Heuristics as Mental Shortcuts in Evaluating Interactive Systems

https://doi.org/10.3991/ijep.v8i4.8054

Amando P. Singun, Jr.
Higher College of Technology, Muscat, Sultanate of Oman
Amando-Singun@hct.edu.om

Abstract—Heuristics refer to the specific "rules-of-thumb" discovered from knowledge or experience which can simplify the complexity of making judgements. Heuristics are mental shortcuts to draw conclusions when evaluating interactive systems. In this study, a set of heuristics had been discovered by end-users while developing a series of prototypes of a test blueprint system. This study suggests that the design process of an interactive system should cater to the following two (2) components, namely: technical heuristics and specialized domain heuristics. Heuristics from these components should be the emphasis during the evaluation of the interactive system that has been designed using a user-centered paradigm of development called the Interaction Design Model (IDM).

Keywords—heuristics, test blueprint, user-centered development, prototypes

1 Introduction

In this study, the end users had been consulted and collaborated with the author to effectively capture the actual requirements for the development of an interactive system. The interactive system under study is a test blueprint designed to judge students' learning performance [1]. The author believes that the users are key resources to capture system's requirements. As shown in Fig. 1, collaboration by a cohesive team of user representatives is evident that ultimately promotes the attainment of a shared goal and continuous improvement of the interactive system. Each user has a specific role to play, thus, actively engaged into the design of the interactive system.

As shown in Fig. 1, the test construction is performed in a two-way line of collaborative process of communication process by the users, namely: Course Lecturers, Course Coordinators, Program Coordinator, and Head of Section – all of whom are directly involved in the several stages of the interactive system design and implementation. The author asserts that in an organization where peer collaboration is valued and supported, the best way to capture the requirements among themselves in the design of an interactive system is to utilize prototypes where they could interact among each other.

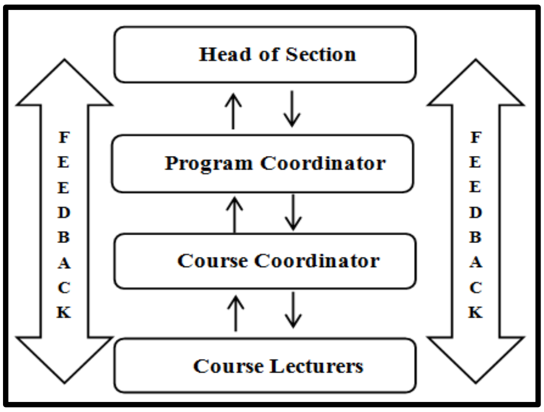

Fig. 1. Organizational Structure and Protocol (Course Level)

2 Conceptual Framework

A robust user-centered paradigm called Interaction Design Model (IDM) has been broadly used in designing interactive system. As shown in Fig. 2, the following four (4) interwoven key stages optimize the involvement of users throughout the design process: a) Establishing Requirements Stage which solidifies a constant evidence-based mode of searching information from a vast-range of sources to form requirements; b) Designing Alternatives Stage which includes the creation of a number of creative ideas that materialize the users' requirements; c) Prototyping Stage which enables user to interact with the system by immersing them through a look-and-feel-walkthrough with the system. According to the author, interaction with the prototype is the most sensible and collaborative way for users to evaluate the interactive system design; and d) Evaluation Stage which is an ongoing appraisal about the acceptability level of the system. Once the system is up-to-acceptable-level, it is going to get released ready for use by the intended users. However, if some requirements require refinements then these are fed back to the preceding stage(s) for appropriate action.

The specific stage which has been emphasized in this study is PROTOTYPING which greatly relies on the insights and feedback of the intended users regarding the prototypes of the interactive system. It allows designers to better understand users' real needs and preferences as it allows them to generate more ideas and articulate constructive feedback. Experts [2] claimed that "[…] prototypes are widely recognized to be a core means of exploring and expressing designs for interactive computer artifacts. It is a common practice to build prototypes in order to represent different states of an evolving design and to explore options. Simulating a design through prototyping can reduce design risks without committing to the time and cost of full production." Prototyping is highly intensive in the design process of interactive systems as it may suggest iterations, if needed.

