Measuring Task Complexity in Information Search from User’s Perspective

Yuelin Li, Yu Chen, Jinghong Liu, Yuan Cheng
Department of Information Resource Management, Business School, Nankai University
94 Weijin Road, Nankai District, Tianjin, China, 300071
yuelinli@nankai.edu.cn, chenyu0812232, liujinghong@mail.nankai.edu.cn, chengyuan0228@126.com

Xuan Wang, Ping Chen, Qianqian Wang
School of Software, Nankai University
94 Weijin Road, Nankai District, Tianjin, China, 300071
wxvqh111@126.com, chenping0727@163.com, wangqianq@mail.nankai.edu.cn

ABSTRACT
Task complexity is a critical task characteristic that influences users’ information seeking and search behavior. In general, researchers differentiate task complexity from objective and subjective task complexity. Both significantly affect task performance. However, few studies have been done to examine how task complexity could be measured from users’ perspective in information science. The present study identifies a set of objective and subjective measures and conducted a survey. The survey asked users to judge the complexity of a task, and then give the reasons why they made that judgment based on the measures. Six simulated task situations were developed for the survey and 168 valid questionnaires (84% return rate) were analyzed. The results indicate that the number of words hard to understand, the number of languages required for search results, and the number of domain areas involved in a task could significantly predict task complexity. The study helps further understand the attributes of task complexity and has implications in research on interactive information retrieval (IIR), task-based information seeking and search, and personalization of information retrieval (IR).

Keywords
Task, Task complexity, information search, objective and subjective measures

INTRODUCTION
Task has drawn more and more attention in information science in recent years. It has been recognized as an influential factor in information seeking and search.
complexity is a psychological experience or perception of task doer (Campbell, 1988). Most of studies developed task situations based on pre-defined objective task complexity, for instance, Maynard and Hakel (1997), Li and Belkin (2010), and so on. Subjective task complexity is also defined in different ways. Byström and Järvelin (1995) defined it based on a priori determinability. It is also measured by asking the participants to assess their perception of the complexity of the tasks or the statements indicating task complexity (Li and Belkin, 2010; Maynard and Hakel, 1997). Because the way to measure work or search task complexity is not consistent, it is very hard to compare the research results from different studies. This not only hinders the integration of the research findings from various studies, but also prevents the understanding of task complexity and its effects on information seeking and search. Thus, it is necessary to examine how to more appropriately measure task complexity and pave a way to a consistent measure.

LITERATURE REVIEW

Objective Task Complexity

Objective task complexity is a characteristic of task (Campbell, 1988). According to Wood (1986), three types of complexity should be taken into account when measuring and calculating task complexity, namely component complexity, coordinative complexity, and dynamic complexity. The total task complexity should be the function of these three types of complexity.

Though Campbell (1988) did not propose measurements for task complexity, he pointed out that objective task complexity “implies an increase in information load, information diversity, or rate of information change” (p.43). He identified four basic complexity attributes: (1) multiple paths to tasks, (2) multiple desired outcomes of tasks, (3) conflicting interdependence among paths and desired outcomes, and (4) uncertain or probabilistic links among paths and desired outcomes. On the basis of these attributes, 16 types of task were identified. Campbell (1988) characterized all these tasks into Simple tasks, Decision tasks, Judgment tasks, Problem tasks, and Fuzzy tasks with the increasing of task complexity.

Maynard and Hakel (1997) examined the effects of objective and subjective task complexity on task performance. In this study, objective task complexity was operationalized as the amount of information that the participant needed to integrate to complete a one-week schedule for a small fictitious film processing store.

Grill and Hicks (2006) synthesized 13 existing task complexity constructs in a comprehensive literature survey, such as degree of difficulty, sum of JCI (Job Characteristics Index) or JDS (Job Diagnostic Survey) factors, degree of stimulation, amount of work required to complete the task or information load associated with the task, amount of knowledge, size, number of paths, degree of task structure, non-routineness or novelty of task, degree of uncertainty, complexity of understanding system or environment, function of alternatives and attributes, and function of task characteristics.

