Evaluation of the Cognitive Dimensions Questionnaire and Some Thoughts about the Cognitive Dimensions of Spreadsheet Calculation

Markku Tukiainen
Department of Computer Science
University of Joensuu
Markku.Tukiainen@joensuu.fi

Keywords: POP-III.C. *All Cognitive Dimensions* POP-III.B. *Spreadsheets* POP-V.B. *Questionnaire Analysis*

Abstract

The Cognitive Dimensions of Notations framework (CDs) in the form of questionnaire has been used for evaluating the usability of different programming tools and systems (Kadoda et al. 1999, Cox 2000, Kadoda 2000). Green and Blackwell (2000) continued on this road by developing a CDs questionnaire that presented all the CDs in general terms. We adopted their questionnaire, translated it into Finnish and administered it to ten spreadsheet users in order to evaluate the general CDs questionnaire approach and to find out the usability problems in spreadsheet systems. Overall, the most of the Cognitive Dimensions were interpreted quite easily but some were more difficult than others. We also give some results concerning the Cognitive Dimensions of spreadsheet calculation.

Introduction

The Cognitive Dimensions of Notations framework (CDs) in the form of questionnaire has been used for evaluating the usability of different programming tools and systems (Kadoda et al. 1999, Cox 2000, Kadoda 2000). The questionnaires used have tailored the CDs questions to specific systems. This has demanded quite a lot of work on behalf of the experimenters designing the questionnaires. Last year at PPIG Blackwell and Green (2000) proposed a generalised questionnaire in which the definitions of the CDs themselves are offered to the users, and respondents can choose for themselves the features of the system they wish to criticise. They evaluated the questionnaire with different programming tools and music typesetting programs. We have adopted their questionnaire and evaluated their approach with Finnish spreadsheet users.

Most of the evaluations used in HCI, e.g. Heuristic Evaluation (Nielsen & Molich 1990) or Formal Usability Inspection Method (Kahn & Prail 1994), have traditionally been performed by the CHI experts or the designers. The end-users have relied on the advertisements and try-outs found on the popular magazines. There have not been many attempts to deliver usability evaluation tools that could be used by the end-users (Ravden & Johnson 1989, Tukiainen & Lempinen 1994).

CDs was originally proposed as a discussion tool with which to discuss the usability tradeoffs that occur in different programming environments (Green 1989, Green & Petre 1996). The evaluator was conceived as a person who would understand the framework and the system to be evaluated (or designed) well. Green (1989) has expressed devout hopes that CDs would provide an addition to the usability vocabulary that the end-users could appreciate while evaluating software systems. Kadoda et al. (1999) took a step toward this goal by devising a CDs questionnaire approach in which the evaluation was performed by system users. Their questionnaire presented a system tailored subset of CDs they thought were the most relevant to the system they evaluated. Green and Blackwell (2000) continued on this road by developing a questionnaire which presented all the CDs in general terms, and let the users decide which ones are the relevant.

We adopted their questionnaire, translated it into Finnish and administered it to ten spreadsheet users in order to evaluate the general CDs questionnaire approach and to find out the usability problems in spreadsheet systems. Green and Blackwell (2000) point out the plus side of the general questionnaire as (i) the users do all the work, (ii) the data only reflects their opinions, and (iii) the same questionnaire can be used for all information artefacts. On the down side, the questionnaire is longer (because of all CDs) and possibly harder to understand (because CDs are presented in general terms).

Cognitive Dimensions Questionnaire Study

We first translated the questionnaire into Finnish. We followed the original Blackwell and Green questionnaire (Blackwell & Green 2000) as closely as possible. The only difference we made was to add some background questions to the Section 1 – Background Information. We asked the subjects how many spreadsheet applications they had done and what was the size (in cells) of the largest application they had done.

The translation was fairly easy. We did have some problems in trying to find the most general terms without losing the original meaning but at the end we were quite satisfied with the result. We tested the questionnaire with a few researchers and students at the department of Computer Science who did not have any background knowledge of CDs and evaluated the translation by letting them to describe how they understood the questions.

We appended the questionnaire with one page containing 20 questions. The questions asked how easy or difficult it was to understand what was meant by the questions in the questionnaire. The answers were given in Likert scale with 1=very easy, 2=fairly easy, 3=neutral, 4=fairly difficult and 5=very difficult. First four questions dealt with explanations for product, notation, helper device and redefinition device. Then followed 14 questions of CDs, and the last two were for unusual usage of the system and improvement of the system.