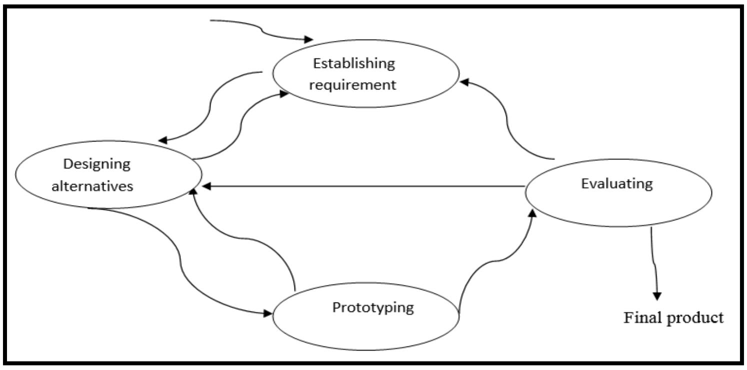

Fig. 2. Interaction Design Model [3]

The author observed that throughout the design of the interactive system, the intended users had been showing a sense of ownership as revealed by their active participation. Also, an apparent collaboration had enriched the interaction and improved communication by 'thinking aloud' their insights, feedback, and concerns towards the improvement of the system.

It is suggested to have a number of prototypes in order to allow users to select the best solution from the set of alternatives. Some designers are using prototypes to evaluate existing ideas. Users' rationale about their needs and requirements greatly influences the design of a more usable interactive system. There are many tools and techniques available for prototyping which ranges from 'paper-and-pencil' to 'more advanced technologies'. Prototypes do not necessarily produce fully-functional alternatives but rather are concrete representations that may envision and reflect the final product. It can be a low fidelity prototype or high fidelity prototype. In this study, two (2) samples of prototypes of the test blueprint have been iteratively developed: (1) test blueprint using paper-and-pencil as the earlier version and (2) test blueprint using spreadsheets as the derivative version. The final product is a web-based test blueprint system.

3 Statement of the Problem

This study is motivated to show how to design an interactive system (e.g. web-based test blueprint system) using a user-centered model of development. Specifically, it aims to discover the appropriate heuristics from the series of test blueprint prototypes that are essential in the development of a usable interactive system. Such heuristics are subsequently used as mental shortcuts to evaluate the said interactive system with emphasis on two (2) components, namely: technical heuristics and specialized domain heuristics.

4 Research Design: Methods and Procedures

This study made use of a descriptive research design with the following methods and procedures.

4.1 Literature Review

The author finds it very helpful to do a thorough review of related literature to establish content validity resulting to acceptable content validity index of the heuristics.

4.2 Focus Group Discussion (FGD)

The FGD session comprises of user representatives who could provide information, insights, feedback, and suggestions that are related to their respective role/tasks. It is a useful qualitative method of capturing specialized domain requirements that may assist the designer. The author of this study served as the moderator of the FGD.

4.3 Prototyping

A series of early versions called prototypes had been used in the development of a usable interactive system. Several iterations had been done in order to accommodate the intended users' needs and requirements.

5 Participants of the Study

There were double experts (n=5) who participated in determining the content validity of the heuristics identified by the designer in consultation with the users. Double experts are those individuals who have multiple areas of expertise, including an area related to the specialized domain-under-study.

The number of participants (n=10) during the Focus Group Discussion (FGD) session was within the recommended range which is six to twelve key informants [4]. The FGD comprises of the Course Coordinators, Course Lecturers, Program Coordinators, and Heads of Section.

Moreover, there were ninety-three (93) intended users who actually interacted with the sample prototypes of the interactive system.

6 Data Analysis

The following tools were used to determine the content validity and reliability of the identified heuristics:

6.1 Content Validity Index (CVI)

The researcher made use of an empirical method called Content Validity Index (CVI) to analyze the feedback of the panel of double experts. The CVI determines the relevance of the heuristics. It is used to check whether or not the items adequately represent the specialized domain of content [5]. The double experts rated the items of heuristics based on their level of relevance, using the rating scale shown in Table 1.

