FQS FORUM: QUALITATIVE SOCIAL RESEARCH SOZIALFORSCHUNG

Volume 6, No. 2, Art. 29
May 2005

Application of Qualitative Content Analysis in User-Program Interaction Research

Christian Spannagel, Michaela Gläser-Zikuda & Ulrik Schroeder

Key words: qualitative content analysis, user behavior, user-program interaction

Abstract: This article describes how techniques of qualitative content analysis (QCA) can be applied to analyze user actions. Criteria for a software system supporting QCA of user-program interactions are systematically developed. A software tool is presented which allows recording and replaying user actions. In addition, it supports the application of QCA on recorded action sequences. Two approaches of category specification in analysis are discussed—inductive category development and deductive category application—which may be transferred to the area of user action analysis. The procedure for the analysis of user behavior and consequences for further development of the software system are discussed.

Table of Contents

1. Introduction

In many of today's empirical studies computers are used as a tool for content presentation and data acquisition. The analysis of the interaction between user and software plays an important role in these studies:

- In usability tests, the analysis of the working process provides information about the usability in addition to interviews and questionnaires.
- In psychological experiments, software is often used to present stimuli and to achieve exact data (e.g. reaction times, error rates).

© 2005 FQS http://www.qualitative-research.net/fqs/

- In evaluations of instructional software, it is often not sufficient to analyze just the learning outcomes. In order to gain information about learning processes and learning strategies the focus must be placed on the user behavior, as well. [1]

To obtain data about the process of software usage, researchers may *observe* this process (*data acquisition*). One method is to sit next to the participants observing their behavior and to keep observation minutes. A significant disadvantage of this approach is that participants knowing that they are observed show a different behavior. Furthermore, it is impossible to observe user behavior at a later time, e.g. for the purpose of controlling the coded data (SWEENEY, MAGUIRE & SHACKEL 1993). To avoid these disadvantages, behavior should be made persistent, for example by video. Videotaped data allow the behavior to be observed many times and by many observers. Because of these considerations video based observation becomes more and more important in research on learning and instruction (AUFSCHNAITER & WELZEL 2001; PRENZEL, DUIT, EULER, LEHRKE & SEIDEL 2001; JACOBS, KAWANAKA & STIGLER 1999). [2]

To study participants' behavior using software, two types of behavior may be observed: Off-line and on-line behavior. *Off-line behavior* refers to the actions in front of the screen like mimics and gestures. *On-line behavior* refers to actions on the screen like mouse motions, menu selections and keyboard inputs (HILBERT & REDMILES 2000). For the observation of on-line behavior, computer programs can be used which, for example, store mouse and keyboard actions in log files. System monitoring of user actions is non-intrusive (SWEENEY, MAGUIRE & SHACKEL 1993) and thus, the influence on the participants by observation is zero. [3]

Once data acquisition is completed, the compiled data set has to be analyzed (*data analysis*). Depending on the research question and purpose, more qualitative or more quantitative methods may be applied. For instance, qualitative methods are applied for the classification of data with respect to categories, whereas quantitative methods allow the calculation of parameters and the application of statistical tests. [4]

In this article we focus on qualitative-oriented analysis of on-line user behavior supported by a computer system. The aim of the analysis is to categorize behavioral sequences based on a systematical approach. For this purpose, qualitative content analysis (MAYRING 2000a) seems to be an appropriate method. It allows to define categories in an inductive or deductive procedure, which is explained in Section 2. Section 3 describes the interaction between user and program in detail and discusses advantages and disadvantages of representation formats for interaction processes. In Section 4, criteria for a system which supports qualitative content analysis of user behavior are derived. Section 5 describes CleverPHL, a "capture & replay" tool, which allows to perform qualitative analyses on recorded action sequences. Finally, implications for further improvement of the system and steps of the application of qualitative content analysis for the analysis of user behavior are discussed. [5]

© 2005 FQS http://www.qualitative-research.net/fqs/

2. Qualitative Content Analysis

Qualitative content analysis (MAYRING 2000a, 2000b) is a qualitative oriented method that applies different techniques for a systematic analysis, mainly of text material gained e.g. by interviews, diaries, observation protocols, or documents. Furthermore, the analysis of other products such as pictures, video-tapes, radio transmission or songs is possible, as described in numerous studies in the last few years (MAYRING & GLÄSER-ZIKUDA in press). Originally, the method was developed to analyze a huge amount of interviews in a qualitative way. To develop a suitable method of analysis, the advantages of quantitative content analysis developed in communication sciences were preserved, and transferred to qualitative-interpretative steps of analysis by a development of specific procedures and techniques (MAYRING, 2000a). [6]

Qualitative content analysis is not limited to a specific discipline. Meanwhile, it is widely applied in psychology, linguistic, sociology, history, arts etc. (e.g. GLÄSER-ZIKUDA 2001a,b; KRIPPENDORFF 1980; MAYRING 1996; MAYRING, KÖNIG, BIRK, & HURST 2000; MAYRING & GLÄSER-ZIKUDA in press; RUST 1983). [7]

As these empirical studies indicate, qualitative content analysis represents an empirical approach of a methodologically controlled analysis of acoustic, verbal or visual contents within their context of communication, following content analytical rules based on a step by step model, that also allows different kinds of quantifications (MAYRING 2000a, b). [8]