Method

We asked ten volunteers to fill out the questionnaire at their own pace without supervision. None of the 10 subjects in the study had any degree of prior familiarity with CDs. The only requirement was that the subject had used spreadsheet calculation system for some time. Five of the subjects were teachers at the department of Computer Science and they all had taught some spreadsheet calculation. Another 5 subjects were Professional Upgrade program students at the department. They all had used spreadsheet calculation for at least six months and implemented at least ten applications.

The teacher group had used spreadsheet calculation on an average of 8.8 years and the student group on an average of 3.1 years. We also asked the general programming experience of the subjects. The teacher group had been programming on an average of 11.0 years and the student group 1.8 years.

Results

There was a difference in the activity profiles of the groups, see table 1. The students spent the most of their time transcribing data from other sources (mean 34.0 % of their use time) and restructuring the spreadsheet (mean 25.0 %). The teachers spent the most of their time adding things to their spreadsheets (mean 34.0 % of their use time).

Activity	Students		Teachers	
Search	14.0	(5.5)%	9.0	(4.2)%
Transcribe	34.0	(18.2)%	17.0	(13.0)%
Add	13.0	(6.7)%	34.0	(16.7)%
Restructure	25.0	(16.6)%	20.0	(10.0)%
Explore	16.0	(8.9)%	13.0	(6.7)%

Table 1 – Mean (and Standard Deviation) percentage of time spent on different activities.

One speculative explanation for this could be that the teachers use spreadsheet for bookkeeping purposes, while the students do not.

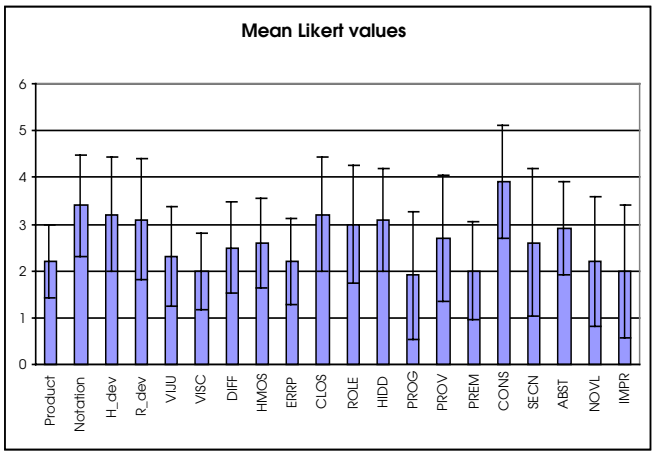

Figure 1 - The Mean Likert-values over all subjects (with Standard Deviations).

Our study found the concepts of notation, helper devices and redefinition devices somewhat problematic. Figure 1 shows the mean Likert-values over all subjects. Two of the teachers said after completing the questionnaire that the term notation was problematic because the word processor example stated that the notation is the way the letter looks on the screen and that in spreadsheets the formulae look like numbers in the sheet most of the time but can be viewed as spreadsheet formula language as well. The subjects did not specify sub-devices despite the word processor example that could have been applied directly to the spreadsheet systems.

Out of fourteen CDs only 3 scored over three in mean Likert-values. They were Closeness of Mapping (mean 3.2), Hidden Dependencies (3.1), and Consistency (3.9). Out these three the Consistency (mean 3.9, standard deviation 1.2) was the hardest question to understand for the subjects. This could be that the general explanation of consistency could not be applied to the spreadsheet systems easily. It could be also that the wording of the question was not the best one. We have changed the wording after the study. Overall, the most of the CDs were interpreted quite easily. For the exact values and the differences between the students and the teachers see Appendix A.

Cognitive Dimensions of Spreadsheet Calculation

Visibility and Juxtaposability

Both groups agreed that the visibility is good in general. The current spreadsheet systems show the content of the cell at the cell position at the cell grid and also at the formula bar while the cell is being edited. When asked what kind of things are more difficult to see or to find, the most of the subjects mention formulae or referencing of the cells in formulae. The question of juxtaposability got divided answers, four of the students said that it is not possible to see two parts at the same time while the teachers said that it is not problematic to see two parts.

Viscosity

Nine out of total 10 subjects answered that changing previous work is easy. Three of them made an explicit comment about changing the contents of a cell. The only subject who doubted the easiness of changes made an example of larger change. It could be argued that the subjects might have thought the changing at the cell level which is easy but overlooked larger changes that could affect for example the layout of the spreadsheet. An example of such small change at the goal but quite a large change at the implementation level could be changing the total sum to a guarded sum (Sajaniemi et

al., 1999). When asked about particularly difficult changes, seven subjects gave positive answers, six of them mentioned larger changes to formula structure at this point.

Diffuseness

All of the subjects agreed that the spreadsheet notation can express what they want reasonably briefly. Four out of five students mentioned commenting formulae as an example to take more space to describe. The teachers did not give examples but one of them mentioned that the operations invoked by menu choices demand more space.

