

Preliminary exams practice

Responding to Guthrie, J.T. (1988). Locating information in documents: examination of a cognitive model. Reading Research Quarterly, 23, 178-199.

This response is to study 1 only. Responses written by Joseph Martineau, April 2002.

A. Theoretical perspective.

1. Critique the author's conceptual framework.

The conceptual framework is a cognitive model for document reading, found on page 183. It was developed using a "cognitive components" approach which performs a sequential task analysis consisting of multiple simple tasks performed in a certain order, for completing a complex task. One problem of this cognitive components model is that it could be re-conceptualized as a non-linear task. While the linear model presented is a reasonable model, it could also be presented with the following modifications.

- i. The iterative process as implied by the returning arrows could return directly to "encode question features," since the features may need to be re-encoded if they have been forgotten or not comprehended.
- ii. The process need not start with encoding the question features. By definition, it was the first mini-task in this model validation study because it was the first task displayed on screen, and participants were instructed to read the question first. Even in this scenario, participants could skip the encoding and go directly to the menu. Reading or scanning the document may actually precede encoding the questions for some readers.
- iii. Goal formation could be broken into two components: (1) encoding question features, and (2) identifying unsatisfied features.
- iv. It could be that category selection (does document contain info in a relevant category) occurs directly after encoding the question features, and determining whether that category has already been used to satisfy a feature occurs after identifying the relevant categories.
- v. It is likely that integration of information is implicit in all the other steps. It may not be a separate step at all. It seems that the combination of goal formation, category selection, and information extraction could be integration itself.
- vi. The distinction between goal formation and recycling is trivial. It seems that recycling and goal formation are the same thing. In fact, the two "Exit"s in the model are redundant with each other. The return arrow to goal formation should go directly from integration to goal formation.

These possible additions to the model make it non-linear, and more complex. This does not necessarily change the necessary analyses for testing contributions of the components of the model. This presents a difficulty for validating the model, because the analyses selected could be used to test the components of this simple model, or the components of the much more complex non-linear model.

The conceptual framework also includes an assumption that efficiency in locating information well enough to successfully complete a task can be predicted by the quality of readers' approaches to each of the components in the conceptual model of document reading. This is spelled out explicitly in describing the statistical methods used in the study, but is not really described before the methods section. It ought to be explicitly connected to the cognitive model description with a defense for why this can provide evidence that the cognitive model is reasonable.

The conceptual framework also includes the assumption that each of the five components of the cognitive model can be learned over time. This is nowhere evident until the description of the statistical methods used in this study. Again, this ought to be described in conjunction with the description of the cognitive model with a defense for why it provides evidence for the plausibility of the cognitive model. Testing this assumption is critical to providing evidence that this model will be at all useful, but this rationale is never provided.

2. Comment on the need for this study and its importance

The authors do not explicitly provide evidence of the need for this study and its importance, but they do provide some implicit evidence in that there are few young adults who can read documents well for extracting specific pieces of data to solve a problem, and that there are few studies of reading documents for obtaining information. There is a logical chain that is not described that goes something like this.

Few young adults can adequately do document reading for information, THEREFORE we need to teach students while they are in school to read documents for information much better than we are, BUT we don't really know what to teach BECAUSE part of the problem is that we don't have a good model for document reading for information, THEREFORE we need to develop AND validate a reasonable model for document reading for information THAT includes teachable components.

This less than obvious logical chain has to be developed by the reader, rather than being spelled out by the authors. In developing this logical chain for myself, I can see the need for this study and its practical importance; but the authors really need to explicitly spell out the chain, since there was quite a jump from the problem to the proposed model (e.g. how is the proposed model going to help educators address the problem).

3. How effectively does the author tie the study to relevant theory and prior research

The study is well connected to the NAEP results showing very little ability of young adults to read documents well, to a definition of the domain of document reading tasks, and to a parallel process of cognitive components theory for describing complex tasks, but because of the jump from problem to a model whose applicability to addressing the

problem is not obvious, a complete tie is somewhat weak. Discussing how other models of reading comprehension have assisted in developing instructional strategies to address comprehension problems would provide a good tie-in to prior research and theory that supports developing a cognitive model to address the problem.

