DOCUMENT RESUME

ED 108 598 IR 002 109

AUTHOR Follettie, Joseph F.
TITLE Experimental and Student-System Interactive Instructional Illustrations Pertinent to IDCMS.
INSTITUTION Southwest Regional Laboratory for Educational Research and Development, Los Alamitos, Calif.
REPORT NO SWRL-TN-1-72-01
PUB DATE 4 Feb 72
NOTE 23p.

EDRS PRICE MF-$0.76 HC-$1.58 PLUS POSTAGE
DESCRIPTORS Audiovisual Aids; Computer Programs; *Computer Science; Computer Storage Devices; Educational Development; *Educational Experiments; *Information Systems; *Instructional Technology; Interaction; Learning Processes; Program Design; Skill Analysis; *Systems Analysis; Systems Approach
IDENTIFIERS IDCMS

ABSTRACT

Southwest Regional Laboratory IDCMS (Instructional Development Control and Monitoring System) will be a flexible hardware system for controlling and monitoring instruction and research in the laboratory setting. This paper seeks to introduce potential users to the system and software designers to representative challenges that system exploitation will pose. (Author/SK)

SWRL

SOUTHWEST REGIONAL LABORATORY
TECHNICAL NOTE

DATE: February 4, 1972

NO: TN 1-7 -01

TITLE: EXPERIMENTAL AND STUDENT-SYSTEM INTERACTIVE INSTRUCTIONAL ILLUSTRATIONS PERTINENT TO IDCMS

AUTHOR: [illegible] Hotte

ABSTRACT

SWRL IDCMS will be a flexible hardware system for controlling and monitoring instruction and research in the Laboratory setting. This paper seeks to introduce potential users to the system and software development and to represent the challenges that system exploitation will [illegible].

SUPPORTING [illegible] [illegible] -- [illegible] INSTRUCTIONAL ILLUSTRATIONS PERTAINING TO IDCMS

Joseph [illegible]. Follettie

The SWRL Instructional Development Control and Monitoring System (IDCMS) is designed to provide broad support for the Laboratory mission. That is, it will be responsive to a variety of specific organizational requirements. This paper will deal only with the system's capability for supporting requirements for 1) conducting educational research and 2) for trying out instruction that stems from such research. Comments on instruction will be further restricted to those sorts of instruction that feature nontrivial levels of student-system interaction. While certain details of the illustrations to be offered may appreciably influence system software designers or even lie beyond capabilities of the system in Version 1 configuration, the intent is only to promote a better understanding concerning how the interactive domain will be bounded during earliest applications of the system. The illustrations are offered as conjecturally consonant with system capabilities as these can be read from preliminary documentation. If facets of the illustrations go beyond what the system in Version 1 configuration will be able to do, then it will be useful to find out in what respects this is so and whether such features can be obtained quickly and cheaply consonant with earliest usage of the system--that is, usage during the first semester of the 1972-73 school year.

Although it is inevitable that the system will not be able to do certain things at the outset that users would like it to do--or, more to the point, not do them with sufficient ease--preliminary documentation suggests that system capability for supporting interactive mode educational research and tryouts of instruction will fall somewhere between one and two orders of magnitude beyond the pre-system Laboratory capability for supporting such efforts. General understanding of system support capability among instructional design-development staffs should be followed almost immediately by a dramatic rise in productivity of the efforts of these staffs. However, the system will be little exploited until such understanding prevails.

This paper seeks to stimulate both prospective users and software designers, for it is necessary that the two groups reach a common understanding concerning support capabilities that it will prove reasonable to expect that the system in Version 1 configuration can supply. Student-system interactive instruction was selected as a first query domain because it is one that will tax any system, will encompass much work that has been [illegible] for years, and which will produce the greatest [illegible] [illegible] [illegible] [illegible] [illegible] capability.

[illegible]

[illegible] controller [illegible] NOVA 1200 device having, in Version 1 configuration, a core storage of 16K (16,384) 16-bit words that can [illegible] in different proportions to Read Only (program) and Read-Write (e.g., analysis functions) memories. Should it prove necessary, the controller can be core-to-core interfaced with larger computer [illegible]. (Although such interfacing currently is viewed as [illegible]-line, it is [illegible] that on-line coupling could be effected if [illegible] level of on-line calculations underlying conditional advance [illegible]. [illegible] on-line analysis might apply sequential analysis [illegible] analysis of a sequence of responses to test [illegible], with test terminating after n items are administered or [illegible] upper or lower decision boundaries are pierced, whichever [illegible] first. Should it be required that such analyses occur under a [illegible] each of 64 Ss responds once every 6 seconds--[illegible] and [illegible] computation per second on the average--then it appears that further computer [illegible] would have to come into the system on an [illegible].)

[illegible] characterization comments address users who may [illegible] little [illegible] awareness concerning hardware assemblies of the system. [illegible] place to begin is with the system's analog-form video disc storage--approximately 1800 frames ([illegible] are lost to system control). This store can [illegible] randomly accessed at acceptably-high speed (for most purposes). The [illegible] video buffer for still display of any [illegible] require the same frame at almost [illegible] time, [illegible] to wait a short while (typically no more [illegible] additional frame accessing-copying time, itself [illegible] frames are off-line loaded using [illegible] part of the IDCMS contract).

[illegible] audio loader, extensive audio [illegible] ([illegible] 12 tape transports, in Version 1 configuration) and audio [illegible]-specified audio programs from the audio store. An [illegible] feature of audio tape storage of programs is that intrin[illegible] controlling selection-synchronization of video frames [illegible] on the audio tape using a pulse-modulated 55 Hz [illegible]. A [illegible] off-line coding of tapes. Master [illegible] and another for [illegible], [illegible] of masters. Code and [illegible] audio storage and [illegible].

[illegible]--a digital [illegible] for voice responses. [illegible] by the NOVA [illegible] capacity [illegible] not all of this capacit[illegible] digital tape [illegible]

[illegible] will [illegible] fully [illegible] specific [illegible] part of the system. This apparently [illegible] behind [illegible] the state of the art.

Self-referenced software is loaded into the system through a teletypewriter which encodes instructions via a paper tape punch device. Potentially, a dynamic augmentation of audio-video presentation materials can be effected using external-source audio and video material which enters the system through patchboards. System outputs referenced to user responses, latencies, and the like are not yet been identified ([illegible], that is).

The version [illegible] system contemplates supporting instruction or research [illegible] complex [illegible] to 6 Ss simultaneously.[1] Each S is provided with a terminal whose principal features are a 12" Trinitron color CRT, a set of earphones, and a response keyboard. Other input-output devices can be readily "jacked" to the terminal. Depending on applicable characteristics of controlling software, the system can be caused to vest certain interactive options in S. (One that it appears to vest routinely in S--or in [illegible] standing beside S--is the item advance option. Desirably we will be able to vest such an option either in S or in the system and, if in the system, either conditional on S's prior performance or not.)

Finally, the system incorporates a character generator (also a part of the television System and procured through that contractor) that can be used to caption video displays. The system apparently also easily can be made to accommodate a line generator or whatever is needed to [illegible] performance plots--e.g., plots that contrast some facet of S's performance over time with criterion performance--to come up on command.

