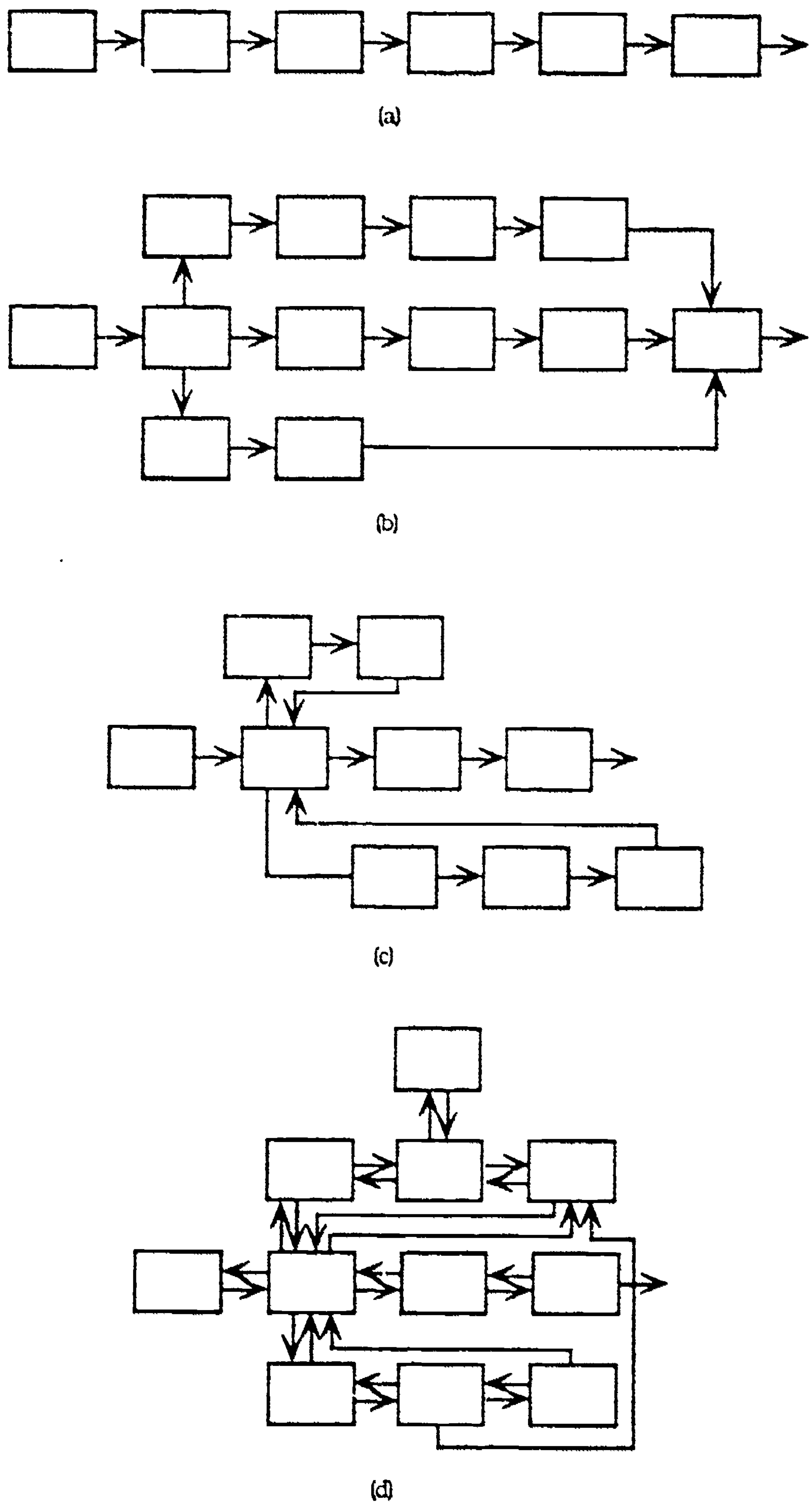

Figure 1. Program structures: (a) linear; (b) learner controlled parallel path branching; (c) classic feedback loop branching; (d) multimedia/hypermedia.