2.1 Procedures and techniques of qualitative content analysis

Qualitative content analysis is embedded in a model of communication. The focus of analysis is related to aspects of the communicator (to individual experiences, thoughts or feelings), to the situation of the production (e.g. interview situation), to the socio-cultural background, to the material itself or to the message of the material. Furthermore, qualitative content analysis represents a rule guided method. The analysis of the material follows a step by step and rule guided procedure dividing the material into content analytical units. Thus, the central analytical units are categories. Following the research question, categories are developed based on specific theoretical aspects. By feedback loops and revisions the conformity of the categories with respect to theory and analytical procedure is ensured (MAYRING 2000a,b). Finally, central empirical criteria of quality, such as reliability and validity are of importance. On the one hand, reliability is proven to ensure to which extent the procedure of analysis is inter-subjectively comprehensible. To check the inter-coder reliability parts of the material are coded by at least two researchers. Generally, the inter-coder reliability is accepted as being sufficient if, for example, KRIPPENDORFF's alpha is higher than .70 (KRIPPENDORFF 1980). On the other hand, validity is a relevant criterion to compare the results gained in the present study with those of other

© 2005 FQS http://www.qualitative-research.net/fqs/

studies in the sense of triangulation (FLICK 2000).[1] To allow a flexible and multi-level analysis, a variety of procedures was developed in qualitative content analysis. Two procedures may be characterized as being central for analysis: inductive category development and deductive category application. [9]

2.2 Inductive category development and the technique of summarization

It is a crucial question how the central analytical units, the categories, are developed. In qualitative content analysis an inductive category procedure is applied in orientation to the strategies of reduction in the psychology of text assimilation (BALLSTAEDT, MANDL, SCHNOTZ & TERGAN 1981). [10]

The main idea of the inductive category procedure is to define in orientation to the theoretical background and research question which parts of the text material are relevant for inductive categorization (MAYRING 2000b). For instance, in a study focusing on learning emotions, the general definition for the inductive category development was: "All specific emotional states of the learner in relation to learning processes" (GLÄSER-ZIKUDA 2001a, b). Following this criterion, the material is analyzed and categories are tentatively developed based on a step by step model. [11]

The analytical technique of summarization is most often applied in qualitative content analysis. It allows for creating inductive categories by reducing, paraphrasing and generalizing relevant text passages (MAYRING 2000 a,b). These categories may be reduced to main categories. If necessary, the general category definition is revised, as well as the main categories for a formative improvement of the inter-coder reliability. Finally, the whole material is analyzed in accordance to the inductive model, and a summative inter-coder reliability is conducted. Furthermore, the inductive category development allows for quantifications (e.g. frequencies of categories and main categories). [12]

2.3 Deductive category application and the technique of structurization

The aim of the deductive procedure in qualitative content analysis is the application of categories based on the theoretical considerations of the study. The analytical procedure follows a methodologically controlled application of the category to the material (MAYRING 2000a, b). The analysis aims at giving explicit definitions, examples and coding rules for each deductive category, determining exactly under which conditions material may be coded with respect to a category. The analytical technique generally applied in the deductive category procedure is structurization. This technique aims at the analysis of the specific structure of the material, at a standardization of a variable, or finally at a rating of specific theoretical aspects. For instance, in the study mentioned above, categories for learning emotions were defined with respect to emotion theory (GLÄSER-ZIKUDA

1 The approach described in this article may not be characterized as totally qualitative. Also quantitative criteria play a certain role. Although it is important to take subjectivity into account—particularly in qualitative approaches—a main purpose of the work described here is to control intersubjectivity between observers.

© 2005 FQS http://www.qualitative-research.net/fqs/

2001a,b). The variables were scaled as three values (much—some—no), explicit definitions, anchor examples and rules for the distinction of categories were given (see Table 1).

		CODING AGENDA	
Variable	**Definition**	**Examples**	**Coding rules**
much anxiety	Strong anxious feelings referring the subjective • feelings of worry • appraisal of thread • excitement/stress in learning situations	"The teacher told us that the test will turn out bad" "I was not able to answer most of the questions" "I have to study very much to get a good grade"	Overall impression from interview data and in addition more than 80% of diary entries supporting much anxiety. Concrete statements for feeling anxiety are reported.
some anxiety	Some aspects of anxiety or only moderate anxious feelings	"I didn't know that" "I had many mistakes"	Impression from interview data and in addition less than 80% of diary entries pointing to some anxiety
no anxiety	Abolutely no anxiety	"Nothing, I understood all" "Should I be anxious about anything?"	Concrete statements declaring having no anxiety concerning the learning (interview data and diary entries, 20% rule)

Table 1: Example for a Coding Agenda for the Deductive Category Application (GLÄSER-ZIKUDA 2001a, b) [13]

As qualitative content analysis offers these various possibilities of analysis, we argue in the following that these specific procedures may be successfully applied to the analysis of user behavior. [14]

3. User-Program Interactions and the Format of Process Storage

As described in Section 1, on-line behavior is one kind of user behavior and describes all actions performed by the user like mouse motions, mouse clicks and keyboard input. Actions performed by the program may be the presentation of a window asking for user input or the presentation of data. The mutual exchange of actions between user and program is called *user-program interaction*. [15]