In information science, objective task complexity is not frequently examined. Li and Belkin (2010) defined objective task complexity based on the quantity of sub-tasks involved (Li and Belkin, 2010). The study found that objective work task complexity significantly affected different aspects of users’ interactive information search behavior. Saracevic and Kantor (1988) used the number of terms in a search query to indicate the complexity of the information task. Gwizdka and Spence (2006) examined how users’ information search behavior could predict task difficulty. They measure objective task complexity of fact-finding web-navigation tasks by using the path length, page complexity, and page information assessment. Path Length refers to the length of the navigation path leading to the target information; page complexity means the complexity of navigation choices on each web page (i.e., visual design, link labels); page information assessment measures the difficulty of relevance judgment or the difficulty of assessing information scent.

Subjective Task Complexity

Subjective task complexity is a variable to measure task doers’ perception of the degree of complexity of a task. Most studies in information science are concerned with subjective task complexity from this perspective. Byström and Järvelin (1995) measure task complexity based on “a simple, uni-dimensional complexity categorisation of tasks based on, from the worker’s point of view, a priori determinability of, or uncertainty about, task outcomes, process, and information requirements” (p.194). They characterized work tasks into five categories. Bell and Ruthven (2004) take a similar stance to develop search task situations with varying complexity and define it based on a priori determinability of information required, the way to locate information, and relevance assessment (Bell and Ruthven, 2004). Li and Belkin (2010) measured subjective task complexity by asking the participants to assess their perception of task complexity. Maynard and Hakel (1997) measure subjective task complexity based on a 4-item scale: 1) I found this to be a complex task; 2) this task was mentally demanding; 3) this task required a lot of thought and problem-solving; 4) I found this to be a challenging task. The purpose of all items is to describe users’ subjective perception of task complexity.

The review indicates that task complexity has not been well conceptualized and it has been measured in different ways. Researchers measure objective task complexity based on the characteristic of task, while subjective task complexity was measured based on task doers’ assessment or from their perspectives. However, task complexity is such an influential factor in information seeking and search that it is necessary to examine how users measure task complexity.
from objective and subjective perspectives, by which to identify significant measures. Therefore, this study specifically investigated:

*How do users measure task complexity in terms of objective and subjective measures?*

It is assumed that objective measures measure objective task complexity and subjective measures measure subjective task complexity. Thus, objective measures should be those describing the characteristics of a task; subjective measures should reflect users’ perception or their subjective judgment of task complexity.

**METHOD**

To explore the issue, the researchers identify a set of objective measures based on the previous studies and then developed corresponding subjective measures. This could help further test the reliability of objective measures as well as see whether these subjective measures are appropriate to measure or predict task complexity. Based on these measures, a questionnaire was developed and a survey was conducted to identify the most effective measures for task complexity from users’ perspective.

**Questionnaires**

*Simulated task situations*

To elicit the participants’ assessment on task complexity, six simulated task situations (See Appendix) that involve the description of work tasks and search goals, were developed based on the target group’s daily work tasks, i.e., students’ daily life tasks and assignments. To reduce the participants’ load, the six simulated task situations were separated into two groups with three tasks each (randomly assigned). The questions or statements to measure task complexity for each simulated task situation are the same. Therefore, for this study the researchers in fact developed two questionnaires, i.e., Questionnaires I and II.

*Measures for task complexity*

For each simulated task situation, the participants were asked to assess the degree of complexity of the task, based on a 5-point scale from very simple (1) to very complex (5). Based on the literature review, the research group identified 11 measures for task complexity. To avoid a too-long survey, the group conducted an informal interview. Eleven students (convenience sample) were asked to select which measure is related to task complexity. Based on the participants’ selection, seven measures and users’ perception of its influence on task complexity, namely objective measures and subjective measures respectively, were finally identified as tested measures. The seven objective measures are as follows:

- Number of keywords in the task description: measured by a 3-point scale, i.e., too more (>=5)(3), fair (3-4)(2), too less (<2)(1).
- Number of sub-tasks of work task: measured by a 3-point scale, i.e., too more (>=5)(3), fair (3-4)(2), too less (<2)(1).
- Number of domains involved in work task: measured by a 3-point scale, i.e., complex (>=3 clauses)(3), fair (1=<clauses<3)(2), too less (no clauses, one simple sentence)(1).
- Complexity of syntax structure in task description: measured by a 3-point scale, i.e., extremely complex (>=5)(3), fair (3-4)(2), too less (<2)(1).
- Complexity of language structure in task description: measured by a 3-point scale, i.e., complex (>=5)(3), fair (3-4)(2), too less (<2)(1).
- Perception of keywords: the extent to which do you think the number of keywords in the task description affects the complexity of the task at hand. It is measured by a 5-point scale from Extremely (5) to Not any (1).
- Perception of terminologies: the extent to which do you think the number of terminologies included in the task description affects the complexity of the task at hand. It is measured by a 5-point scale from Extremely (5) to Not any (1).
- Perception of terminologies: the extent to which do you think the number of languages involved in search results affects the complexity of the task at hand. It is measured by a 5-point scale from Extremely (5) to Not any (1).
- Perception of domain areas: the extent to which do you think the number of domain areas involved in the task affects the complexity of the task at hand. It is measured by a 5-point scale from Extremely (5) to Not any (1).
- Perception of sub-tasks: the extent to which do you think the number of sub-tasks in the task description affects the complexity of the task at hand. It is measured by a 5-point scale from Extremely (5) to Not any (1).
- Perception of hard words: the extent to which do you think the number of words hard to understand in the task description affects the complexity of the task at hand. It is measured by a 5-point scale from Extremely (5) to Not any (1).
- Perception of language: the extent to which do you think the number of languages involved in the search results affects the complexity of the task at hand. It is measured by a 5-point scale from Extremely (5) to Not any (1).
- Perception of complex tasks: the extent to which do you think the number of complex tasks in the tasks at hand. It is measured by a 5-point scale from Extremely (5) to Not any (1).
- Perception of syntax structure: the extent to which do you think the syntax structure in task description affects the complexity of the task at hand. It is measured by a 5-point scale from Extremely (5) to Not any (1).
- Perception of easy tasks: the extent to which do you think the number of easy tasks in the tasks at hand. It is measured by a 5-point scale from Extremely (5) to Not any (1).
- Perception of hard tasks: the extent to which do you think the number of hard tasks in the tasks at hand. It is measured by a 5-point scale from Extremely (5) to Not any (1).
- Perception of any tasks: the extent to which do you think the number of any tasks in the tasks at hand. It is measured by a 5-point scale from Extremely (5) to Not any (1).

**The survey**

The survey was conducted in one of the top universities in China. Questionnaires were randomly distributed in the
campus. Questionnaires I and II were distributed to 100 students respectively. Totally 200 questionnaires were distributed. For Questionnaire I, 85 questionnaires were returned and 83 were returned for Questionnaire II. In total 168 questionnaires were completed and returned. The return rate is 84%.

RESULTS

SPSS 19.0 was used for data analysis. Reliability of the questionnaire was tested (Cronbach’s Alpha=.775). The result indicates that the questionnaire is reliable to measure task complexity in this study.

This section first reports the characteristics of the participants, and then presents the results of correlation and multiple regression analysis.

Participants

Figure 1-4 reports the participants’ gender, level, age, and major distribution. Among the students who completed the survey, 68.5% of the students are female; 93.4% of the students are undergraduates; 89.2% of the students are between 19-22 years old; 57.7% of the students are from social science.

The students evaluated their capability of online information searching. 26.8% of students evaluated their capability as “excellent” and 66.1% of students as “fair”; In general, 9.5% of students are extremely satisfied and 82.1% of students are fairly satisfied with online information searching process. Overall, the participants of this study are familiar with online information searching.

Task complexity and the objective measures

Pearson correlation analysis was performed to examine the correlation between task complexity and objective measures. Figure 5 shows the correlation coefficients of each objective measure to task complexity. It could be seen that the number of keywords in a task description has no significant correlation with task complexity. All other measures are significantly correlated to task complexity. Based on the correlation coefficients, task complexity is more closely related to the number of words hard to understand, number of languages required for search results, and number of domain areas involved in the task.
R Square of Model 3 indicates that it is the best one to significantly explain the variance of the dependent variable, and Table 2 indicates that in this model all three predictors are significant predictors of task complexity by t-tests.

![Figure 4. Distribution of the participants in terms of major](image1)

![Figure 5: Correlation between task complexity and the objective measures](image2)

**Table 1. Task complexity and the objective measures: models**

<table>
<thead>
<tr>
<th>Model</th>
<th>Predictors</th>
<th>R</th>
<th>R²</th>
<th>AD R²</th>
<th>Df/RE</th>
<th>F</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Constant</td>
<td>.39</td>
<td>.16</td>
<td>.15</td>
<td>1/488</td>
<td>89.77 **</td>
</tr>
<tr>
<td></td>
<td>Hard words</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>Constant</td>
<td>.43</td>
<td>.18</td>
<td>.18</td>
<td>2/287</td>
<td>54.48 **</td>
</tr>
<tr>
<td></td>
<td>Hard words</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Languages</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>Constant</td>
<td>.44</td>
<td>.19</td>
<td>.19</td>
<td>3/486</td>
<td>38.90 **</td>
</tr>
<tr>
<td></td>
<td>Hard words</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Languages</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Domains</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