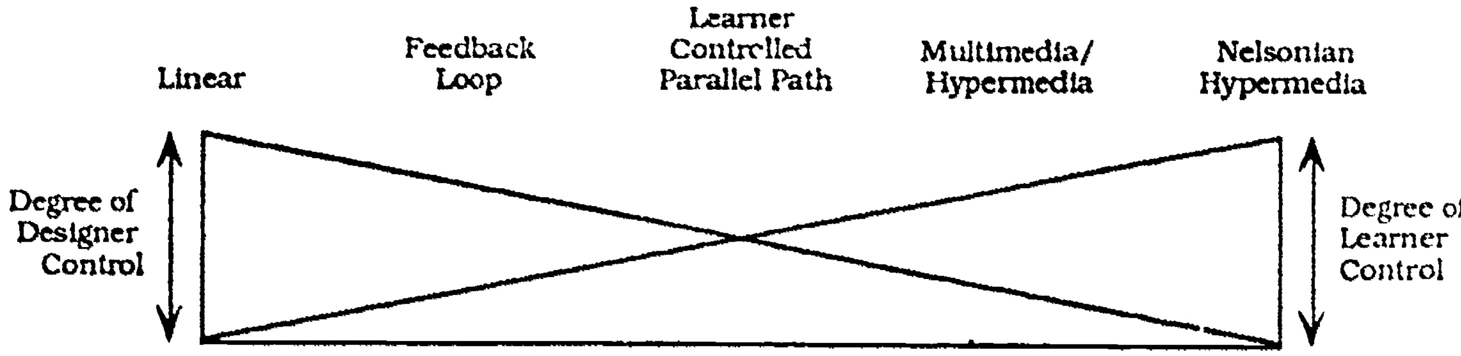

Figure 2. Degree of designer and learner control in various structures.

Subject 1	2	2	2	2	1	1	1	1
Subject 2	2	2	2	2	1	1	1	1
Subject 3	2	2	2	2	1	1	2	1
Subject 4	2	2	2	2	1	1	2	1
Subject 5	2	1	2	2	1	1	1	1
Subject 6	1	3	2	2	2	1	1	1
Subject 7	2	2	2	2	2	1	1	1
Subject 8	2	1	2	2	2	1	1	1
Subject 9	2	3	2	2	1	1	2	2
Subject 10	2	1	2	2	1	1	2	2
Subject 11	2	3	2	2	1	1		
Subject 12	2	3	4	4	1	1		
Subject 13	2	3	4	4	1	1		
Subject 14	2	3	4	4	1	1		
Subject 15	2	2	2	2	1	1	1	
Subject 16	2	2	4	2	1	1		
Subject 17	1	1	2	2	1	1		
Subject 18	2	1	2	2	1	1		
Subject 19	2	1	2	2	1	1		
Subject 20	2	1	2	4	1	1		
Subject 21	1	3	4	1	2	1	2	
Subject 22	1	3	4	1	2	1	2	
Subject 23	1	3	4	4	2	1		
Subject 24	2	3	4	2	2	1	2	
Subject 25	2	2	2	2	2	1	1	
Subject 26	2	2	4	2	2	1	2	
Subject 27	2	1	4	2	2	1	2	
Subject 28	1	3	2	2	1	2	1	1
Subject 29	2	3	2	2	1	2	1	1
Subject 30	2	2	2	2	1	2	1	1
Subject 31	2	2	2	2	1	2	1	1
Subject 32	2	2	2	2	1	2	1	1
Subject 33	2	1	2	2	1	2	1	1
Subject 34	2	3	2	2	2	2	1	1
Subject 35	2	3	2	2	2	2	1	1
Subject 36	2	3	2	2	2	2	1	1
Subject 37	2	2	2	2	2	2	1	1
Subject 38	2	2	2	2	2	2	1	1
Subject 39	2	2	2	2	2	2	2	1
Subject 40	2	1	2	2	2	2	1	1
Subject 41	2	1	2	2	2	2	2	1
Subject 42	1	3	2	2	1	2	2	2
Subject 43	2	3	2	2	1	2	1	2
Subject 44	2	3	2	2	1	2	1	2
Subject 45	2	3	2	2	1	2	2	2
Subject 46	2	3	2	2	1	2	2	2
Subject 47	2	1	2	2	1	2	2	2
Subject 48	2	3	2	2	2	2	1	2
Subject 49	2	3	2	2	2	2	2	2
Subject 50	2	3	2	2	2	2	2	2
Subject 51	2	2	2	2	2	2	1	2
Subject 52	2	2	2	2	2	2	1	2
Subject 53	2	2	2	2	2	2	1	2
Subject 54	2	2	2	2	2	2	2	2
Subject 55	1	1	2	2	2	2	2	2
Subject 56	2	1	2	2	2	2	1	2
Subject 57	2	1	2	2	2	2	2	2
Subject 58	2	3	2	2	1	2		
Subject 59	2	3	4	2	1	2		
Subject 60	2	3	4	4	1	2		
Subject 61	2	3	4	4	1	2		
Subject 62	1	2	4	2	1	2		
Subject 63	2	2	2	2	1	2		
Subject 64	2	2	2	2	1	2		
Subject 65	2	1	2	2	1	2		