For analysis, it is necessary to *record* user-program interactions and to *store* them. Therefore, computer-based methods should be used for many reasons. First, the computer minimizes the influence of observation on subjects using the software. Second, data are gathered by the same method for all subjects. Therefore, the collection of data is not influenced by subjective decisions of

© 2005 FQS http://www.qualitative-research.net/fqs/

different observers. Third, after data acquisition, data sets are available to be automatically analyzed with computational methods. [16]

There are two classic ways of storing information about software usage: log files and screenshot sequences. Both methods have advantages and disadvantages. In log files, important events are represented as textual descriptions and stored in a linear list. Those event lists can easily be analyzed later on. Disadvantageous is the fact that information stored in log files is often just a very coarse extract of user-program interaction (low resolution). Many details are lost which may become important during the analysis of the user behavior. For example, in many programs there are various ways to achieve one goal, e.g. loading a file in a text processor by choosing an item in the file menu, by clicking on a specific icon, or by using a keyboard shortcut. A log file would only document which file was loaded, but not how this goal was achieved. But this might be of interest in a usability test. A second disadvantage is that the interaction cannot be watched later on. Furthermore, log mechanisms are often implemented within the software and cannot be changed by experimenters to accomplish special demands. [17]

On the other hand, taking screenshot sequences is an image-based way of representing on-line behavior. This method allows for observation of the interaction afterwards in analogy to a video. Unfortunately, it is practically impossible to automatically analyze behavior which is stored in this kind of format. Thus, it has to be coded by human observers, which can be very time-consuming, especially when many participants take part in a study producing hundreds of hours of user behavior. [18]

The representation format of interaction processes should not limit the possibilities during the qualitative content analysis. Therefore, both formats described above are not sufficient for this task. Behavior stored in log files cannot be observed, and interactions captured in a sequence of screenshots cannot be analyzed with computational methods. In the next section, criteria for a format avoiding both disadvantages are derived. In addition, criteria for a software system which supports qualitative content analyses with computational methods are specified. [19]

4. Criteria for a System Which Supports QCA of User-Program Interactions

Based on the considerations described above, the following criteria for a system can be derived which allows the recording and analysis of user behavior with qualitative content analysis[2]:

- *Recording functionality:* The system captures the interaction between user and program and represents it in a format, which is suitable for qualitative content analysis.

2 Some of the criteria have already been described in KLAUDT and SPANNAGEL (2004), and in SCHROEDER and SPANNAGEL (2004), in both cases for different purposes.

© 2005 FQS http://www.qualitative-research.net/fqs/

- *Category definition:* The system supports the definition of categories based on specific theoretical aspects. Inductive category development and deductive category application are supported.
- *Symbolic representation:* The interaction process is represented symbolically as in log files. Thus, on-line behavior can be automatically summarized and structured regarding to the previously defined categories.
- *High resolution:* Especially for inductive category development, the types of elements which belong to a category are not known until the analysis takes place. Therefore, filtering of information during data acquisition is minimized.
- *Playback of interactions:* The definition of categories and the automatic categorization of user-program interactions can be checked by selectively replaying them. This allows researchers to redefine categories in the light of observed user behavior.
- *Support of recoding:* The system allows recoding of the same data after category definitions have been revised.
- *Support of inter-coder reliability checks:* If at least two coders have analyzed the data, the system offers methods to perform inter-coder reliability checks.
- *No need to adapt programs:* Logging, analyzing and replaying user actions is possible without the need of adaptation to the program the user interacted with.
- *Quantitative analysis:* Based on categorized user-program interactions, quantitative analyses can be performed. [20]

Summing up, the representation format of interaction processes combines the advantages of log files and screenshot sequences. On-line behavior can be summarized and structured automatically, and it can be observed repeatedly to check the categorization conducted by the system. And finally, the system allows a revision of the category definitions, if required. [21]

5. Qualitative Content Analysis with CleverPHL

In the following, we describe a software tool which meets most of the criteria described in the previous section. It's acronym *Clever* is formed by the letters of its features (capture, log, edit, visualize, evaluate, and replay user actions). The software originally was developed as a software engineering tool for the purpose of capturing and replaying interactions with prototypes in order to visually specify system functionality of software to be developed (SCHROEDER 2000). Then it was enhanced as a tool supporting automatic tests of system functionality and it was applied to instructional scenarios. CleverPHL is part of the Jacareto capture & replay toolkit. Besides its analysis tools, it was designed to support functionality tests of software and to implement action-oriented concepts and constructivist models like the cognitive apprenticeship model in E-Learning scenarios (SCHROEDER & SPANNAGEL 2003; SCHROEDER & SPANNAGEL 2004). The toolkit is an open source project and can be downloaded from http://jacareto.sourceforge.net/. [22]

5.1 Description of the system

CleverPHL has been developed inter alia to perform qualitative analyses on user-program interactions. The system combines tools for data acquisition and methods for data analysis. Specific techniques allow the logging of actions (also called *events*) on programs (*recording functionality*) and to store these interactions in an analyzable format (*symbolic representation*). This format is called *record of user-program interaction*, or just *interaction record*. The program the user interacts with is called *target application*. A target application must be written in Java and can be used together with CleverPHL without having access to its source code (*no need to adapt programs*). Actions are stored chronologically, combined with time information. The linear sequence of actions can be viewed in CleverPHL, and attributes of actions can be inspected (see Fig. 1). All events occurring during the process of software usage are captured and stored; no information is filtered out while acquiring the data. That means that every mouse motion, mouse click, keyboard input etc. is saved (*high resolution*).