Interactive Instruction

It is possible to characterize interactive instruction and interactive instructional research within the same framework. That is, different groups can receive different versions of a given instructional program, with variations across versions tapping one or more factors of research interest. The illustration to be presented is at a level of [illegible] that is consonant with ready extrapolation to the research [illegible]. [illegible] view of this paper is that by the time we are ready

---

[1] The illustration to be presented will assume that up to 6 Ss time-[illegible] while participating in the same instructional program [illegible] experiment. The alternative--running up to 6 instructional programs [illegible] experiments simultaneously--entails assigning portions of core to [illegible] experiment. It is apparent that this can occur only if [illegible] instructional programs or experiments (or a mix of the two) [illegible] relatively reduced core.

[illegible] traits of interactive instruction, [illegible] will have laid the groundwork for that instruction in consequence of [illegible] interactive educational research. The instructional [illegible] of the illustration, therefore, is intended to be generated.

Gross Anatomy of Interactive Instruction. Although the contractor has characterized student-system interaction as the simple-minded sort that went out of vogue over a decade ago--i.e., as referenced to a one-response decision domain--IDCMS is potentially much more powerful than the contractor's view implies. To begin with, mainline versions of any instructional program can be viewed as linear progressions. We will deal here with just one mainline version of a program. It will have a start point, followed by a linear progression of lessons and decision points, with exit--e.g., to following program--through a terminal decision point. The decision point permits I (or the system) to evaluate a given skill using 2-20 items, to diagnose the cause of unacceptable performance (whether referenced to a current decision point or to a sequence of decision points culminating in the one now visited by S), and to prescribe instruction designed to raise unacceptable performance to an acceptable level.[2] If S responds as intended at each decision point, then he will negotiate the program along the instructional path traced by the mainline version of the program.

There are no set or acceptable conventions for labelling the alternatives to negotiating the program along the instructional path traced by its mainline version. We are concerned here with a requirement that [illegible] exit through a single terminal decision point to [illegible] following. Worth noting in passing, a program that truly [illegible] includes [illegible] through the same terminal decision point. True branching programs offer S a choice, at one or more decision points, concerning whether he will travel to a specified terminal decision point or to an alternate terminal decision point. A program that truly branches does not ever converge (at least in any finite [illegible]) its branches, once these are formed. Hence, reference to branching in an instructional context cannot signify true branching when the intent is to [illegible] all [illegible] through the same terminal decision point. The apparent forms of departure from a mainline linear progression [illegible] loops, skips, extended alternate paths, and [illegible].

[illegible] dimensional. The first--a simple [illegible] return to an earlier mainline lesson in [illegible] negative outcomes experienced at a decision

---

[2] [illegible] "[illegible]-Prescription [illegible]"

[illegible] reverts to the mainline instructional path. The [illegible] loop to the same lesson seems pedagogically unattractive. It [illegible] to represent the lesson in exactly the same form it was negotiated [illegible]. Apart from the unacceptable rote-memorization implications of [illegible], we are still with the fact that earlier instruction [illegible] ineffective for the particular S under consideration. Diagnosis [illegible] indicate that modifications are in order; prescription should indicate the address of suitably-modified instruction addressing the [illegible] skill as the failed mainline lesson. We return to such alternative instruction in a moment.

The second sort of loop might be called an *extra-program loop*. All instruction presumes certain prerequisite skills occur in S's skill-repertoire. These typically are called entry skills. S moves forward [illegible] a program start point only if the assumption appears tenable that [illegible] possesses the entry skills on which the program is predicated. While [illegible] assumptions can be evaluated to a degree, it is impossible to [illegible] and evaluate every sort of prerequisite skill that might apply, although informal evaluation of major approximations is possible. We [illegible] here that extra-program loops reference to the major perceived entry skills and that initial tenability of an assumption that S possesses such skills ought not close out on the possibility that later, [illegible]-mainline evaluation will show holes in his entry skills repertoire. [illegible] a hole is detected at a given decision point, it is possible that [illegible] that will be required is that S loop to supplemental instruction addressing the deficient entry skill, with exit from the loop being [illegible] to the decision point at which the loop originates.

[illegible] loops [illegible] to an earlier program skill but not to the same lesson [illegible] negotiated--but rather an alternate lesson, perhaps [illegible] elaborates on what was presented earlier--then the loop shares [illegible] characteristics of the recycling loop and of the extra-program loop. That is, [illegible] recycles to a program skill for which he received [illegible] instruction but he does not simply recycle to the same lesson [illegible] to an alternative lesson. We characterize this departure [illegible] loop departure.

[illegible] linear progression that--installed in the schools--[illegible] require no skips. However, experimental mainline [illegible] progressions might either accept skips as an undesirable interim [illegible] or as a phenomenon whose effects upon [illegible] are worthy of study in their own right. A *skip* occurs when [illegible] decision point indicates that S should skip one [illegible] lessons in the progression. Setting aside rest-review-[illegible], a skip decision could only be reached if testing at [illegible] reference both to skills negotiated earlier [illegible]. It appears tenable that a skip most often [illegible] of earlier lessons applying [illegible] departure that carries S back [illegible] skill [illegible]

[illegible] interpretation. If these program skills are independent in the sense that either is treatable independently of treating the other, then a mixed loop departure might be followed by a skip that returns to the decision point wherein unacceptable performance originally was detected. This would constitute a mixed loop plus skip form of departure from the mainline progression.

When a mixed loop retrogression is followed by progression along an alternate path to the mainline path until the original decision point departed from is reached or when S proceeds from such a decision point forward to an alternate path that closes with the mainline path at a terminal or intervening decision point, then the departure might be characterized as an extended alternate path departure. Particularly when such departures are in the forward direction, some are tempted to call them branches, even though it is clear in light of the earlier discussion that they do not represent true branches. Alternate paths might arise for a variety of reasons--e.g., differential demands across Ss for elaboration on instruction, as revealed by accuracy and rate of response on achievement or diagnostic tests. (A decision to assign S to such alternate path is not irrevocable. The decision might be reevaluated at each decision point along the progression, with reclassification a continuing option.)

The variations in interactive instruction referenced to a mainline [illegible] of such instruction take the form primarily of departures from the mainline linear progression of the following types: extra-program loops, nested loops, mixed loops with skips, simple skips forward, and extended alternate paths, whether or not following retrogression. The anatomy of interactive instruction--at the skeletal level if not the histological-cytological--is that which is afforded by these options for forming departures from mainline instruction conditional on performance. In the current era we do not imagine a system of the IDCMS type having artificial intelligence capabilities for perceiving useful departures from mainline instruction or for acting on such perception to form suitable departure materials. Another way to put this is that I, during the current era, will identify all entertainable forms of departure, prepare materials sets that can be used under given departure conditions, prepare all achievement, diagnostic, and other material sets that are of interest, and do whatever else is required before IDCMS can be asked to do its thing. This does not preclude the probe model made famous by operant conditioners, wherein several treatments in succession are tried [illegible], so long as it is understood that I must foresee all the alternative treatment sets he will study under such conditions [illegible] IDCMS.

As noted earlier, specification of the anatomy of an interaction [illegible] that can be used to tryout instructional programs can be [illegible] step in specification of interaction research. [illegible] instructional programs having all of the interact[illegible] [illegible] in an interactive instructional system, [illegible] the performance of different groups of Ss on [illegible]

[illegible] for [illegible] instruction [illegible] administered but [illegible] options, [illegible] characteristics [illegible] used to define within-condition [illegible], and the like are [illegible] across groups. We [illegible], then, that illustrative interactive instruction simply need to be carried [illegible] [illegible] to [illegible] illustrative interactive instructional research.