Subject 66	2	1	2	2	1	2		
Subject 67	2	1	2	2	1	2		
Subject 68	1	3	4	1	2	2	2	
Subject 69	1	3	4	2	2	2	2	
Subject 70	1	3	4	4	2	2		
Subject 71	1	3	4	4	2	2		
Subject 72	1	3	4	4	2	2		
Subject 73	2	3	2	2	2	2		
Subject 74	2	3	2	2	2	2		
Subject 75	2	3	2	2	2	2		
Subject 76	2	3	4	4	2	2		
Subject 77	2	3	4	4	2	2		
Subject 78	2	3	4	4	2	2		
Subject 79	2	3	4	4	2	2		
Subject 80	2	3	4	4	2	2		
Subject 81	2	3	4	4	2	2		
Subject 82	2	3	4	4	2	2		
Subject 83	2	3	4	4	2	2		
Subject 84	2	3	4	4	2	2		
Subject 85	2	3	4	4	2	2		
Subject 86	2	3	4	4	2	2		
Subject 87	2	3	4	4	2	2		
Subject 88	2	3	4	4	2	2		
Subject 89	2	3	4	4	2	2		
Subject 90	2	2	4	1	2	2	2	
Subject 91	1	1	2	2	2	2		
Subject 92	2	1	2	2	2	2		
Subject 93	2	1	2	2	2	2		
Subject 94	2	1	2	2	2	2		
Subject 95	2	3	2	2	1	3	2	1
Subject 96	2	1	2	2	1	3	1	1
Subject 97	2	1	2	2	1	3	2	1
Subject 98	2	1	2	2	2	3	1	2
Subject 99	2	1	2	2	2	3	2	2
Subject 100	2	2	2	2	2	4		
Subject 101	1	3	4	4	1			
Subject 102	1	2	4	1	1			
Subject 103	1	2	4	1	1			
Subject 104	1	2	4	2	1			
Subject 105	2	3	4	4	2			
Subject 106	1	2	4	2	2			
Subject 107	1	4	4	4	4			
Subject 108	1	4	4	4	4			
Subject 109	1	4	4	4	4			
Subject 110	1	4	4	4	4			
Subject 111	1	4	4	4	4			
Subject 112	2	4	4	4	4			
Subject 113	2	4	4	4	4			
Subject 114	2	4	4	4	4			
Subject 115	2	4	4	4	4			
Subject 116	2	4	4	4	4			
Subject 117	2	3	4	4	4			
Subject 118	1	2	2	2	4			
Subject 119	1	2	4	1	4			
Subject 120	1	2	4	2	4			
Subject 121	2	2	4	2	4			
Subject 122	2	2	4	2	4			
Subject 123	1	1	2	2	4			
Subject 124	1	1	2	2	4			
Subject 125	1	1	2	2	4			
Subject 126	1	1	2	4	4			
Subject 127	2	1	2	2	4			
Subject 128	2	1	2	2	4			
Subject 129	2	1	2	4	4			
Subject 130	2	1	2	4	4			
Subject 131	2	1	2	4				

Figure 3. Sample raw data matrix.