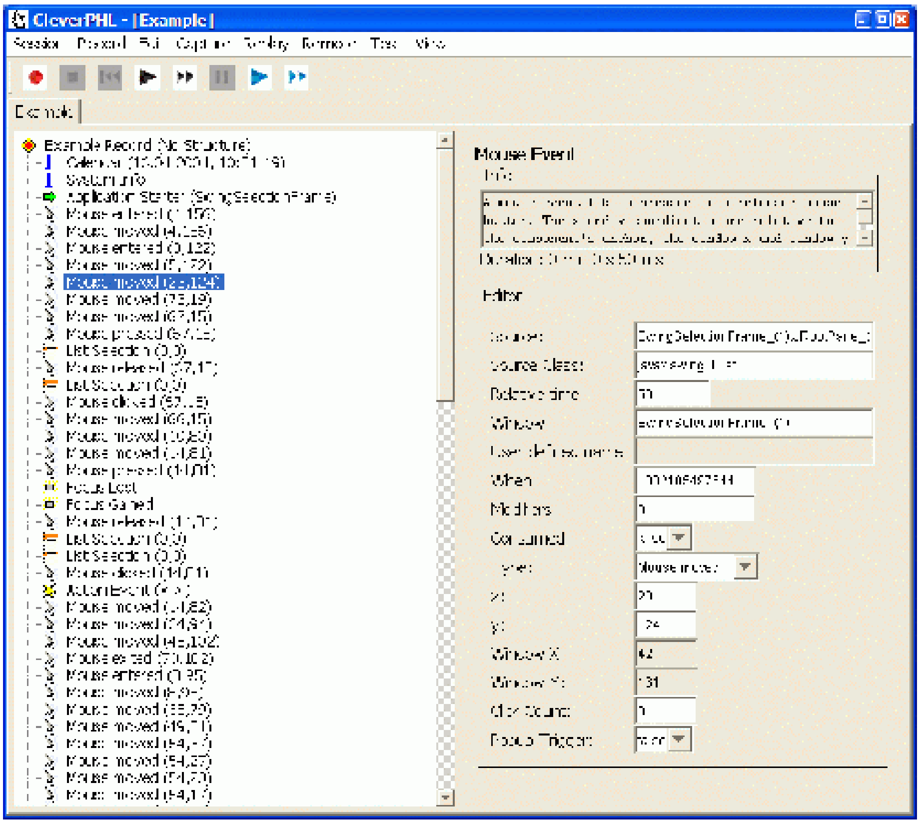


Fig. 1: The linear interaction record on the left, and the editor for a selected record element on the right. Please click here for an increased version of Figure 1. [23]

The user-program interaction contained in a given interaction record can be replayed in order to observe and evaluate it (*playback of interactions*). For this purpose, the target application is started anew, and all stored mouse and key events are dispatched to it. As a result, the mouse pointer is automatically moved atop the target application, and mouse clicks and keyboard inputs are performed

© 2005 FQS http://www.qualitative-research.net/fqs/

by the system instead of a human user. No screenshots or images are stored in addition to the symbolic interaction record. The visual impression of the behavior is reconstructed from the interaction record and replayed on a real instance of the target application so that it can be observed repeatedly. During the replay process, the location of the important actions can be highlighted in order to attract the observer's attention (Fig. 2).

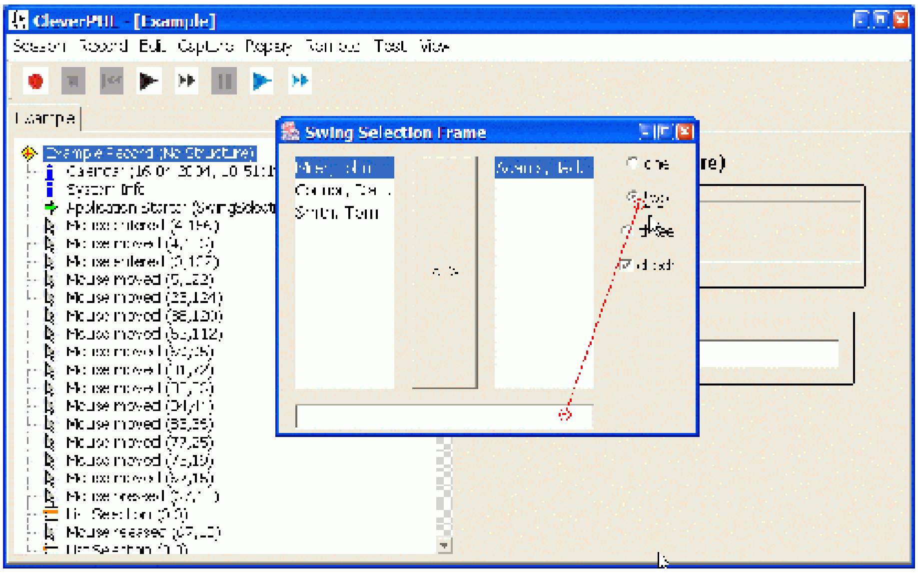


Fig. 2: Replaying actions on the target application. The location of the next important action can be highlighted (red line). Please click here for an increased version of Figure 2. [24]