Hard Mental Operations

All the teachers gave formula creation as an example of hard mental operation. The students answered more variable answers, only one of them gave formula creation explicitly but two mentioned the figuring out of what are the correct references in formulae to be hard. All gave some examples of hard operations. When asked about things especially complex to work out in one's head, nine said no and one mentioned physically disperse data to be connected in computation.

Error Proneness

Three subjects mentioned the syntax of formula expressions particularly easy area to make mistakes, another three mentioned making cell references hard and only one said no to this question. When asked about slips, three mentioned mistyping, three mentioned wrong selection from menus and one mentioned the automatic formatting between numeric values and dates.

Closeness of Mapping

Six subjects said that the notation is either closely or very closely related to the result. Three subjects (2 teachers and one student) said that the notation is the formulae and the result is the numbers at the bottom line so they are not closely related.

Role Expressiveness

Eight subjects responded positively to the overall role expressiveness question. They gave explanations like "I use sensible titles on the columns" or "I have a standard way of constructing my applications". Two of the teachers responded negatively and one of them explained: "The problem is that I can see only one formula at the tome so it is hard to see all the roles easily". The answers to the other two role expressiveness questions were all negative stating that there are no especially difficult or non-understandable parts in the spreadsheets.

Hidden Dependencies

Nine subjects answered that there are hidden dependencies at the spreadsheets. Six out of them mentioned formulae at this context. When asked about particularly large descriptions and hidden dependencies eight subjects responded that the effect gets worse. All the teachers and one of the students responded that one has to be very careful with cell references when implementing or updating large spreadsheets. When asked about do the dependencies stay the same, or are there some actions that cause them to get frozen, four replied that the dependencies stay the same, four did not answer the question at all and two responded they did not understand the question.

Progressive Evaluation

All 10 subjects agreed that it is very easy to check progressing at any time. They also agreed that evaluating the amount of work done is easy and possible at any time. Nine subjects answered that they can try out partially-completed versions of the product and one subject responded that in principle it is possible to try out the product but partially-completed product does not offer the same capabilities as the final product.

Provisionality

Seven subjects responded positively to provisionality question. As the features of notation that help at this, they mentioned the direct manipulation nature of the interface, easy to use cut-copy-paste operations, and using sub-totals in calculations.

Premature Commitment

Nine subjects responded that there are no premature commitments in spreadsheet calculation. One student thought that the input of the data must be done before specifying the computations. Two teachers added that the system demands no premature commitments but the task usually sets some order into the process.

Consistency

Consistency was a hard question for the subjects. Six of the subjects did not answer at all to this question. Out of four subjects that did answer this question, two said that the notation is not consistent giving no examples what they meant by that and two said that the notation is consistent giving as an example the way the average is computed and saving the application with different user interface operations.

Secondary Notation

Three subjects answered that there is no possibility to use secondary notation in spreadsheets. Five said that it is possible and two did not answer at all. When asked what they would do if the notation was on the paper, five answered that they would depict cell references between formulae using lines, arrows etc. One student answered that she would color different parts of the application with different colors. When asked do they ever add extra marks etc. to the notation, six replied positively giving extra textual cell contents as an example.

Abstraction Management

Seven subjects said that it is possible to define new facilities or terms within the notation, three of them giving macros as an example. When asked does the system insist defining new terms before anything else, eight subjects answered no and two did not answer anything.

Conclusions

This preliminary study shows that the generalised Cognitive Dimensions questionnaire is applicable in Finnish, for spreadsheet calculation systems, and for Finnish university teachers and students. Although the size of the subject population was small, the results were quite coherent.

One impression we had after this study was that maybe it would be beneficial to change the example given at the generalised CDs questionnaire into the specific example taken from the system under the study, e.g. in our study the spreadsheet system. This would not demand too much work on behalf of experimenters but would probably guide the subjects to better understanding of the concepts of product, notation and sub-devices.

Another observation of the questionnaire was that in a case of notation system like spreadsheet calculation it is possible that the subjects fix they analysis at a certain level of the notation. For example in our study it seemed that a lot of the time the subjects thought about working at a local, cell level (e.g. changing or writing a formula) and not so much at a level of larger structures or at a level of an application. This could also be because the current spreadsheet systems do not support well working with larger structures than one cell. For a discussion of other approaches see e.g. Sajaniemi et al (1999).

References

Blackwell, A.F. & Green, T.R.G. (2000). A Cognitive Dimensions Questionnaire Optimised for Users. In A.F. Blackwell & E. Bilotta (Eds.) *Proceedings of Twelfth Annual Meeting of the Psychology of Programming Interest Group (PPIG-12),* 1(2), 137-152.