The different key formulae used in Content Validity Indexing are listed in Table 2.

Table 1. Relevance Scale for Content Validity Indexing

Rating	Description
1	Not Relevant
2	Somewhat Relevant
3	Quite Relevant
4	Highly Relevant

Table 2. Key Formulae for Content Validity Testing

Entity	Formula
I-CVI (Item Content Validity Index)	I-CVI = No. of Agreement / No. of Raters
S-CVI/Ave (Subscale Content Validity Index / Average)	S-CVI/Ave = Average of I-CVI
S-CVI/UA (Subscale Content Validity Index / Universal Average)	S-CVI/UA = Total Agreement / No. of Items
No. of Agreement	The count of raters whose response is greater than or equal to 3, 3 being the marginal response
Total Agreement	The count of No. of Agreed Items by the raters Total Agreement = Total number of items agreed by raters

6.2 Cronbach's Alpha

The Cronbach's alpha, which is sometimes called the coefficient of reliability, is used to measure the internal consistency of a test or scale [6], which had been used to determine the reliability of the items of heuristics identified in this study.

The rules of thumb corresponding to a Cronbach's alpha coefficient is shown in Table 3. Generally, it ranges from 0 to 1. The closer Cronbach's alpha coefficient is to 1.0 the greater the internal consistency of the items in the scale.

Table 3. Cronbach's Alpha Coefficient [7]

Cronbach's Alpha Coefficient	Description
0.91 – 1.00	Excellent
0.81 – 0.90	Good
0.71 – 0.80	Acceptable
0.61 – 0.70	Questionable
0.50 – 0.60	Poor
Less than 0.50	Unacceptable

7 Results and Findings

7.1 Prototypes

The author found out that prototypes used by the intended users can be "proofs-of-concept" geared to discover the appropriate heuristics derived from the sample prototypes through the use of Literature Review. It was revealed that the longer these prototypes have been exposed to users, the greater ideas are generated by them that could contribute to the enhancement of the design of the interactive system. From the feedback received from users, the procedure for test construction preparation and examination approval of the institution had also been improved. To put it simply, the logical and physical designs of the interactive system, as well as the business process of the institution, had been enhanced.

Below were the two samples of prototypes of the test blueprint that had been iteratively developed: (1) test blueprint using paper-and-pencil as the earlier version, shown in Fig. 3, and (2) test blueprint using spreadsheets as the derivative version, shown in Fig. 4.

TABLE OF SPECIFICATIONS
E-Commerce for IT (ITBS2203)
First Semester, AY 2011-2012

Course Code: ITBS2203 Course Title: E-Commerce for IT

TOPICS TAUGHT/ CHAPTER TITLE	LEARNING OUTCOME NO.	INSTRUCTIONAL OBJECTIVES	BLOOM'S TAXONOMY COGNITIVE LEVEL	NO. OF HOURS SPENT	% OF TOTAL HOURS SPENT	NUMBER OF MARKS PER TOPIC	Allocated Marks Per Question/Cognitive Level
IA. Introduction to Electronic	3, 4, 7, 10	Explain the differences between E-Commerce and E-Business.	Comprehension				1

Course Code: ITBS2203 Course Title: E-Commerce for IT

CHAPTER NO.	Knowledge	Comprehension	Application	Analysis	Synthesis	Evaluation	Total
Chapter 1T		9		3			12
Chapter 2T		2		4			6
Chapter 3T	1	4					5
Chapter 1P			27				27
TOTAL	1 (2%)	15 (30%)	27 (54%)	7(14%)	0 (0%)	0 (0%)	50(100%)
	16 (32%)		34 (68%)				

Note: Synthesis and evaluation levels are tested in their course project/assignment.