4. Evaluate the clarity and appropriateness of the research questions or hypotheses.

The research questions and hypotheses are not clear. They are vaguely stated as “to examine the relative contribution of the five components to document search performance (p. 184).” They are implicit in the statistical procedures used to model the data, as mentioned in point 1. Explicitly spelling out the research questions would make it much easier to link the data analyses to sections of the cognitive model, and to the implicit assumptions that the components can be taught, and that their quality of performance on each component leads to efficiency in performing the task.

The statistical hypotheses tested are reasonable if we buy the implicit assumptions just mentioned, but they don’t really seem reasonable given that another stated purpose of the study is to “suggest a cognitive processing model of document search that will account for the proficiencies and deficiencies found in this aspect of literacy (p. 181).” It is not clear how to devise a study that would accomplish such a task (fuzziness is the issue here).

B. Research Design and Analysis

5. Critique the appropriateness and adequacy of the study’s design in relation to the research questions or hypotheses.

The research design was a one-shot case study, but it was labeled as an experiment. The purpose of the study was actually a sort of construct validation. Specifically stating some researchable questions in plain language terms would be helpful in determining whether the research design actually addresses the questions the author wanted to ask, or whether it is simply fortuitous that we can construct a logical link between the conceptual framework and the research design.

Can a one-shot case study using correlational data validate the constructs implied by the cognitive model? It seems that it can provide some evidence, but the evidence is weak. The evidence provided really answers whether efficiency can be accounted for by component task performance quality, and whether the components can be learned, not whether the cognitive model is adequately specified (including the number of components, conflation of multiple components into a single component, splitting components into multiple components, linearity of the model, and ordering of the components). It only provides some limited evidence for the existence of the components.

6. Critique the adequacy of the study's sampling methods (e.g. choice of participants) and their implications for generalizability.

Given that the author would like to validate this cognitive model for document reading for information, with no apparent exceptions, the sample selected seems woefully inadequate. The sample was a convenience sample that was self-selected from undergraduate students in educational psychology and child development courses. Especially given the considerations of the literature review that included young adults of widely varying educational backgrounds, it seems that this can be only a rudimentary start on a cognitive model, unless we are willing to make the assumption that the model cannot include or exclude different components from one subpopulation to another.

7. Critique the adequacy of the study's procedures and materials (e.g. interventions, interview protocols, data collection procedures)

The data collection procedures seem to be moderately adequate to obtain measures spelled out. There were two major issues in this portion of the study. First, in real life, the material on flight information is available all on the same screen, but in this study each category was selected one at a time. The author claims face validity for the whole data collection process, but this disconnect between authentic and inauthentic situations calls this claim into question. Second, the author attempts to force the respondents into the first stage of the cognitive model (goal formation) by starting them into reading the question before looking at any document information. The big issue was the validity of the measures used, as discussed below.

8. Critique the appropriateness and quality (e.g. reliability, validity) of the measures used.

The authors used the number of times the participants returned to the question as an indicator of integration, but did not include any measure of the goal formation. It seems that the number of returns to the question could just as easily be a measure of goal formation; since goal formation is part of an iterative process. High numbers of returns is seen as a negative indicator of integration, but it could actually be a positive indicator of goal formation; with participants returning explicitly to evaluate the next mini-task to be performed. A think-aloud protocol (especially with only 26 participants) would not be an overly burdensome substitute that would allow the researchers to better distinguish between these two mini tasks. Using the number of returns to the question also builds in a relationship with efficiency in complete task performance, since the number of returns is also a measure of mini-task efficiency.

By operationalizing category selection as the number of categories selected, the authors are building in a relationship between complete task efficiency and this variable by defining it as mini-task efficiency. The appropriateness of the category selection given a think-aloud protocol would be a much more valid way to measure whether participants know how to select a category to achieve a mini-goal.