Because the interaction record is represented in a symbolic format, qualitative analyses can be performed with computational methods. CleverPHL has a mechanism, which enables the classification of user-program-interactions regarding to defined categories. Based on a system of categories, CleverPHL bundles events as a category element. Categories may be defined on different levels, thus forming a hierarchy: category elements of a lower level may be part of a category element of a higher level, and so on. For example, a category element "*actions on a print dialog*" may consist of category elements "*format chosen*," "*page numbers chosen*," and "*print started*," which may consist of sub-elements such as mouse and keyboard actions. [25]

Categories are defined by programming detection algorithms for specific action sequences (*category definition*). The category system consists of a bundle of detection algorithms, and parts of the linear interaction record are classified by the interplay of those algorithms. The result is a hierarchical view of the user-program interaction (Fig. 3). The process of categorizing parts of the interaction record is based on methods derived from compiler design (AHO, SETHI & ULLMAN 1985). It is similar to the creation of a syntax tree regarding to a grammar of a formal language (SCHROEDER & SPANNAGEL 2003).

© 2005 FQS http://www.qualitative-research.net/fqs/

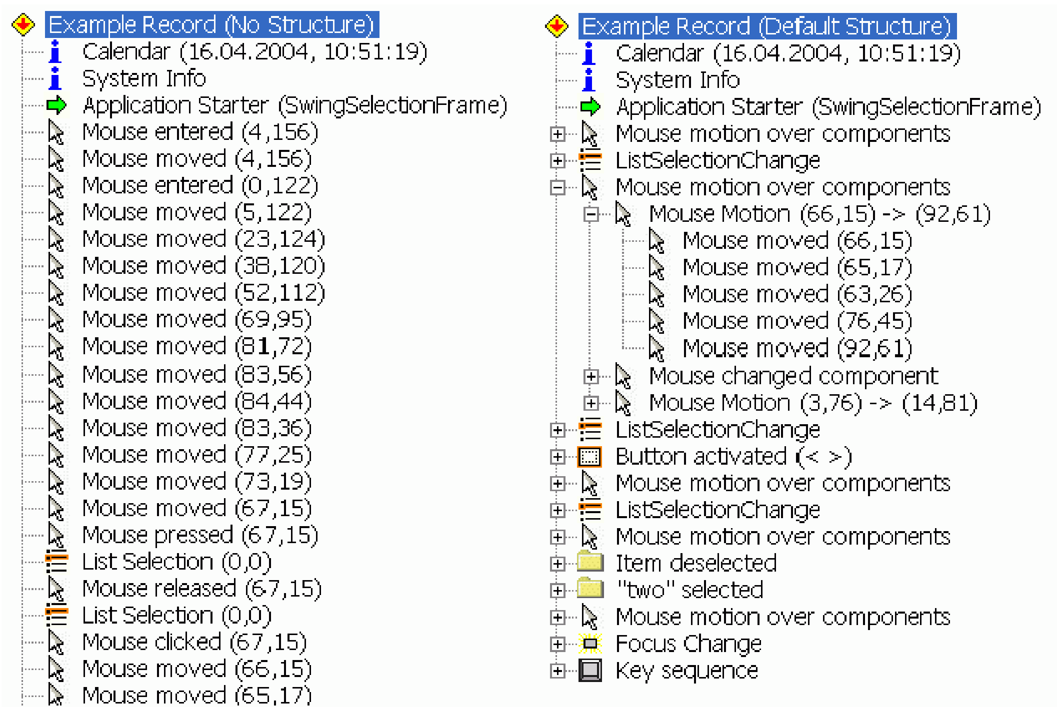


Fig. 3: The linear interaction record on the left (just the first part), and the same record after categorization on the right [26]

If a new category of a specific user behavior has to be integrated, a new detection algorithm for the action sequence must be implemented: Then, captured data can be recoded using the new set of category definitions (support of recoding). Because recoding is done automatically, it can be performed with almost no effort after categories have been revised or new categories have been included. [27]

Once an interaction record has been analyzed, it is possible to replay only one behavioral element found. For instance, if an element *actions on a print dialog* has been found, just the part of the record belonging to this element can be replayed. Thus the view of the analyzed record (see Fig. 3) serves as index to video sequences of the behavior (SPANNAGEL 2003). [28]

Crucial to the process of categorization is the definition of categories. Systematic procedures which are part of the qualitative content analysis can be applied for this purpose. In the following two sections, the inductive and deductive procedures are described that define behavioral categories. [29]