An Interaction Scenario for Spelling Instruction. Short instructional [illegible] preclude formulating interactive research that is much interactive. Conversely, extended instructional programs tend to imply interactive research whose hardware requirements cannot yet be met or can be met only at prohibitive cost. Instructional programs of intermediate length permit formulation of interactive research that is subroutine rich (and [illegible]ing) while making only reasonable demands on hardware in light of state of the art and available resources. Moreover, system-referenced [illegible] employment of an interactive system addressing instruction of intermediate length has the realistic effect of raising time-preserved [illegible]-retrieval burdens--a time-share facet of the system--above trivial levels, since a system that is responsive to student rate [illegible] positions must in time cause students to spread out over an [illegible] stretch of the program. For these reasons, we assume a unit of instruction that a child of instruction-referenced median rate will [illegible] in 15 30-minute periods--e.g., in 3 weeks, 5 days a week, 30 minutes a day.

So that we can get instruction far enough along to warrant extensive generalization responses based on prior acquisition of rules of an instructional knowledge system, we assume that the illustrative program unit is the third such in an ordered program of spelling units. We assume further that children entering this third unit of spelling instruction have similar entry skills profiles, although perhaps varying in chronological age and educational experience, including time spent [illegible] first two spelling units.

One could considerably unburden the system regarding on-line interaction [illegible] by having children sit around whenever the system found it convenient for them to do so. The pony express and bicycle [illegible] "interactive" systems manifest this predisposition toward "[illegible]-time interaction." It is possible that student-generated demand for real-time instruction will reduce [illegible]ership, thus lowering the on-line burden placed on the system. However, system-generated demands for "[illegible]" are to be tolerated only as a last-resort concession to state of the art, since the object is to increase educational productivity [illegible] [illegible] can hardly be done by sacrificing student instructional time [illegible] the machine. The objective is to [illegible] instruction [illegible] student characteristics. [illegible] state of the art, we [illegible] [illegible], rather than allowing [illegible] [illegible] [illegible] [illegible], [illegible].

to be particularly true for IDCMS, that the hardware will not stand in the way of a broad field of sophisticated uses of the system. [illegible] remarks of this paragraph could be interpreted to mean that we contemplate fully challenging the ingenuity of software people.

Our characterization of the anatomy of interactive instruction reveals that the system will react to a response sequence--or even to a set of such sequences--rather than to a single response to a single test item. Its reaction to a decision formed on a response sequence will be at the same level. That is, following a decision, the system will direct S to a lesson--or to an instructional sequence--rather than to a single instructional or test item. This might be contrasted with the item-to-item interactive models introduced over a decade ago by programmed instruction researchers. The item-to-item model assumes two things that do not seem reasonable: a) that a single response to a single item is a reasonable basis for evaluation, diagnosis, or prescription, and b) that the fact that S is wrong at the same time signifies why he is wrong. It may be a little slower going, but one can advance through Crowder-type programs simply by throwing darts at response alternatives; it is like playing a slot machine that comes up only 0 or 1, with 0 signifying loop and 1 signifying advance. The spelling program to be sketched will be interactive on a segment-to-segment, rather than an item-to-item, basis. We will characterize segments as of 2-20 item length, with abrupt termination of the segment a possibility whenever its items require item-to-item response through the keyboard and on-line evaluation of performance.

IDCMS permits student-system interactive studies using one or a combination of the following presentation modes: Video Only (VO), Audio Only (AO), Audio + Video (AV). Spelling instruction entails that many or most audio elements have durations on the order of 1 second--the time it takes to pronounce words of one- or two-syllable length. It will be assumed that the illustrative spelling unit for the most part deals with spellings of one-syllable words. The illustrative program will be made of AV items during instructional segments and AO items during test segments (but with complications to be revealed). Some or most instructional items will require keyboard responses. Responses to instructional items will be evaluated from a standpoint of appropriateness [illegible] to determine whether perceptual-mechanical skills assumed prerequisite to instruction are really there. Responses to test items will be evaluated against spelling criteria (which may differ depending on whether the response refers to a rote-memorized program word or to a rule-generalized novel word and on such factors as point in the instructional [illegible] the criterion will be applied).

[illegible] IDCMS [illegible] terminal [illegible] instructional [illegible]

[illegible] matter [illegible] different [illegible] line. [illegible] at the [illegible] by [illegible] performance [illegible]. [illegible] instructional [illegible]. [illegible] appropriate [illegible] redundant instruction and [illegible] to reach higher rates [illegible] that [illegible].

[illegible] segment of instruction [illegible] will [illegible] entail presentation of [illegible] items. [illegible] the ordering of such a series [illegible] not be required, [illegible] first presentation of the series could be presented in a specified [illegible] order without doing harm to pedagogical intent. [illegible] several alternatives [illegible] sequences from single audio and [illegible] elements. Setting [illegible] the use of external sources for [illegible] fixed-sequence segment [illegible] use implies development [illegible] that [illegible] these sources (which enter the [illegible] through [illegible]), the major remaining implication is that [illegible] system [illegible] only with longer program [illegible], even though most audio elements used will be quite short. [illegible] applications of [illegible] fixed-sequence segments [illegible] purposes [illegible] the illustrative inter[illegible] [illegible].

[illegible] view [illegible] instruction is that instructional segments, or [illegible], alternate with test segments, or decision points, with evaluation, [illegible], [illegible] prescription confined to the test segments. All know, however, that any instructional segment or item therein may [illegible] of responses—both covert and overt. If one wishes, one may [illegible] that S's eyes [illegible] during presentation of instructional item[illegible]. [illegible] S' demand rate for item presentation. Or one [illegible] define [illegible] segment to accept responses that reflect elements of skill, [illegible] skilled test responses. Illustrative instructional segment [illegible] require imitation (or copying) [illegible] referencing to the V element during each presentation. [illegible] with presentation of such a segment, the system will [illegible] series of responses against a criterion for acceptability. [illegible] acceptable, the system will select a most-entertainable hypothesis [illegible], based on special instructional software. This [illegible] will be tested and, if found tenable, will result in [illegible] suitable *prescriptive segment* of instruction—[illegible] items, [illegible]. Prescriptive segment [illegible] instructional [illegible] up to [illegible] segment [illegible] that [illegible] to be.

[illegible] segment [illegible] typically [illegible] entail presentation of [illegible] items, [illegible] the [illegible]. [illegible] need not be completed [illegible] decision can be reached [illegible] and [illegible] decision information. First presentation [illegible] test [illegible] would typically [illegible], later presentation [illegible] segment [illegible] could be [illegible].

[illegible] ertainty [illegible] acceptability [illegible] rule w [illegible] st-entertainable [illegible] instructional action [illegible] tenable, [illegible] instruction. Presumably [illegible] reference [illegible] in [illegible] specificity. [illegible] progress still [illegible]

An illustration [illegible] sequence [illegible] in [illegible]

1. Present Instructional Segment I, [illegible] data. Evaluate [illegible] where conditions for entertaining each are [illegible] asserts that performance is acceptable [illegible] H_0 [illegible] tenable, advances to next Segment I. [illegible] H_0 is rejected [illegible] the data suggest [illegible] H_1--instructional pace [illegible] fast. If data reveal [illegible] all [illegible] responses are correct [illegible] as they go but incomplete, then H_1 is tenable and Instructional Segment I is presented again at a slower pace. [illegible] H_1 also is rejected, but that [illegible] tenability of H_2--perhaps [illegible] unacceptable [illegible] information-processing [illegible] (key-pressing) deficiency. [illegible] sort of [illegible] be imagined. Let us suppose that the [illegible] to supply the next to last letter in [illegible] word whose letter sequence is [illegible] e.g., [illegible] P, [illegible] P for [illegible] P.