Were the use of the notepad obligatory, the measure of information extraction would be reasonable. However, because it was not obligatory, the measure seems to take on different meanings for some participants than for others. A think aloud protocol would also help to address this problem.

The kicker is the measure of recycling. First, finding a unique measure for this is an impossible task, since it is really the goal formation component all over again. Second, defining recycling as the quality of the category selection sequence makes no sense whatsoever. How does one determine whether the goal has been met by measuring the quality of the category selection sequence? This measure is simply nonsensical. No wonder it had no significant effect either in predicting efficiency or in being predicted by time point. Furthermore, this assumes that there is a set, most efficient path. This path has been defined by the researchers as winnowing the field the most quickly at each step. However, if there are three pieces of information that have to be used in winnowing, three category choices is the most efficient route, regardless of whether the winnowing occurred most rapidly in the first step or not.

9. Critique the adequacy of the study's data analyses. For example: Have important statistical assumptions been met? Are the analyses appropriate for the study's design? Are the analyses appropriate for the data collected?

There was no discussion of assumptions. Some mention of concern would have been a reasonable requirement, especially since time taken to complete tasks, number of menu choices, and number of returns to question were probably highly positively skewed.

The analyses performed are analyses that can validly be performed with this kind of cross-sectional one-shot case study data. Reasonable analyses were performed given the assumption that counts can be treated as continuous outcomes in a traditional parametric statistical analysis.

This judgment only evaluates the appropriateness of the analyses for the data and design, not the appropriateness of the design for the desired construct validation.

C. Interpretation and implications of results

10. Critique the author's discussion of the methodological and/or conceptual limitations of the results.

The author briefly discusses and dismisses another method (eye movement analysis) to get at some of the constructs he is trying to measure. He ignores a think-aloud protocol (which comes with its own limitations) as another alternative. The discussion itself excluded any limitations of either the design or the interpretation of the results. It did not discuss the limitations of the measures used, nor of the design to validate the cognitive processing model. The discussion was centered around only the results of the statistically defined research hypotheses. No discussion of generalizability was present, but should

have been included given that this study is a construct validation with no apparent exceptions.

11. How consistent and comprehensive are the author's conclusions with the reported results?

Given that the author makes few conclusions, and those he makes stick to the research hypotheses defined by the statistical analysis, he does not go outside the ability of the analyses to answer the questions posed. There are two problems with his approach. First, he assumes that his measures of the outcomes were entirely valid, when there are serious problems with assumptions surrounding the instruments. Second, he makes no explicit connection to theory. All connections to theory have to be assumed by the reader.

12. How well did the author relate the results to the study's theoretical base?

As stated above, the author makes little to no connection to the literature cited in the literature review, and makes no real connection to a construct validation of his cognitive processing model. This is a serious weakness of the study. Its conclusions are that the operationalizations of the components of the model (regardless of whether the model is correct) account for 68% of the variation in overall task efficiency, when those relationships have been built in automatically (see point 8), and that the components can be learned over time. This provides a little contaminated evidence for the existence of the components, but provides no construct validation for the model itself that he proposes.

13. In your view, what is the significance of the study, and what are its primary implications for theory, future research, and practice?

The significance of the study is that the study provides some contaminated evidence for the existence of the proposed cognitive components, and some evidence that the components (as we are willing to accept their operationalizations) can be learned.

The implications for theory are that these components may be helpful in creating and validating one or more cognitive processing models for document reading for information.

The implications for future research are (1) that another study that attempts to measure these components should be done using operationalizations that do not contaminate the relationship between component performance quality and overall task efficiency, (2) an other study that attempts to more validly measure the components using think aloud protocols may provide additional evidence for their existence and relationship with overall task efficiency, and (3) that true construct validations be attempted now that there is some evidence for the existence of these components, and (4) that future research should investigate the teachability (as opposed to learnability) of these constructs.

The implications for practice are not yet defined, since the teachability of the components has not been investigated, and since the model has not begun to be construct validated.