5.2 Inductive category development with CleverPHL

Inductive category development with CleverPHL refers to the implementation and improvement of detection algorithms based on given user-program interactions. First, it has to be decided which actions or action sequences are theoretically relevant for the categorization. For instance, actions performed on a special window could be the only relevant actions. Those conditions can then be implemented in the detection algorithm to separate relevant and irrelevant record elements. Rules can be specified in the detection algorithm to define which action sequences belong to a category definition and which do not. Those sequences can be taken from given records of user-program interactions, and the category definition can be validated with those records. For instance, if a given record contains a sequence *format chosen, page numbers chosen,* and *print started*, a detection algorithm can be implemented which categorizes this sequence as *actions on a print dialog*. The categorization can then be validated by performing the algorithm on the given record and by evaluating whether the action sequence mentioned above is categorized correctly or not. As a consequence, detection algorithms can be improved incrementally. Main categories can be derived based on given inductive categories by combining detection algorithms to one single algorithm. Intercoder reliability can be proven if at least two coders have defined detection algorithms. Highest intercoder reliability is found when the two algorithms are semantically equivalent. This is the case if the algorithms detect exactly the same action sequences. After the entire set of information has been categorized by the implemented algorithms, data sets can be extracted to perform quantitative analyses. [30]

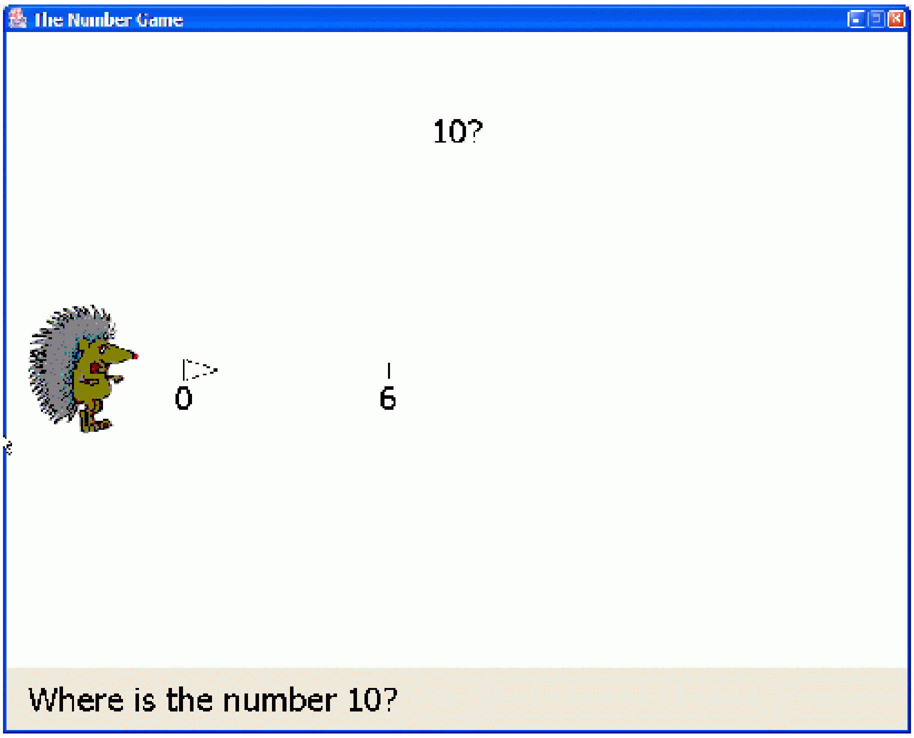
5.3 Deductive category application with CleverPHL

Deductive category application may be applied if, for example, different types of user behavior have been theoretically defined. It may be proven to which extent these types are found in a set of recorded user-program interactions. Detection algorithms can be implemented which categorize records regarding the types of user behavior. Given examples of those types, detection algorithms can be revised and improved in the same way as described above. Again, intercoder reliability can be checked by finding action sequences which are categorized differently by two or more detection algorithms. After the entire set of information has been categorized, again quantitative analyses may be performed after data sets have been extracted from the record. [31]

5.4 Example

In this section, both approaches of defining categories for user-program interactions are illustrated by an example. In the project CEKA, children in primary school worked with software written in LOGO showing number lines (KLAUDT 2003). The children's task in this project was to find a given number (the *target number*) on the number line. The number lines were only labeled at certain points, e.g. the edges. For instance, given the number line from 0 to 6, children had to find the number 10 (Fig. 4). Finding a number was defined as "clicking with the mouse on the correct location." The objective of the project CEKA was to infer the students' mental representations of numbers by observing how they direct the mouse pointer to the correct spot.

Fig. 4: The software used in the CEKA project [32]



Experiences made with the data analysis features of the LOGO system led to the development of CleverPHL's mechanisms to analyze user-program interactions. To demonstrate the possibilities of CleverPHL, the number line software has additionally been translated to Java. Actions performed on the Java version of the program can be stored as interaction records using CleverPHL. As described above, it is possible to replay those records with CleverPHL to get an impression of the recorded actions. [33]

Inductive category development: Given a recorded user-program interaction, its parts can be categorized with the categories included in CleverPHL by default. Those are general categories which can be applied to interactions with any kind of program. But categories for interactions with special software like the number line program have to be developed and implemented first. For example, it is desirable to combine elements which represent a mouse motion with a click or a pause at a position on the number line to one behavioral element (*Number Game Motion and Click resp. Number Game Motion and Pause*), and to combine those elements which belong to a single task (*Number Game Task*). For instance, given the target number 10, a subject has clicked on 5, 7, 12 and then 10. All actions belonging to this sequence were combined as one element of type *Number Game Task*, which consisted of four elements of type *Number Game Motion and Click*. Detection algorithms for these two categories can be defined and added to the CleverPHL categorization engine. The definition of the algorithms is an incremental process where categories can be refined step by step. The detection algorithms, implemented so far, can be validated by categorizing given interaction records with them. If parts of the records are not categorized correctly, the algorithms can be improved or enhanced. The result is shown in Fig. 5.