2. Present [illegible] Segment [illegible] (faster and reduced rate). [illegible] replicates the Instructional Segment I data (even though [illegible] listing that omissions extend [illegible] consistent [illegible] first letter position), then H_2 [illegible] accepted. [illegible]

3. Present Pre[illegible] Segment I.2.([illegible] instruct [illegible] the child to attend to all letters in the sequence. If the word is [illegible]IP, this might be accomplished by blocking all keys except I at the outset. Only then [illegible] P be unblocked (and I blocked). [illegible] the [illegible] keypresses. Advance to the next item [illegible] follows the pressing of P. This segment might be terminated either [illegible] acceptability criterion for the overall response [illegible] interpre [illegible]

[illegible]. Present Instructional Segment [illegible] performance data. [illegible] that data [illegible] acceptable performance.

[illegible]. Present Test Segment [illegible] to collect spelling performance data. Two sorts of test items might be distinguished: a) those that test for spelling of program words (words whose spellings were copied during instruction), which could be accomplished [illegible] on the basis of retrieving spellings from memory, and b) those which test for generalization of spelling rules by using as test items novel words (words [illegible] during instruction) that are consonant with spelling [illegible] thus far introduced. Unacceptable performance referring to program and to novel words should have different prescriptive implications.

Fastest and Average Item Consumption Rates. What value should the system evaluation-command-switching sequence of intervals have? The question is relative. Sometimes 5-20 seconds could be allowed; sometimes such delay would be unacceptable. The value of delay between presentation can be defined on the ratio of up to down time. Up time here signifies the time spent by the system presenting items and accepting responses; that is, up time is presentation-response time. Down time is student-waiting time, used by the system to evaluate the situation, issue commands, and execute these commands (to [illegible] by accessing items). The [illegible] instructional domain (e.g., spelling, phonics) invites few presentations of speech length. Responses should occur more than occasionally. Since unautomated spelling can feature rather high item presentation densities per unit time, we ask the system, among other things, to improve on such densities. We do not ask for switching in the [illegible] sense of speed. As time is measured in 1972, we make the apparently reasonable demand that the down portion of a short down-up "trial" not exceed a few seconds. [illegible] we propose a down-up ratio that is favorable to up.

A video system available permits the sequence of operations culminating in presentation of a TV frame on a CRT to occur in 1 second or less following detection of the controlling digital code under all but the most-impossible heavy-use conditions. The video system can [illegible] through required to true random access from 1800-frame video storage.

When the [illegible] wherein we find it impossible to do computer-assisted interactive work unless able to true random access from storage the size provided for IDCMS video but characterizing both video and audio, then we probably will need augment the system--e.g., through addition to it of analog-form audio storage that parallels that for video, but without losing the synchronization feature inherent in intrinsic digital codes associated with audio messages. However, the illustrative program requires audio storage at the level of the segment, [illegible] at the level of the audio element. Even though the illustrative program is not trivially short, we will argue later that the maximum number of such segments required will be far short of 1800, that audio storage can handle the illustrative program's audio component, and that we can tolerate the relatively slow segment-accessing times so long as the intrasegment interitem down times are of short duration. Given that a segment of interest is in S's audio buffer, we ask the system to step to a following item 1 second following completion of the response to the item that comes before it and to initiate video frame accessing by the end of that second. Assuming, then, a rather large intersegment accessing time, it still would be true, if intrasegment and interitem down times were averaged, that trial down time would average to less than 2 seconds for audio and less than 3 seconds overall. Hence, the illustrative down interval for instructional items is taken as 3 seconds. Acceptability of an interval of this magnitude depends on magnitude of the up interval.

Let audio come up on earphones and video come up on the CRT at the beginning of a 4th second. Audio will go off at the end of the 4th second. Conceivably, there will be younger children who could perform a 3 or 4 letter imitation response through the keyboard during 5th and 6th second. If so, then these fastest children reflect a worst case (a 3-sigma above the mean case). A median-paced child more likely will use 4 seconds (5th through 8th) to complete such responses (under system encouragement at that). Hence, we might think of the presentation-response interval as being 3 seconds under worst-case instructional conditions and 5 seconds on the average. Thus, worst-case item consumption rate for AV items will be 10 per minute (per terminal). The average item consumption rate for such items will be 7-8 per minute (per terminal).

Testing features AO mode presentation. Hence, V retrieval time drops out during testing, reducing down time to 2 seconds. If the argument is correct that the spelling response should require more time than the imitation response (I am not sure that it is), then the second saved from down time might be used during testing to prolong the up interval. (There is little point in relaxing A item retrieval time during testing unless it can be shown that different subsystems might profitably be employed, depending on whether AV or AO mode applies.)

Instruction probably will on occasion be punctuated with little speeches. Such speeches should improve the down-up ratio; whether they would do so in an instructionally productive sense is unknown.

The picture sketched above yields a worst case 1:1 ratio for AV and 1:2 ratio for AO; it yields an average-case 1:1.67 ratio for AV and 1:3 ratio for AO. (Should one wish to evaluate probable 6-terminal traffic under normal distribution, a 3-sigma below the mean case could be viewed as using three times as much presentation-response time as worst case.)

Were we five years into the future--that is, if pedagogical implications of pacing factors had been pretty well established in consequence of putting IDCMS to extensive use--then it might be agreed that ratios much less favorable to up time might suffice. However, IDCMS is the instrument one will use to evaluate such matters. We cannot test *toward* the upper limits for speeded instruction if hardware is allowed to limit the empirical attack so as to preclude such tests.

Foregoing remarks establish desirability of a worst-case item consumption rate of 10 per minute, referenced to one terminal. Magnitude of the storage requirement underlying support of such a rate remains to be determined.

Random Access Requirements. When one has a large collection of items any one of which one may wish to display at any time, then true random accessing to such items is required. If items of the collection do not class under a mnemonically-useful linear or hierarchical taxonomic scheme, then the sort of storage we see in Video Master Disc becomes required. Even if mnemonically-useful classification is possible, it has no useful implications for Video Master Disc storage when this is used in connection with IDCMS terminals, because multiple usership of IDCMS precludes a movable head staying in one place over item presentations referencing to a given S. Thus, it would be no advantage to group video frames physically on a disc just because they are likely to be used in the same segment of instruction or even in a fixed order in that segment.

When we transfer over to Audio Master Reproduction (the audio store), it again becomes difficult to see how a classification scheme could aid equipment response time per user under the condition of true random accessing from a large number of alternatives. However, the audio and video subsystems differ in one important respect. Audio Buffer can receive a sequence of audio elements that is physically grouped in Audio Master Reproduction (although I am not sure that such a fixed-sequence segment then can be made to run as the audio portion of a series of AV presentations whose pace is controlled by the system, at least not until different software is provided than apparently currently is contemplated by the contractor). Video Buffer can only receive one frame at a time. Two sorts of classification for the audio system might [illegible].

Assume first that children will be at or near the same point at the outset of the instructional program but at different points later. [illegible] different users entering (or wishing to enter) the same tape transport of Audio Master Reproduction at the same time, this might necessitate having 2-3 copies of the "first quarter" of the program available in Audio Master Reproduction at the outset--that is, 4-6 transports might contain one copy of the first quarter of the program (where 12 transports are available). In consequence, during this phase of instruction 2-3 terminals might be tied to the 4-6 transports containing [illegible] of the instructional segments for the phase. Queuing to [illegible] the same or adjacent segments would only be a problem at the outset while S [illegible]. Following an initial phase, the whole-copies feature of storage would drop out. Thereafter, the different transports might be used in linear organization to reflect the most-useful linear organization of remaining segments. Whether early or late, segments might be stored according to a system which minimizes predicted queuing to audio tape transports.

