

Assignment 7: The Task Model for your example assessment

Hannon and Daneman’s Reading Comprehension Assessment

Description of the Task Model—Tasks

The tasks in Hannon and Daneman’s (2001) assessment consist of *passages* that examinees read and true/false *questions* that they later answer. The tasks are highly structured; so much so that they could easily be generated by an algorithm (cf. Embretson, 1998).

Fixed Features

Tasks. The assessment consists of 1 practice passage with questions and 6 test passages with questions.

Passages. Each passage consists of 3 sentences, each containing 2 terms and 1 semantic feature. In each passage a total of 2 real terms and 3 nonsense terms are used. All real terms and semantic features are selected from those known to college students (the population that the assessment was designed for and piloted on). Although not specified, each sentence begins with a nonsense word. Although not explicitly stated, all semantic features are real words, not nonsense words.

Sample passage:

[TERM] [TERM] [semantic features]

“A TOLP resembles a MARB but is more colorful, larger and lives in a colony.

A MARB resembles a BUTTERFLY but is more colorful and larger.

A JERP resembles an ANT but is less colorful and larger.” (p. 126)

In the example above, TOLP, MARB, and JERP are the nonsense terms, BUTTERFLY and ANT are the real terms, and “colorful, larger” and “lives in a colony” are the semantic features.

Questions. All questions are true/false statements. One half of the statements are true, and one half are false, created by reversing the order of the terms in a correct statement.

Sample questions:

TM: “A MARB is more colorful than a BUTTERFLY.” [T]

TI: A BUTTERFLY is more colorful than a TOLP. [F]

KI (low): “A MARB is larger than an ANT.” [T]

KA (high): COCKROACHES have queens, whereas ANTS don’t. [F] (p. 126-127)

Variable Features

Tasks. The number of features per paragraph is 2 in paragraphs one and two, 3 in paragraphs three and four, and 4 in paragraphs five and six.

Passages. Each passage uses 2-4 semantic features; there does not seem to be any order in which they are repeated across the passage. When there are more than 2 features, which sentence has 3 features also varies from passage to passage, in no apparent order. Among the 2 real terms and 3 nonsense terms, 1 term is reused across the 3 sentences, but it can be either a real or nonsense term.

Questions. The design and number of questions for each type of question is complicated, and is represented in the Table 1. Each 3-sentence passage has 42-54 questions.

For example, a passage with 2 semantic features has 42 questions: [TM = 12] + [TI = 4] + [KI (low) = 4] + [KI (middle) = 6] + [KI (high) = 8] + [KA (low) = 4] + [KA (high) = 4].

Although not specified, questions follow one of 4 formats: 1) A [TERM1] is/does [semantic rel]. 2) A [TERM1] is/does [semantic rel] than a [TERM2]. 3) Like [TERM1], [TERM2] is/can [semantic rel]—used for KI(high) questions. 4) A [TERM1] has/is [semantic rel], whereas a [TERM2] doesn't/isn't—used for KA(high) questions.

Conditions of Administration

Each passage is presented on a computer screen 1 sentence at a time in a fixed order, and is replaced when the examinee pushes the “+” key. Each question is presented 1 at a time in random order, for a maximum of 12 sec per question. Examinees may not scroll backwards or forwards, either within the passages, within the questions, or between the questions and passage.

Stimulus Material and Work Product Specifications

Stimulus materials are presented on a computer screen; there are no specifications regarding the font size or color, background color, screen size, or programming language. The real and nonsense terms are always presented in CAPITAL letters, but the semantic relations are in lowercase letters. The “work product” is a key press on the computer keyboard signifying Yes for True or No for False. Failure to answer is counted as an error. There are no specifications regarding criteria for key presses (e.g., how to interpret an ambiguous key press).

Rationale—Relationship of the Task Model to Other Aspects of the Assessment

With respect to the student model, each question is designed to tap one (and only one) of

the student model variables, with questions differentiated among low, middle, and high levels of student model variables. (This suggests a 7-variable student model as opposed to a 4-variable student model.) The TM, TI, KI, and KA formats are clearly linked to the student model. However, it is open to doubt whether each question taps only one SV—text memory, as our ETS counterparts pointed out in the April 29th class, is surely involved in all of the questions. Further, while the design of the questions seems linked to the student model, the complexity of the number of questions does not appear to be linked. That is, what is it about text memory that demands 12 questions, not more or less, and what is it about and middle knowledge integration that demands 6 questions, not more or less? Likewise, what is it about the student model that dictates, for example, that nonsense words come first or that only KI(high) questions can take the format shown in example 3 above?

With respect to the statistical model, each answer has only 1 observable variable (unlike many performance assessments (e.g., Clauser et al., 1997). The answer provides the True/False evidence required by the evidence interpretation model and what appears to be a CTT model.

With respect to the psychological model, the tasks clearly fail to tap the vital comprehension components of vocabulary and world knowledge that Hannon and Daneman cite in their own literature review.

With respect to the purpose of the assessment—psychological theory building—again the task design fits Hannon and Daneman’s restricted theory of comprehension, but not mainstream comprehension theories.

Table 1: Task model variables and numbers of items with given TMV values

	Level	Is it used in the text?			Number of questions	
		Term	Term	Semantic feature	If this many semantic features are used	...there are this many questions (½ true, ½ false)
Text memory	—	Y	Y	Y	2	12
					3	14
					4	16
Text inferencing	—	Y	Y	Y	2	4
					3	6
					4	8
Knowledge Integration	Low	Y	Y	Y	Any number	4
	Middle	N	Y	Y	Any number	6
	High	N	Y (nonsense)	N	Any number	8
Knowledge Access	Low	Y	Y	Y	2	4
					3	6
					4	8
	High	Y	N	N	Any number	4