Fig. 3. Paper-and-Pencil Prototype of the Test Blueprint

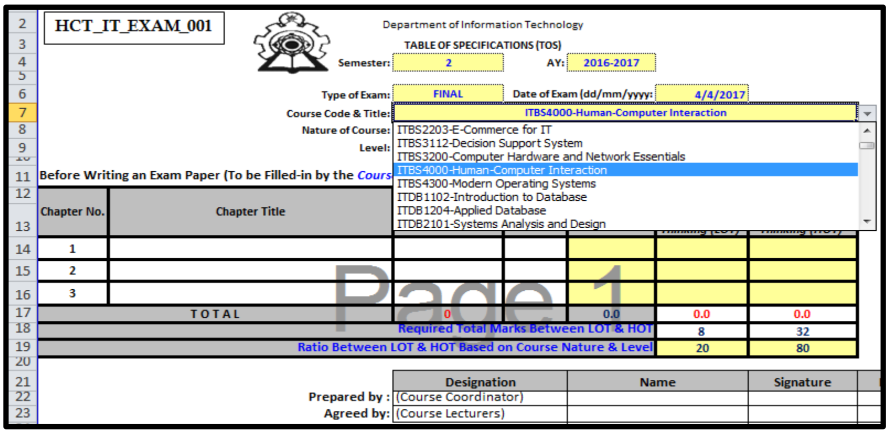

Fig. 4. Spreadsheet Prototype of the Test Blueprint

7.2 Identified Specialized Domain Heuristics through Literature Review

A substantial and rigorous literature review had supported in framing heuristics that pertain to the specialized domain i.e. Classroom Assessment vis-a-vis web-based test blueprint system. The following specialized domain heuristics had been identified as summarized in Table 4, namely: (1) Content Validity, (2) Fairness and Comprehensiveness, (3) Accountability, and (4) Flexibility. Under each heuristic, there are sub-heuristics (n=5) written in clear statements. The specialized domain heuristics contain twenty (20) specific heuristics in total. These heuristics may serve as mental shortcuts when evaluating the usability of the interactive system.

Table 4. Specialized Domain Heuristics

Test Blueprint Heuristics
1. Content Validity
1.1 The Test Blueprint improves content validity of my students' exam [8]. 1.2 The Test Blueprint ensures that my students have achieved a specified standard of achievement or learning expectation at the end of their exam [9], [10]. 1.3 The Test Blueprint makes it easier for me to identify what types of test questions are required based on the cognitive levels of LOT and HOT [11]. 1.4 The Test Blueprint helps me to construct a test which focuses on the key areas or topics which are weighted according to importance or significance [12], [13], [14]. 1.5 The Test Blueprint provides a link between what is taught and what is tested both in content and in skills required [8].
2. Fairness and Comprehensiveness
2.1 The Test Blueprint ensures that all the required course outcomes to be tested have been adequately covered [11], [14]. 2.2 The Test Blueprint ensures that there is a representative sample of questions from each chapter required to be covered [8]. 2.3 The Test Blueprint measures appropriate varied cognitive levels of my students which test different learning skills and types of exams [11], [14]. 2.4 The Test Blueprint creates a balance of testing between lower-ordered thinking skills and higher-ordered thinking skills that are appropriate to the level of my students [11], [14]. 2.5 The Test Blueprint produces an exam paper that does not discriminate the different types of learners [8].

3. Accountability
3.1 The Test Blueprint helps me to create a degree of my own accountability [8]. 3.2 The Test Blueprint makes me confident to answer students' complaints of dissatisfaction or any kind of exam appeals at the end of the exam [12], [13], [14]. 3.3 The Test Blueprint can create quality exam resulting in a high GPA's credibility of my students [11], [14]. 3.4 The Test Blueprint is my effective exam paper preparation tool [12], [13], [14]. 3.5 The Test Blueprint is more effective when collaboratively prepared by a team [9], [10].
4. Flexibility
4.1 The Test Blueprint makes the teachers creative in writing an exam paper [8]. 4.2 The Test Blueprint preparation is a simple task after having been used to it [11]. 4.3 The Test Blueprint can prepare a common exam paper among multiple sections [9], [10]. 4.4 The Test Blueprint format can be modified according to the needs of the institution. [12], [13], [14]. 4.5 The Test Blueprint is following the existing exam procedures of the institution [8].