Assume second that most segments of instruction can be assembled off-line ahead of instruction in consequence of fully-explicated design-development specifications. That is, items of a segment can be ordered without respect to performance data and, if alternative orders are required, then differently-ordered but otherwise identical copies of the segment can be assembled, again without regard to performance. This need not mean that a segment must run its course. The segment could be terminated at any point in consequence of S reaching an upper or lower criterion bound defined through application of sequential analysis. The important questions then become a) How many audio sequences and reordered copies of these sequences would inhere in the largest portion of the program that needs to be stored on a worst-case day? b) Will we allow the system to assemble such sequences (when properly instructed) on-line?

Earlier consumption rate assumptions warrant the view that Day 10 will be the worst-case day for illustrative instruction. At the end of Day 10, the fastest child will complete the program, the average child will be 2/3 through, and the slowest child perhaps 1/3 through. Hence, worst-case storage is 2/3 of the program (although without whole-program duplication).

The amount of audio sequence storage determines system response speed when the audio tape system is used. If we assume sequences will run on the order of 2 minutes in the worst case, then we might well be willing to tolerate the almost-20 second rewind time it takes any tape transport to reach out to either end of its tape and to transfer 150 inches of storage to Audio Buffer--a worst case for storage position. What remains then to determine is how many such sequences will occur in the terminal 2/3 of the program.

If we define the central instructional path on the fastest child, we see it will be 600 minutes long, including central instructional path treatment and loops. Let us imagine that this child will loop one time in five. The 600-minute program yields 300 2-minute segments. Of these, the 4:1 ratio between mainline and loops apportions 240 segments to mainline and 60 to loops. Let us interpret this as 120 instructional segments, 120 test segments, 15 instruction-referenced diagnostic segments, 15 instruction-referenced prescriptive segments, 15 test-referenced prescriptive segments, and 15 alternative test segments. Let us imagine that the average child will show a 3:1 ratio between mainline and loops, necessitating apportioning 80 segments to loops. Further, let us imagine that the slowest child will show a 2:1 ratio between mainline and loops, necessitating apportioning [illegible] segments to loops. Hence, the fastest child will use [illegible] 300 ([illegible] + 60) segments, the [illegible] ([illegible]40 + 80), and the slowest child [illegible] (240 + 1[illegible]). [illegible] defines worst case for storage.

We indicated above how much looping we will assume children will do, but not then the looping segment storage requirement, [illegible] terms. For present purposes, let us assume that the slowest child's diagnostic, prescriptive, and alternative test segments are drawn from a pool of segments that is twice the size used--[illegible] segments. The unused segments (which might be used by other users) are there because we cannot be sure just what segments will be needed. If several sets of users revealed that some are never used, they could be deleted. Conversely, experience might also show that new segments were needed.

On the worst-case day (Day 10), the slowest child is approximately 1/3 through the program and the fastest child is completing the program. Assuming that the looping level is constant across the program--probably a worst case--then on Day 10, the system would need to store just in excess of (2 × 480)/3, or 320 segments of digitally-coded audio material. Let us assume 20 digitally-coded audio elements per segment plus material needed to address the audio portion of the segment. Let us assume that each segment can be stored on approximately 170 inches of tape ([illegible]5 seconds at 3.75 ips, with digital codes occurring *ahead* of their audio elements).

Based on the classification assumption made earlier, it is assumed that two students never will go to the same Audio Master Reproduction tape transport at the same time. S's own audio buffer contains two tape transports. We assume here that while one of these tape transports is in use, the system always will anticipate the most-probable next segment and cause it to be delivered to the other tape transport before the playing segment is completed. At worst, then, the system will sometimes be wrong. Only in those instances will interactive instruction suffer something on the order of a 20-second worst-case rewind-transfer-forward-rewind penalty. Such a penalty should be encountered no more than 1 time in 3 for the slowest child and 1 time in 5 for the fastest child, even if the system were operating on the basis of chance rather than on the basis of prior performance by S, including on the segment being negotiated. To summarize, the mean such penalty would be approximately 10 seconds, experienced one time in four (or less), [illegible].

Audio Master Reproduction consists of 6 modules, each containing [illegible] tape transports. Each transport is reel-centered on 490 feet of tape whose normal-play speed is 3.75, with bidirectional drive over either of two tracks. Each transport will take [illegible] minutes of audio material, which it will high-speed search at 150 ips. Each track in either direction from reel center measures 2940 inches. Taking the segment as 170 inches long, each track in either direction can store approximately [illegible]. A single tape transport, then, can carry on the order of [illegible]. [illegible] contains [illegible] tape transports, half of AMP storage will accommodate the worst-case day storage requirement ([illegible] segments).[3]

It appears possible that placing a large number of small program segments on a tape might necessitate minor modification of system hardware or software. The system typically employs a "2 program" per direction per track view of audio storage in AMP. Whether this is merely for expository purposes or reflects a system design commitment remains to be determined.

If a segment can be loaded into Audio Buffer and if the Buffer then can be made to start, pass over and relay the digital code in approximately one second, stop while video catches up, then pass over and transmit the audio element in synchrony with video, then stop until an instruction from the system controller turns off video, then start through the next item, then perhaps the sketched instructional program is consonant with the audio tape subsystem. If that is so, then the problem of referencing to number of stored audio "items" and accessing them gives way to other problems that research software design must address. While we have hinted at the magnitude and forms of such problems in the illustrative sketch, they are not central to the present paper and will be handled elsewhere.

[illegible] of the segments discussed above are AO, rather than AV. These are less a problem than AV, so warrant no special comment. (Whether AO segments, not of interest here, constitute a problem depends on whether a coded 55 Hz subcarrier is audible when unmasked. It depends on amplitude of the subcarrier.)

We probably would not take kindly to closing out on a high-speed retrieval capability referencing to a large collection of items in this system. Conceivably, careful study will reveal a need to augment such a program as has been sketched with a capability for retrieving, perhaps, one item from a large collection all of whose items are quickly retrievable. While the casual treatment presented above does not clearly reveal such a requirement when the equipment is under interactive research use, further study might show the research need for it; moreover, short-term research use clearly implies such a capability. Hence, we [illegible] should [illegible] pursue the question of an audio subsystem that [illegible] "start" audio elements [illegible] the analog-form video

---

[3] It is not clear from available contradictory materials whether transport storage should be [illegible] minutes or [illegible] minutes. Assuming the [illegible]-minute figure, then total AMP storage would need be used on the [illegible].

[illegible] later handle--although it remains to determine just how many items [illegible] would permit [illegible] need true random accessing to.

Interactive Research

One reason why I believe that IDCMS will increase SWRL's interactive research and instruction capability from tenfold to one hundredfold is that the state of the art base is low. A study now in formulation by John Koehler has follow-on interactive research implications, although they study itself will not require interaction in the sense of prescriptions that lead S off his treatment-defined mainline. One purpose of the Koehler study is to determine what sorts of prescriptive departures from mainline instruction in the phonics domain should occur. Such information, to be obtained during the second half of the current school year, should show the way toward follow-on research each of whose treatments features student-system interaction in the sense of performance-contingent departures from the mainline.