Cox, K. (2000). Cognitive Dimensions of Use Cases – Feedback From a Student Questionnaire. In A.F. Blackwell & E. Bilotta (Eds.) *Proceedings of Twelfth Annual Meeting of the Psychology of Programming Interest Group (PPIG-12),* 99-121.

Green, T.R.G. (1989). Cognitive Dimensions of Notations. In A. Sutcliffe & L. Macaulay (Eds.), *People and Computers V.* Cambridge University Press, 443-460.

Green, T.R.G. & Petre, M. (1996). Usability Analysis of Visual Programming Environments: A 'Cognitive Dimensions' Approach. *Journal of Visual Languages and Computing,* 7, 131-174.

Kadoda, G. (2000). A Cognitive Dimensions View of the Differences Between Designers and Users of Theorem Proving Assistants. In A.F. Blackwell & E. Bilotta (Eds.) *Proceedings of Twelfth Annual Meeting of the Psychology of Programming Interest Group (PPIG-12),* 123-135.

Kadoda, G., Stone, R. & Diaper, D. (1999). Desirable Features of Educational Theorem Provers - A Cognitive Dimensions Viewpoint. In T.R.G. Green, R. Abdullah & P. Brna (Eds.) *Collected Papers of the 11th Annual Workshop of the Psychology of Programming Interest Group (PPIG-11),* 18-23.

Kahn, M.J. & Prail, A. (1994). Formal Usability Inspections. In Nielsen, J. & Mack, R.L. (eds.) *Usability Inspection Methods,* 141-171.

Nielsen, J. & Molich, R. (1990). Heuristic Evaluation of User Interfaces. *Proceedings of ACM CHI'90 Conference (Seatle, WA, April 1-5),* 249-256.

Ravden, S. & Johnson, G. (1989) Evaluating Usability of Human-Computer Interfaces: a practical method. *Ellis Horwood Limited, Chichester.*

Sajaniemi, J., Tukiainen, M., Väisänen, J. (1999). Goals and Plans in Spreadsheet Calculation. University of Joensuu, Department of Computer Science, Technical Report, Series A, Report A-1999-1.

Tukiainen, M., Lempinen, R. (1994). Muokattu Ravdenin ja Johnsonin menetelmä ohjelmiston käytettävyyden arviointiin (Modified Ravden & Johnson method for evaluating the usability of a computer system). University of Joensuu,, Department of Computer Science, Tech. Rep. B-1994-1.

Appendix A:

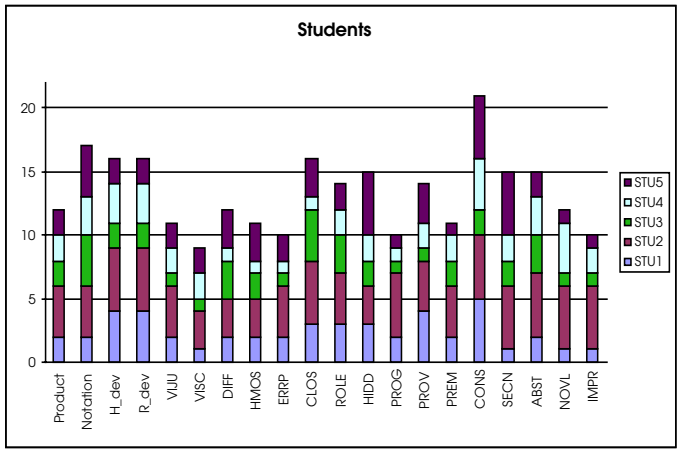

Figure 2 – Likert-scale values for the students group

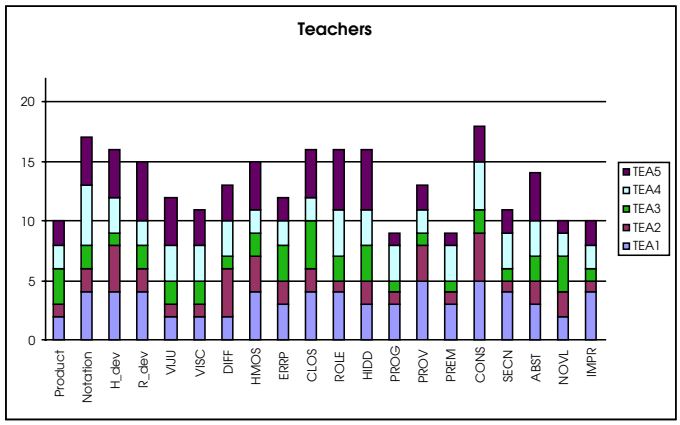

Figure 3 - Likert-scale values for the teachers group