7.3 Content Validity Index and Reliability of Heuristics

The heuristics listed in Table 4 were tested by 3-5 expert evaluators using Content Validity Index (CVI). With this number of expert evaluators, the most acceptable value for I-CVI is 1 and the ideal value for S-CVI/Ave is 0.90 or higher [15].

As reported in Table 5, all the expert evaluators agreed that each item is relevant as shown by the S-CVI/Ave which is equivalent to 1, which exceeded the ideal value for S-CVI/Ave which is 0.90. All the items under each subscale are correlated well with each other as shown by the S-CVI/UA which is equivalent to 1.

Table 5. Summary of Content Validity Testing Results using Content Validity Index (CVI)

Domain Heuristics (Test Blueprint Heuristics)			
Subscale	***S-CVI/Ave***	***Total Agreement***	***S-CVI/UA***
1. Content Validity	1	5	1
2. Fairness and Comprehensiveness	1	5	1
3. Accountability	1	5	1
4. Flexibility	1	5	1

On the other hand, the heuristics are deemed reliable as reported by the reliability coefficient of each item which is greater than 0.7. As shown in Table 6, the Cronbach's alpha coefficient is 0.76 which is considered acceptable.

Table 6. Reliability Statistics

Cronbach's Alpha	**Interpretation**	**No. of Items**
0.76	Acceptable	20

7.4 Final Product (Web-based Test Blueprint System)

Using the Interaction Design Model, the designer in consultation with the intended users, was able to craft the final product (web-based test blueprint) as shown in Fig. 5.

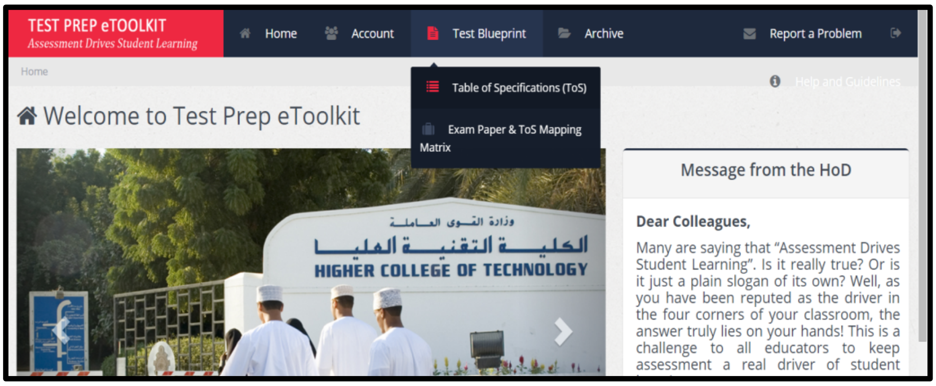

Fig. 5. Web-based Test Blueprint System Home Page

Fig. 5 shows the interface of the web-based test blueprint system which contains several menus that make it even more usable. There are two (2) submenus, namely: Table of Specifications (TOS) and Exam Paper & TOS Mapping Matrix. The test blueprint comprises of the chapter number and chapter title, the percentage of chapter weight, the weighted mark, the actual allotted mark, and the level of difficulty (LOT/HOT ratio). On the other hand, the 'Exam Paper and TOS Mapping Matrix' shows how items should be mapped against the cognitive levels of the Bloom's Taxonomy of Learning. The exam paper should perfectly match with the approved TOS.