The Koehler study will evaluate each of three entry skills underlying the segmentation-blending level of phonics instruction, will assign S's to one of three training treatments consonant with entry proficiencies, and will retest entry skills (using novel rule words) following training. This sequence will be repeated through three cycles--one each for CVC, CCVC, and CVCC type items. Two training strategies will be used: one that makes the letter-sound rule the basis for segmentation and blending, one that employs bipartite patterns.

The entry skills pretesting will be as follows:

1. Pretest ES-1. 6 items.

1.1 Illustrative item. A+V = /kæt/+cat, imitate A (10 sec);
A+V = /bæd/+cat bad, imitate A (10 sec);
A+V = "Now say this word" + bat in the context cat bad. R = /bæt/ (15 sec).

1.2 Maximum time. Each item presentation will involve presenting three different AV pairs. The first two of these will require imitation responses in oral form, the third, a test response in oral form. I will evaluate these responses. If the response is made in less time than shown above, I will make a keyboard response that advances the ES-1 segment. If no such response is made in the time allowed, the system will automatically step the ES segment to the next AV pair. Maximum time is: 5 sec (ES-1 segment retrieval + duplication) + 6(3)(2) sec (down time) + 6(35) sec (max time) = 251 sec. = 4.2 min.

[illegible]

[illegible] = [illegible] (1 [illegible]).

[illegible] procedure [illegible] retrieval + duration [illegible] + [illegible](2) + [illegible] 1[illegible] min.

[illegible]

[illegible] [illegible] + [illegible] [illegible] = [illegible] ([illegible] sec).

[illegible] procedure and time [illegible] 2.[illegible] min.

[illegible] and below, [illegible] segments, segment compositions, [illegible] materials lengths, and time durations are meant to be [illegible] in the illustrative sense. Finer details remain to be specified [illegible] are unimportant here). The assumption is that IS-1, -2, and -3 segments for CCVC and CVCC cycles will have the same [illegible] characteristics as are shown above for CVC entry skill pretesting contents.

Training will emphasize analysis, synthesizing, or both, depending on entry skills pretest outcomes. Koehler and his associates distinguish a set of training subskills appropriate to proficient word attack of the segmentation-blending variety. If entry skills proficiencies indicate an analytic deficiency then Subskills 1-3 are germane (or so it is assumed), if pretesting indicates synthetic deficiency, then Subskills 3-5 are germane (or so it is assumed); if deficiencies of both sorts are indicated, then all of the subskills may need be addressed during training. Koehler indicates that all children will receive training directed to all subskills, but that the emphasis will vary depending upon the hypothesis reached in consequence of pretesting. Thus, for example, one hypothesized to be deficient in analytic skills would have Subskills 1-3 emphasized and Subskills 4-5 treated more perfunctorily. Because all training treatments feature Subskill 3--which uses the most [illegible]--for illustrative purposes we simply will assume that those suspected of being analytically deficient will receive essentially no time on Subskills 4-5 and those suspected of being synthetically deficient will receive essentially no time on Subskills 1-2. This permits spreading [illegible] different treatments in such a way as to more-realistically [illegible].

[illegible] training for the combined analytically-synthetically deficient [illegible]:

1. [illegible] 1. [illegible] 2 presentations = 4 trials.
   [illegible] spit; imitate A. (Illustration)
   [illegible] 5 + 4(2) = 13 sec, up = 4(10) = 40 sec, total time = 1 min.

2. Train 2. 2 items x 2 presentations = 4 trials.
   VA = p-i-t + /p/+/ɪ/+/t/; imitate A. (Illustration)
   Down = 5 + 4(2) = 13 sec, up = 4(10) = 40 sec, total time = 1 min.

3. Train 3. 6 items x 5 presentations = 30 trials.
   VA = a + /æ/; imitate A. (Illustration)
   Down = 5 + 30(2) = 65 sec, up = 30(3) = 90 sec, total time = 3 min.

4. Train 4. 2 items x 2 presentations = 4 trials.
   VA = m-a-p + /m/+/æ/+/p/; imitate A. (Illustration)
   Down = 5 + 4(2) = 13 sec, up = 4(15) = 60 sec, total time = 1.5 min.

5. Train 5. 2 items x 2 presentations = 4 trials.
   VA = m-a-p + "Say the word;" R = /mæp/. (Illustration)
   Down = 5 + 4(2) = 13 sec, up = 4(15) = 60 sec, total time = 1.5 min.

Again, we assume that CCVC and CVCC cycles will have the same system-hardening characteristics for Train-1 through Train-5 as does the CVC cycle. Training materials will not be the same for letter-sound and bipartite patterns; the same items will be used but segmented and blended differently in terms of units used. For the present we assume that the different strategies will be consonant with training coverage illustrated above for the letter-sound pattern strategy.

Posttests will test for the same skills as do entry skills tests except that novel rule words will be used. Because the letter-sound rule set will consist of only 6 items after Cycle 1 training, increasing to 12 items after Cycle 2 and to 18 items after Cycle 3, possibilities for generating novel rule words will be meager following Cycle 1 training and greater with completion of training on succeeding cycles. Hence, it is assumed here that there will need be fewer items in the CVC posttest than in later ones. While this could be remedied by teaching the most rules during Cycle 1 and the least rules during Cycle 3, there is no way to approach the experimental design problem that is entirely satisfactory. Here it is assumed that Cycle 1 posttests each will contain 4 items, Cycle 2 posttests 7 items, and Cycle 3 posttests 10 items. The assumption is that posttests will use the same amount of time per item as do entry tests.

The fixed segment view typically would not sit well with F, particularly for testing segments but, to a lesser extent, also for training segments. Before determining how many alternate random-ordered versions of a segment might be employed, let us inventory segments. There will be 3 ES test segments and 3 posttest segments for each cycle, or 18 different test segments. There will be 5 training segments for each cycle for each strategy, or 30 different training segments. Since segments will be of different lengths, it is necessary to calculate audio storage based on segment lengths. We will assume that maximum times given above entail that amount of audio storage, which usually will not be so. The 3 3-segment ES tests will cost 8.5 minutes each, or 25.5

minutes if stored together (or at the same time). The [illegible] 3-segment [illegible] will, respectively, require approximately 6, 9, and 1[illegible] [illegible], [illegible] minutes if stored at the same time. The 6 5-segment training segments each will require 8 minutes, or 48 minutes if stored at the same time. Assuming 26 x 12 minutes of audio storage, or 31[illegible] minutes [illegible], then if only a single version of each needed segment were [illegible] with [illegible] versions stored in the audio store throughout the study, just [illegible] of total storage would be used. Under the circumstances, [illegible] could afford 3 versions of each segment (or, if the tape transport capacity figure of 52 minutes is used, 6 versions of each segment). The research [illegible] probably precludes the degree of spreading that the instructional illustration presented earlier invites. If the worst-case day necessitates [illegible] 1/3 of the segments needed during the life of the study, then 26-minutes storage per tape transport is consonant with 9 randomly-ordered versions of each segment (and 52-minutes storage with 18 versions).