© 2005 FQS http://www.qualitative-research.net/fqs/

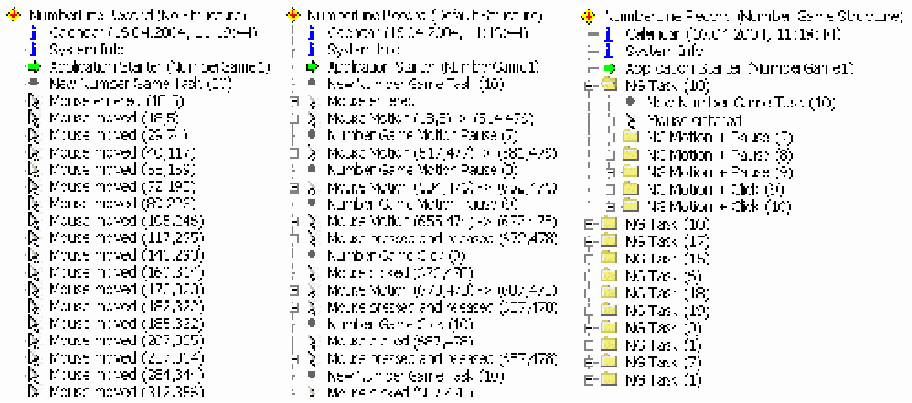


Fig. 5: The linear record on the left, the default categorization in the middle, and the CEKA specific categorization on the right. Please click here for an increased version of Figure 5. [34]

Thus the material is reduced by grouping actions together to build a higher-order action and by abstracting from detailed actions like mouse motions. The result is a concise view of the material which represents the relevant information. It should be emphasized that details are not lost by the process of summarization. Lower-order elements can still be inspected by clicking on the higher-order element in the tree view (see Fig. 5). [35]

Deductive category application: Based on these categorized user-program interactions, strategies of how children work with a number line can be found. For instance, some children always count stepwise from one end of the number line to the given target number. Others use more complex strategies like using multiplication and subtraction to find a number. In orientation to a theory of strategy types, those types can be defined as detection algorithms. Those detection algorithms can be implemented using the categories defined before (*Number Game Task*, for example). Given a single record of user-program interactions which has been categorized with the new detection algorithms, it can then be counted how often each strategy type occurs in this record. Given a set of records, a data set can be extracted which contains data about how often each strategy type occurs in all records (Fig. 6).

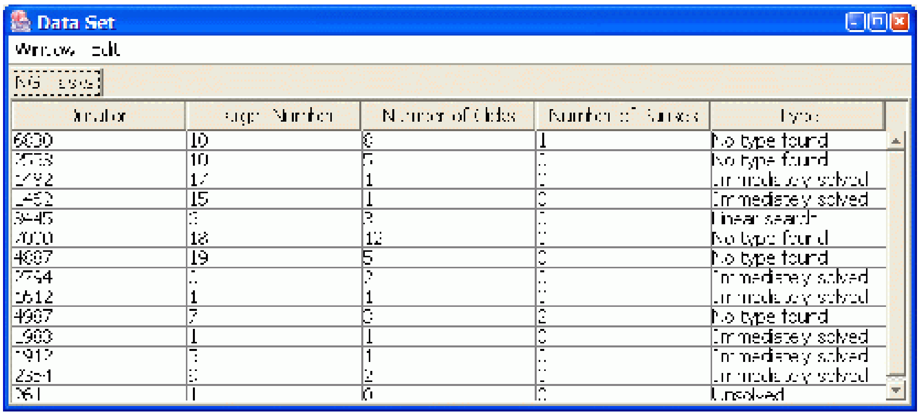


Fig. 6: A data set can be extracted from categorized interaction records. Please click here for an increased version of Figure 6. [36]

6. Discussion

Qualitative content analysis is in many respects fruitful for the research on user behavior. First, it offers a variety of analytical techniques, like for example the technique of summarization and structurization that may be applied in research on user behavior. Second, the procedures of inductive category development and deductive category application offer the possibility to analyze user-program interactions in a theoretically guided, adaptive, rule-guided but as well flexible way. Third, based on these procedures and techniques of qualitative content analysis, quantifications with respect to frequencies, and rankings of specific aspects of user behavior, or correlations of these aspects with further variables are possible. Finally, due to standards of quality, qualitative content analysis allows by the rule-guided procedures to document analytical steps and to verify reliability checks. [37]

The effort for analyzing a huge amount of material with QCA is immense. Using the computer for analysis in the way presented in this paper saves time, because the recorded interactions need not be observed personally and categorized manually. Automatic categorizations of user behavior, defined by a set of detection algorithms, have great advantages. First, when a category definition is revised and improved at a later point in the analysis process, all records can be recategorized with just a few clicks. No material has to be observed in time-consuming ways. Second, the category definition is very explicit. It is formally defined as an algorithm and can easily be evaluated by others. [38]

Beside these advantages, CleverPHL does not yet satisfy all criteria described above. Some features must be improved, e.g.:

- *Category definition:* Defining categories by implementing detection algorithms is a difficult and error-prone process. It would be preferable to give CleverPHL some examples of user behavior and to prompt CleverPHL to learn a rule to automatically assign these examples to a category. One possibility of achieving this goal is applying methods from *machine learning* to learn classification rules for the categories from user supplied examples.