8 Conclusion and Recommendations

Based on the summary of findings of the study, (1) Heuristics can be discovered while interacting with prototypes. Such heuristics may serve as helpful mental shortcuts or benchmarks when evaluating the usability of an interactive system; (2) Prototypes are cost-effective and easy to use in generating more ideas from the intended users. Prototyping activities should always be included in the design process of an interactive system. It can be utilized as "proofs-of-concept"; (3) The following two (2) essential domains should be considered in the development of interactive system, namely: technical domain (e.g. user interface design) and specialized domain (e.g. classroom assessment); (4) The users acquire a sense of ownership, support, and goodwill over the final product because of their level of involvement throughout the design. The use of the user-centered approach to development called Interaction Design Model is recommended in designing interactive systems; and (5) The interactive system and the other prototypes presented in this study can be used as patterns for developing test blueprints by other educational institutions. The format of the test blueprint may vary depending on the needs and requirements of the educational institution.

9 References

[1] Singun, A. (2016). Application-based Test Blueprint for a Summative Classroom Assessment, Proceedings of the 10th International Management Conference "Challenges of Modern Management", November 3-4, 2016, Bucharest, Romania.
[2] Houde, S. & Hill, C. (1997). What do prototypes prototype?, Elsevier Science B. V. Amsterdam.
[3] Rogers Y., Sharp H. & Preece J. (2011). Interaction design: Beyond human-computer interaction (3 rd ed). A John Wiley and Sons, Ltd, Publication.
[4] Maughan, P.D. (2003). 'Focus Groups' in Developing Web-Based Instruction: Planning, Designing, Managing, and Evaluating for Results, eds E.A. Dupuis, London: Facet Publishing.
[5] Polit, D.F. & Beck, C.T. (2004). Nursing Research: Principles and Methods (7th Ed.), Philadelphia: Lippincott, William, & Wilkins Portland, OR: ETS Assessment Training Institute.
[6] Cronbach, L. (1971). Test Validation. In R. Thorndike (Ed.), Educational Measurement (2nd Ed., pp. 443-507). Washington, DC: American Council on Education.
[7] George, D., & Mallery, P. (2003). SPSS for Windows Step by Step: A Simple Guide and Reference. 11.0 update (4th ed.). Boston: Allyn & Bacon.
[8] American Educational Research Association, American Psychological Association, & National Council on Measurement in Education (2014). Standards for educational and psychological testing. Washington, DC: American Educational Research Association.
[9] Mehrens, W.A. & Lehmann, I.J. (1991). Measurement and evaluation in education and psychology. (3rd Ed.) New York: Holt, Rinehart, and Winston.
[10] Nortar, C.E., Zuelke, D.C., Wilson, J.D. & Yunker, B.D. (2004). The Table of Specifications: Ensuring Accountability in Teacher Made Tests. Journal of Instructional Psychology, 31, 115-129.
[11] Linn, R. L. & Gronlund, N. E. (2000). Measurement and assessment in teaching. Columbus, OH: Merrill.
[12] Wolming, S. & Wikstrom, C. (2010). The concept of validity in theory and practice. Assessment in Education: Principles, Policy, and Practice, 17, 117-132. https://doi.org/10.1080/09695941003693856
[13] Brookhart, S. M. (2004). Assessment theory for college classrooms. New Directions for Teaching and Learning, 2004(100), 5–14. https://doi.org/10.1002/tl.165
[14] Gronlund, N. E. (2006). Assessment of student achievement (8th ed.). Boston: Pearson/Allyn & Bacon.
[15] Lynn, M.R. (1986). Determination and Quantification of content validity. Nursing Research. https://doi.org/10.1097/00006199-198611000-00017

10 Author

Dr. Amando P. Singun, Jr. is with the Higher College of Technology Muscat, Sultanate of Oman as an IT Lecturer, Quality Assurance Officer, Former Head of the Information Systems Specialization of the IT Specialization Sub-Committee, Member of the Institutional Standards Accreditation Assessment Application Support Team of the College, the College Representative to the Oman Quality Network in Higher Education (OQNHE), and the Editor-in-Chief of the IT Department Newsletter.

This article is a revised version of a paper presented at the International Conference on Interactive Collaborative Learning (ICL2017), held September 2017, in Budapest, Hungary. Article submitted 02 December 2017. Resubmitted 20 December 2017. Final acceptance 21 December 2017. Final version published as submitted by the author.