Foregoing remarks suggest that maximum times for analytically trained groups (2 groups, each N = 12), by cycle, will be 19, 22, and 25 minutes, [illegible] 66 minutes in all. Maximum times for synthetically trained groups ([illegible] groups, each N = 12), by cycle, will be 20, 23, and 26 minutes, or 69 minutes in all. Maximum times for combined analytically-synthetically trained groups (2 groups, each N = 12), by cycle, will be 22, 25, and [illegible] minutes, or 75 minutes in all. Assuming use of 20-minute periods only, [illegible] minutes of which was usable, then under the conditions of 6-terminal usage of IDCMS, one [illegible] per terminal, and good administrative support regarding scheduling of Ss, the study as illustrated could be accomplished using the system 4 hours per day (9-12, 1-2) for 5 successive school days. Assuming appropriateness of the character generator for all video segments, the study materials could be constructed without recourse to artists or other printshop personnel. If we made it a requirement that the system (and augmenting equipment) be used for purposes of constructing materials only during normal down time (e.g., before 9 a.m., and after [illegible], during lunch hour, on weekends), then perhaps a single technician could construct, organize, and load in everything needed to support a study of the sort outlined above in 8-16 hours without hindering system exploitation for research purposes. Doubling that figure to provide for checking and editing and doubling it again to accommodate study-tailored production of software (if that proves needed), the system might need to be used on the order of 48 hours during its down time for purposes of supporting study execution and 20 hours for purposes of executing the study and performing whatever analyses were required. The personnel bill--exclusive of study formulation and write-up--would be 48 technician hours, 120 research assistant hours, and perhaps 20 IDCMS supervisor hours. My prediction is that this would represent at least a fivefold savings on what the study actually will cost when conducted using mobile laboratory facilities at the schools and pre-IDCMS capability for providing graphic and analysis support. We need to do better than that; I suspect that we could do appreciably better the second time around.

It is much easier to make the case for [illegible], short-element, [illegible] when the task is to obtain [illegible] than when it is simply to tryout-evaluate instruction. While the latter task for the most part can be accomplished with fixed-order segments, experimental tasks almost always require that alternate orders be used to insure that an order bias is not operating. There do not [illegible] be so many different short audio elements in the Koehler study outlined above--a rough calculation suggests that there are between [illegible] and 250 such elements. However, it would be quite easy to increase this number 4 to 8 times in a slightly more ambitious study of essentially the same sort--e.g., by extending cycles to all possible one-syllable [illegible], all pertinent letter-sound rules, and all novel rule words consistent with these extensions. Given high-speed audio accessing, segment formation and randomization could be placed on the synthesizing basis that precludes appreciable investments in forming segments and alternate [illegible] of segments (or storing them in segment form). In that event, software would control segment formation and randomization. Whether this would mean simply that the materials construction savings would be obtained for a like (or greater) dissavings in programmer time should be evaluated.

DOCUMENT RESUME

ED 108 599 IR 002 110

AUTHOR	Follettie, Joseph F.
TITLE	On-Line IDCMS Evaluation of Different Categories of Response.
INSTITUTION	Southwest Regional Laboratory for Educational Research and Development, Los Alamitos, Calif.
REPORT NO	SWRL-TN-1-72-02
PUB DATE	7 Feb 72
NOTE	11p.
EDRS PRICE	MF-$0.76 HC-$1.58 PLUS POSTAGE
DESCRIPTORS	Computer Programs; Educational Development; Educational Research; Electronic Equipment; Evaluation; Information Processing; Input Output Devices; Instructional Media; *Instructional Technology; *On Line Systems; Patterned Responses; Primary Education; *Response Mode; *Taxonomy; Tests
IDENTIFIERS	IDCMS; *Manipulandum Referenced Taxonomy

ABSTRACT

A manipulandum-referenced taxonomy for response categories appropriate to primary education is presented. The tenability of automatic on-line evaluation of the different types of response when processing equipment of the sort that probably will be available to Southwest Regional Laboratory is preliminarily evaluated. (Author/SK)

SWRL

SOUTHWEST REGIONAL LABORATORY TECHNICAL NOTE

DATE February 7, 1972

NO TN 1-72-02

ON-LINE IDCMS EVALUATION OF DIFFERENT CATEGORIES OF RESPONSE

Joseph F. Follettie

ABSTRACT

A manipulandum-referenced taxonomy for response categories appropriate to primary education is presented. The tenability of automatic on-line evaluation of the different types or response when processing equipment is of the sort that most probably will be available to SWRL during the next year or so is preliminarily evaluated.

This document is intended for internal staff distribution and use. Permission to reprint or quote from this working document, wholly or in part, should be obtained from SWRL, 11300 La Cienega Boulevard, Inglewood, California, 90304.

ON-LINE IDCMS EVALUATION OF DIFFERENT CATEGORIES OF RESPONSE

Joseph F. Follettie

Whether on the basis of its own Version 1 capabilities or in consequence of augmentation based on core-to-core interfacing and use of supplemental terminals, we wish in time to transform the SWRL Instructional Development Control and Monitoring System (IDCMS) into one whose student-system interactive capability is not bounded by the category of response employed. This paper presents a taxonomy for response categories appropriate to primary education and preliminarily assesses the degree to which we can expect such a system soon to perform interactively as a function of response category. If we can visualize a relatively straightforward way of arming the system both to accept responses of a given category and to evaluate them on-line against a criterion for response acceptability, a preliminary view can be taken that the system has or can be made to have an interactive capability referencing to that response category. Otherwise, such capability will be realized only following extensive and perhaps costly effort. Unsurprisingly, the general conclusion is reached that those response categories whose responses cannot yet be on-line machine evaluated pose a greater challenge to linguists and semanticists than to those whose efforts reference to hardware design.

For present purposes, the taxonomy of response categories can be viewed as grounded upon three mutually-exclusive and, taken together, exhaustive superordinate categories. These are: a) motor responses using characters or codes (keyboard-referenced), b) voice responses (microphone-referenced), and c) spatial motor responses. In this system, motor and voice categories contrast. Hence, a visual tracking response would class as a motor response, since it clearly is not a voice response. Physiological responses below the molar motor level are not of present interest.

Motor Responses Using Characters or Codes

IDCMS is most clearly applicable to student-system interactive use when the response is formed on a keyboard containing alphanumeric characters and supplementing commands and codes. Three subcategories of such responses are distinguished.

Surface syntactic constructed response. Any phrase or sentence or unordered, semiordered, or ordered set of phrases or sentences which a test item invites and given instruction, if effective, constrains will class here if made through an alphanumeric or similarly coded keyboard. Sequential responses to complex arithmetic problems--e.g., response sequences which show steps enroute to problem solution--also class here. On-line processing of a surface syntactic constructed response presupposes existence of an effective parsing program (and perhaps augmenting programs addressing semantic and set-relational problems). Oversimply perhaps, the system will have to reduce the response to a "base" form and to compare this base form with a criterion base form response retrieved from an auxiliary store. While on-line processing of responses of this type are hardly a "no sweat" proposition, SWRL's Language Analysis Package effort currently is appreciably addressing the problem. It appears not imprudent to assume that we will be able to impute to IDCMS an on-line evaluative capability regarding an appreciable portion of such responses before too much longer, whether on the basis of Version 1 configuration of the system or in consequence of augmentation of the system's processing capability using a suitable auxiliary processer.

Alphanumeric item constructed response. Responses to simpler arithmetic problems, word-spelling responses, and item or concept naming responses will class here. Whatever the criterion response characteristics regarding accuracy or speed, on-line evaluation of such responses appears to pose no insurmountable challenge to software designers. Moreover, it should be possible to cause the system to model or shape such response sequences by progressively unblocking the next correct keypress in the sequence and, if desired, backlighting it. The alphanumeric item constructed response is the type of constructed response that the system can be made to on-line evaluate most easily.