(M: Model; AD R²: Adjusted R²; RE: Residual)

![Table 2. Task complexity and the objective measures: predictors](image3)

**Table 2. Task complexity and the objective measures: predictors**

<table>
<thead>
<tr>
<th>Model</th>
<th>Predictors</th>
<th>B</th>
<th>Beta</th>
<th>t</th>
<th>Tolerance</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Constant</td>
<td>1.76</td>
<td>.39</td>
<td>16.26**</td>
<td>1.00</td>
</tr>
<tr>
<td></td>
<td>Hard words</td>
<td>.59</td>
<td>.39</td>
<td>9.47**</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>Constant</td>
<td>1.39</td>
<td>.35</td>
<td>9.85**</td>
<td>.92</td>
</tr>
<tr>
<td></td>
<td>Hard words</td>
<td>.52</td>
<td>.35</td>
<td>8.09**</td>
<td>.92</td>
</tr>
<tr>
<td></td>
<td>Languages</td>
<td>.37</td>
<td>.17</td>
<td>4.05**</td>
<td>.92</td>
</tr>
<tr>
<td>3</td>
<td>Constant</td>
<td>1.22</td>
<td>.48</td>
<td>7.84**</td>
<td>.86</td>
</tr>
<tr>
<td></td>
<td>Hard words</td>
<td>.48</td>
<td>.48</td>
<td>7.20**</td>
<td>.87</td>
</tr>
<tr>
<td></td>
<td>Languages</td>
<td>.31</td>
<td>.15</td>
<td>3.35**</td>
<td>.87</td>
</tr>
<tr>
<td></td>
<td>Domains</td>
<td>.17</td>
<td>.11</td>
<td>2.55*</td>
<td>.84</td>
</tr>
</tbody>
</table>

** p<.01; * p<.05

Normality assumption was tested by examining normal probability plots of residuals. Also, the tolerance (see Table 2) and Durbin-Watson test (Durbin-Watson=1.759) were conducted. No violation of normality was detected. Also, no multicollinearity was found and adjacent observations are independent (Durbin-Watson is between 1.5 and 2.5). In terms of the independent variables, the number of words hard to understand, the number of language required for search results, and the number of domain areas involved in
the task could significantly predict the dependent variable, i.e., task complexity.

**Task complexity and the subjective measures**

Pearson correlation analysis was also performed to test the correlation between task complexity and the subjective measures. Figure 6 indicates that all subjective measures are correlated with task complexity, among which users’ perception of hard words has the highest correlation coefficient with task complexity. Except the measure ‘perception of keywords’, the results indicate that the users’ response to the objective measures is credible.

![Figure 6: Correlation between task complexity and the subjective measures](image)

**Figure 6: Correlation between task complexity and the subjective measures**

To further examine to which extent the subjective measures could predict task complexity, a stepwise multiple regression was conducted between task complexity and the subjective measures. Only one model was produced ($F(1, 488)=31.55$, $p<.01$) with one significant predictor ($t=5.62$, $p<.01$), which is users’ perception of hard words. However, the R Square is only .061. The model is weak to explain the variance of the dependent variable, that is, task complexity.

**DISCUSSION AND CONCLUSIONS**

This study examined how a set of objective and subjective measures could measure task complexity from users’ perspective. The results indicate that in general objective measures are more closely related to task complexity than the subjective measures. The study found that three objective measures could significantly predict task complexity, including the number of hard words included in the task description, the number of languages required for search results, and the number of domain areas involved in the task. In terms of the capability of predicting task complexity, the objective measures are more powerful than the subjective measures in this study. One of the reasons may be that the participants evaluated the subjective measures based on their perception, which was varied across different participants, but it was easier for the participants to select relatively consistent items based on the objective measures. This may indicate that the objective measures are more reliable than the subjective measures.

To some extent, the study provides empirical evidence to support Campbell’s (1988) definition for task complexity, and also support the way Li and Belkin (2010) to define objective task complexity. Though the number of sub-tasks is not a significant predictor of task complexity, it is really significantly correlated to task complexity. The study also supports Byström and Järvelin (1995) and Bell and Ruthevan (2004). Task complexity is significantly related to a task’s information requirements, specifically, the number of languages required for search results in this study.