Node	Resp.	1																							
1	Freq.	32																							
Node	Resp.	1						2						3						0					
2	Freq.	7						8						12						5					
Node	Resp.	2			0			2			0			2			0			2			0		
3	Freq.	7			0			1			7			3			8			0			5		
Node	Resp.	1	2	0	1	2	0	1	2	0	1	2	0	1	2	0	1	2	0	1	2	0	1	2	0
4	Freq.	0	6	1	0	0	0	0	1	0	3	4	0	0	3	0	3	1	5	0	0	0	0	0	5
Node	Resp.	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
5	Freq.	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.

Node	Resp.	2																							
1	Freq.	99																							
Node	Resp.	1						2						3						0					
2	Freq.	28						25						41						5					
Node	Resp.	2			0			2			0			2			0			2			0		
3	Freq.	27			1			20			5			18			23			0			5		
Node	Resp.	1	2	0	1	2	0	1	2	0	1	2	0	1	2	0	1	2	0	1	2	0	1	2	0
4	Freq.	0	23	4	0	1	0	0	20	0	1	4	0	0	18	0	0	2	21	0	0	0	0	0	5
Node	Resp.	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
5	Freq.	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.

Figure 4 Tabular representation of raw frequencies of paths chosen in sample data. Although the sample data actually have eight decision nodes, only four are presented here, for simplicity. Note how quickly the size and complexity of the table grows as the number of decision nodes and/or the number of choices available at each node increases.

Node																									
Node 1	Resp.	1																							
	Prop.	24																							
Node 2	Resp.	1						2						3						0					
	Prop.	22						25						38						15					
Node 3	Resp.	2			0			2			0			2			0			2			0		
	Prop.	100			0			13			87			25			75			0			100		
Node 4	Resp.	1	2	0	1	2	0	1	2	0	1	2	0	1	2	0	1	2	0	1	2	0	1	2	0
	Prop.	0	86	14	0	0	0	0	100	0	42	58	0	0	100	0	33	11	56	0	0	0	0	0	100
Node 5	Resp.																								
	Prop.																								
.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.

Node																									
Node 1	Resp.	2																							
	Prop.	78																							
Node 2	Resp.	1						2						3						0					
	Prop.	28						25						42						5					
Node 3	Resp.	2			0			2			0			2			0			2			0		
	Prop.	96			4			80			20			44			56			0			100		
Node 4	Resp.	1	2	0	1	2	0	1	2	0	1	2	0	1	2	0	1	2	0	1	2	0	1	2	0
	Prop.	0	85	15	0	100	0	0	100	0	20	80	0	0	100	0	0	9	81	0	0	0	0	0	100
Node 5	Resp.																								
	Prop.																								
.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.

518

Figure 5. Tabular representation of proportions of paths chosen in sample data for first four decision nodes. Note how, as one progresses through more decision nodes, smaller frequencies of responses produce deceptively larger proportions.

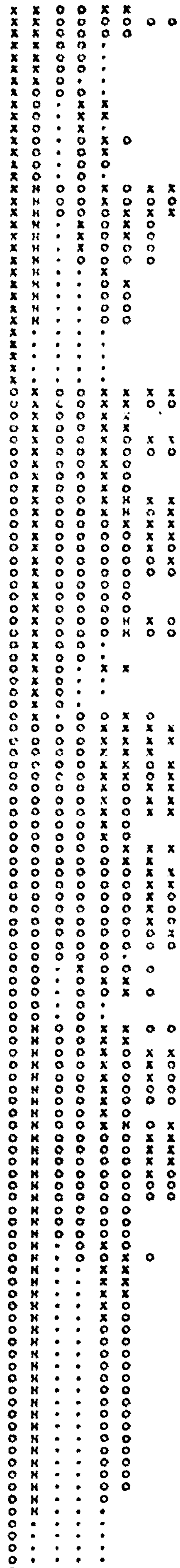

Figure 6. Petit-Point Pattern method of representing an audit trail. An X represents the choice of the first possible path from a node, an O represents the second, and an H represents the third. A • indicates no response.