Selected response. Constructed motor response options are constrained only by instructional effectiveness and surface alternatives inherent in syntax or mathematical structure. Selected response options are more narrowly constrained by the test item. The response must be selected from a sharply-constrained field--e.g., the alternatives of a multiple-choice item. Typically, such responses will be made by depressing a single key, whether alphanumerically coded or color or picture coded. Typically also, a portion of the keyboard will be blocked off for use in making such responses. The keyboard originally proposed for IDCMS was such a reduced keyboard--one consonant with making selected responses to multiple-choice items.

The selected response need not be restricted to a single keypress. Some interactive systems compel S to use a sharply-reduced vocabulary and syntax for purposes of interacting with the system. Often in such situations, the effect is to restrict S's response to one of a few alternatives--e.g., CORRECT, WRONG, DON'T KNOW. Even though S may have to learn the perceptual-mechanical skills underlying loading such responses into the keyboard, the responses themselves class as selected rather than constructed in the present context because a constructed response is constrained only by instructional effectiveness and structure of an interaction language of nontrivial complexity. That is, we call responses constructed only if ineffective instruction makes a large set of response alternatives entertainable. If the set of entertainable response alternatives must be small even though instruction is ineffective, then the response classes as selected. On-line evaluation of selected responses made through a keyboard appears well within IDCMS capabilities.

Voice Responses

The speech and speech element responses classed under this heading parallel motor responses using characters and codes excepting that input into a "written-form" analyzer must be output from a speech analyzer. This heading also subsumes trained voice responses--whether in the sense of diction or of music.

Surface syntactic speech response. Evaluation of speech responses at the syntactic level presumes everything that evaluation of motor responses at this level presumes and, in addition, a front end analysis performed by a speech analyzer that transform spoken into written form. It is probable that on-line evaluation of speech responses at the syntactic level will not be performed by IDCMS either in Version 1 configuration or in an augmented configuration that we can obtain soon. The alternative is to have E rather than the system perform real-time evaluation and to signal the result to the system. To the extent that E can do this quickly and reliably, then the E-augmented system gives us on-line evaluative capability regarding such responses.

Alphanumeric item speech response. Comments made immediately above apply. Letter sounds form an important subclass of the responses that class here. While on-line evaluation of responses of this subcategory by IDCMS seems a less challenging requirement than system on-line evaluation of speech responses at the syntactic level, here too the prognosis for automatic on-line evaluation very soon is not an optimistic one. Again, the E-augmented system can be used to achieve on-line evaluation during earliest use of IDCMS.

Selected speech response. This mode of responding simply is alternative to motor responding through a keyboard under the selected response condition. Having such a capability might prove desirable if it can be obtained cheaply. As an alternative to on-line evaluation using the E-augmented system, SWRL engineering staff might seek to design a sharply-constrained speech analyzer (if a cheap shelf item does not exist) which would discriminate between the voiced items of such series as 1, 2, 3, 4, 5 and white, black, red, green, pink. Such a device really would be a speech pattern discriminator rather than a speech analyzer.

Trained voice response. Whether this is a diction or music type response, the essential feature of the response is that it should match a standard. Perhaps this simply comes down to a comparison of two frequency x amplitude plots, one provided by S and one contained in system storage, with departure from standard guaged against allowable departure from standard. It is probable that early capability for on-line evaluation of such responses comes down to how finely differences between response and standard are to be discriminated. SWRL engineering staff has invested preliminarily in the problem. It appears tenable that on-line evaluation at least of certain types of musical response could be obtained early in the life of the system.

Spatial Motor Responses

We distinguish between drawing and tracking spatially-referenced responses and between "freeboarded" and "keyboarded" drawing responses.

Freeboard-referenced drawing responses. The RAND Tablet and devices currently in prototype at the University of Illinois, University of Pittsburgh, and elsewhere exemplify required hardware. Functionally, this is a terminal whose response area is a flat two-dimensional surface such that stylus-writing on the surface will bring that portion of the surface into contact with a sensing surface. Such devices permit copying responses, free form drawing responses, and intermediate responses. The mechanical analogue, for those unfamiliar with these devices, is the "blackboard" toy wherein a child applies a wood stylus to waxed paper over a black tarlike material. Following composition on such a device, it is returned to the preresponse state.

If a pattern is placed on the writing surface (in proper alignment), then a copying response can be evaluated on-line against a criterion for absolute or average permissible departure from pattern. Such patterns might include line drawings exemplarizing concepts, letters, or words. Given such a terminal, then on-line evaluation of copying responses should pose no insurmountable challenge to software designers.

However, if we allow the response to be free form, then on-line evaluation becomes a problem in pattern recognition. On-line system evaluation of free form drawing or printing responses should pose a considerable challenge for software designers, assuming that hardware in present or quickly-attainable form is up to the requirement. To the extent that on-line evaluation of free form responses by the system itself is not presently feasible, we can again fall back on the E-augmented system, which will be able to function in the desired way if E is sufficiently quick and reliable.

Intermediate to copying and free form responses are relational responses that reference to a pattern but do not involve copying the pattern. An illustration is the diagonal drawing response. A constraining square or rectangular outline is provided and S is required to line a diagonal--with or without a constraint on orientation--through the pattern. Similarly, one might provide S with the outline of a human face in front view and require placement of eyes, nose, and mouth. In this instance, we could avoid a pattern recognition requirement by having responses judged on placement rather than shape. Intermediate responses seem as capable of on-line evaluation as copying responses.

Keyboard-referenced drawing response. Imagine a terminal whose response area is an x, y matrix of keys elevated above a flat two-dimensional surface. Keys should be backlightable by the system or by S responding under system-defined conditions and desirably alternatively backlightable using different colors of light. Such a keyboard could be used either to require S to provide matrix-constrained free form exemplars of geometric concepts--probably with scale constraint imposed--or to complete partially patterned exemplars of such concepts. It is apparent that criteria for sequencing keypresses could be applied--e.g., that the diagonal be constructed progressively left to right. Given such a terminal, on-line evaluation of responses appears to pose no insurmountable challenge to software designers.

Tracking response. Tracking responses can be referenced to various sorts of terminal displays and to various sensorimotor modalities. The pursuit rotor tracking response is illustrative. Time on target is one measure of pursuit rotor proficiency. On-line evaluation of such responses appears to pose no insurmountable challenge to software designers.

* * * * *

Recursive and combinatorial usages of the foregoing categories will yield more involved responses. A child asked to order a set of pictures or sentences to tell a story will select an order and perform a response sequence that indicates the order selected. An identification response may take either selected or constructed response form. While a time-pressured identification response--e.g., at tachistoscopic levels--may necessitate finer resolution than a CRT provides and perhaps shorter interval exposure times than the system finds convenient, the required augmenting devices are shelf items and required software efforts do not seem undue. Discrimination threshold studies similarly might require special peripheral equipment but would turn up no new forms of response as responses are defined above--which is in terms of manipulanda. Responses signifying "There it is" or "I see it now" lend themselves readily to keypress expression. Where latency of such a motor response adds undesirably to a duration of interest, S's mean latency can be evaluated and the system can be required to take this out of all such durations.

The test, then, of the taxonomy presented above is not whether one can come up with responses whose names or sensorimotor involvements are different than the names and molar forms used as exemplars above but rather whether one can come up with responses requiring different sorts of manipulanda than those cited above. In this connection, the tracking response category contemplates various sorts of unnamed peripheral equipment--for example, an eye movement camera if visual tracking or information searching is to be studied.