The study selected more objective measures than previous studies to measure task complexity. The results indicate that objective task complexity has more than one indicator. This result has implications in information retrieval research. It helps construct task situations at different levels of task complexity. Task types and situations are important for information retrieval research. Previous studies usually constructed task types with varying complexity based on researchers’ understanding of task complexity. How users measure task complexity is usually ignored. This study attempted to fill the gap and shed light on the issue. The results indicate that the significant predictors of task complexity should be taken into consideration when constructing task types with varying complexity in IR evaluation. In this way we could get more reliable research results in task complexity and information seeking or search behavior.

This study identified several measures as predictors of task complexity. That means a causal relationship between the hard words, languages required, and domain areas and task complexity. More hard words included, more languages required for search results, and more domain areas involved in the task lead to more complex task. Hard words included and multiple domain areas involved indicate to which extent a task is understandable, and more languages required for the search results indicate to which extent the search results are understandable. This suggests that task complexity reflects users’ capability of understanding a task and its outcomes. This adds one more point to Campbell’s (1988) four dimensions of objective task complexity and Gill and Hicks’ (2006) summary of the existing task complexity constructs.
The limitation of this study is that the investigation did not cover all subjective measures used by previous studies. All subjective measures are users’ perception of the extent to which the objective measure could influence their task complexity judgment. Also, the scale for each measure is arbitrary and need to refine in the future. For the language measure, only a two-point scale was set up. This is because the study was conducted in China, where users usually need to search for documents written in different languages, especially in Chinese and English. But if in the USA, the language measure may not be salient. Because the study is based on a survey, we could not know the reasons why the participants select a certain item; for example, why they judge a task as a complex one if its task description includes words hard to understand. Future studies will enlarge the investigation to include more measures for task complexity. In addition, a follow-up interview will be conducted to further reveal how users judge the complexity of a task.

Future studies will also attempt to explore how to personalize IR based on task complexity. Personalization of IR has been a hot topic in IR community in recent years. To articulate the relationships between task or task characteristics and information search behavior and implement the results to system design has been an effective approach to the personalization, as Wu, Liu, and Chang (2008) and Liu (2009) have presented. Future studies will be based on the results of this study and develop simulated task situations with varying complexity to explore the relationships between task complexity and users’ information search behavior, based on which it is possible to construct a mathematical model that could calculate the complexity score of a task. Then, the system could automatically respond to the score and provide necessary help for users. In this way, it is possible to personalize IR based on task complexity.

ACKNOWLEDGMENTS

This study is sponsored by “National Undergraduate Students’ Innovative Experiment Plan” (2010-2012, No.101005565) at Nankai University, China.

REFERENCES


APPENDIX

Simulated task situations in Questionnaire I

1. You are a senior who is preparing for graduate enrollment examination. You need to buy some mathematic books written for the examination, which should cover calculus, probability and statistics, and linear algebra. These books should include main points in these areas and the exam for graduate enrollment examination each year. Please do a search to locate appropriate books, including the book title, publishers, and editors (at least three books should be found).

2. You are a freshman in the Mathematical department. You are planning to apply for double-degree in economics. So, you need to search for some information, including different majors in the School of Economics, curriculum, course schedule, job marketing, and enrollment threshold for different majors.

3. You would like to take part in CAD exam, and need to know the scope of the exam, the time and place for the exam, requirements for taking part in the exam, and the fee. So, you need to do a search for gathering the information and also other similar exams for comparison.

Simulated task situations in Questionnaire II

1. You are planning for a travel from Tianjin to Ganshu. Please search for the following places: Dunhuang Caves, Taer Temple, Sun-Moon Mountain, and Qinhai Lake. You need to locate the exact address of these places, learn the features of these places, search for flight and bus information, and order the hotels online.

2. You are preparing for the application of “National Undergraduate Innovative Experiment Plan.” You need to know related studies on personalization of information retrieval. Please search for papers related to task complexity and personalization of IR in the journal “Library and Information”, and list all English references in the papers you located.

3. You are preparing for an examination on computer science. You need to know information about the types of modes in which the computer is started, and make a contrast between a safe mode and a VGA-driven mode, figuring out both distinctions and relations of the two. Please do a search to collect related information.