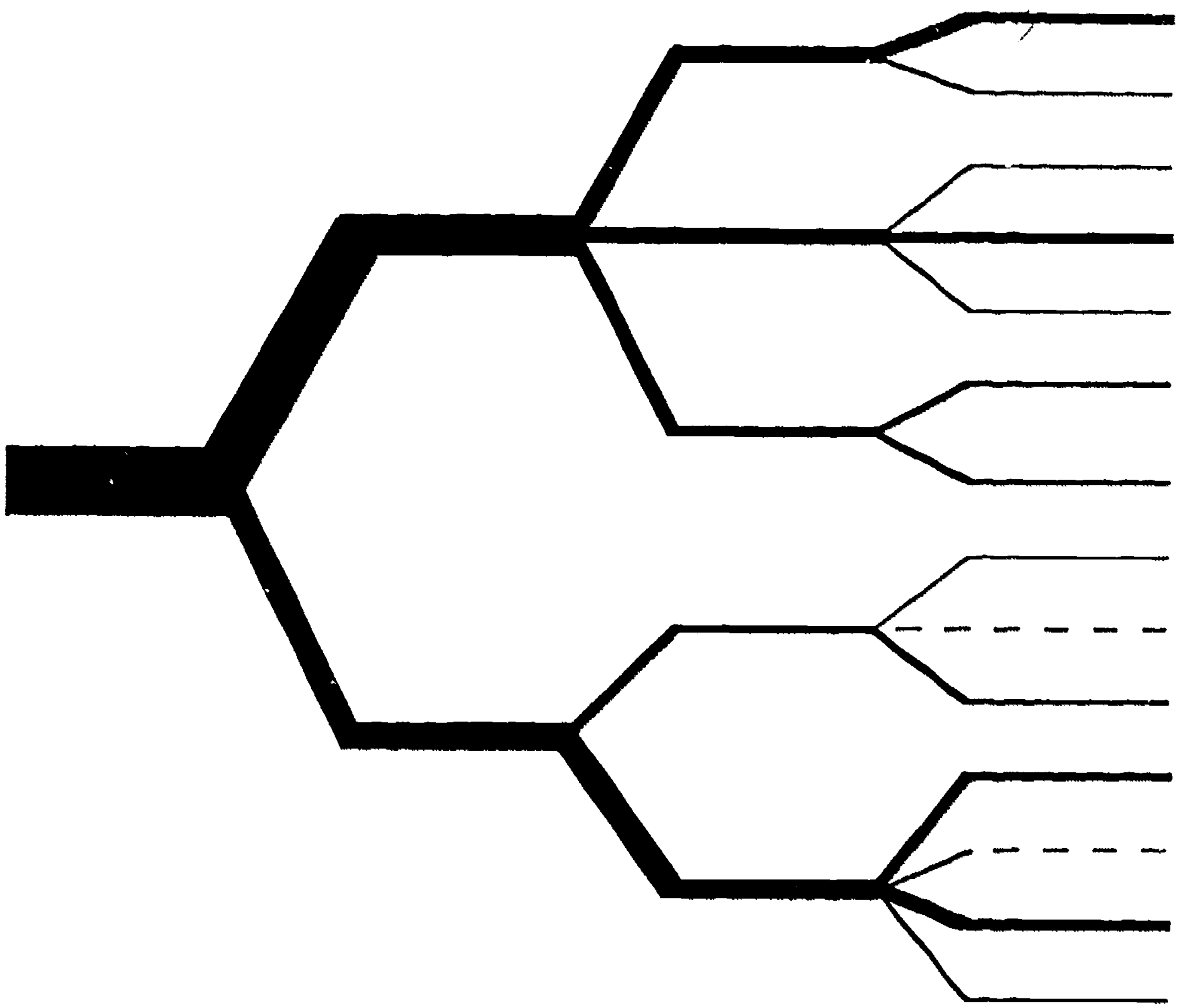

Figure 7. Audit Trail Tree of example data. The width of the line represents the number of learners taking any given path. A dashed line indicates no learners took the path.