- *Support of intercoder reliability checks:* When two or more coders have implemented detection algorithms, CleverPHL should offer methods to automatically check those algorithms for equivalence. Results from computation theory show that the equivalence problem is decidable only for a subset of all formal languages. It should be researched in which cases intercoder reliability checks can be performed automatically, and methods which evaluate reliability should be implemented in CleverPHL. [39]

Acknowledgments

We would like to thank Markus BREITENBACH, University of Colorado, Boulder, for helpful suggestions regarding machine learning, and Markus VOGEL, University of Education Ludwigsburg, for proofreading.

References

Aho, Alfred V.; Sethi, Ravi & Ullman, Jeffrey D. (1985). *Compilers: Principles, techniques and tools*. Reading, MA: Addison-Wesley.

Aufschnaiter, Stefan von & Welzel, Manuela (Eds.). *Nutzung von Videodaten zur Untersuchung von Lehr-Lern-Prozessen*. *Aktuelle Methoden empirischer pädagogischer Forschung*. Münster: Waxmann.

Ballstaedt, Steffen-Peter; Mandl, Heinz; Schnotz, Wolfgang & Tergan, Sigmar-Olaf (1981). *Texte verstehen, Texte gestalten*. München: Urban & Schwarzenberg.

Flick, Uwe (2000). Trinagulation in der qualitativen Forschung. In Uwe Flick, Ernst v. Kardorff & Ines Steinke (Eds.), *Qualitative Forschung. Ein Handbuch* (pp.309-318). Reinbek: Rowohlt.

Gläser-Zikuda, Michaela (2001a). *Emotionen und Lernstrategien in der Schule.* Weinheim: Beltz.

Gläser-Zikuda, Michaela (2001b). Emotions and learning strategies at school—opportunities of qualitative content analysis. In Mechthild Kiegelmann (Ed.), *Qualitative research in psychology* (pp.32-50). Schwangau: Huber.

Hilbert, David M. & Redmiles, David F. (2000). Extracting usability information from user interface events. *ACM Computing Surveys, 32*(4), 384-421.

Jacobs, Jennifer K.; Kawanaka, Takako & Stigler, James W. (1999). Integrating qualitative and quantitative approaches to the analysis of video data on classroom teaching. *International Journal of Educational Research, 31*, 717-724.

Klaudt, Dieter (2003). Exploring number sense—a research project in Primary School using Logo. In Gerald Futschek (Ed.), *EUROLOGO 2003, Be creative ... re-inventing technology on education,* Proceedings of the 9th European Logo Conference, Porto, Portugal (pp.62-68). Coimbra, Portugal: CNOTINFOR.

Klaudt, Dieter & Spannagel, Christian (2004). Computerunterstütztes Operieren am mentalen Zahlenstrahl. *Erziehung und Unterricht – österreichische pädagogische Zeitschrift, 154*(3-4), 246-257.

Krippendorff, Klaus (1980). *Content analysis. An introduction to its methodology*. Beverly Hills: Sage.

Mayring, Philipp (1996). *Einführung in die qualitativer Sozialforschung. Eine Anleitung zu qualitativem Denken* (3rd ed.). Weinheim: Psychologie Verlags Union.

Mayring, Philipp (2000a). *Qualitative Inhaltsanalyse. Grundlagen und Techniken* (7th ed., 1st ed. 1983). Weinheim: Deutscher Studien Verlag.

Mayring, Philipp (2000, June). Qualitative Content Analysis [28 paragraphs]. *Forum Qualitative Sozialforschung / Forum: Qualitative Social Research* [On-line Journal], *1*(2), Art.20. Available at: http://www.qualitative-research.net/fqs-texte/2-00/2-00mayring-e.htm [April 27, 2005].

© 2005 FQS http://www.qualitative-research.net/fqs/

Mayring, Philipp; König, Joachim; Birk, Nils & Hurst, Alfred (2000). *Opfer der Einheit. Eine Studie zur Lehrerarbeitslosigkeit in den neuen Bundesländern*. Opladen; Leske & Budrich.

Mayring, Philipp & Gläser-Zikuda, Michaela (in press). *Die Praxis der Qualitativen Inhaltsanalyse*. Weinheim: Beltz—UTB.

Prenzel, Manfred; Duit, Reinders; Euler, Manfred; Lehrke, Manfred & Seidel, Tina (2001). *Erhebungs- und Auswertungsverfahren des DFG-Projekts "Lehr-Lernprozesse im Physikunterricht – eine Videostudie"*. Kiel: IPN.

Rust, Holger (1983). *Inhaltsanalyse. Die Praxis der indirekten Interaktionsforschung in Psychologie und Psychotherapie*. München: Urban & Schwarzenberg.

Schroeder, Ulrik (2000). PTAH—validation driven software development. *Proceedings of the 5th International Conference on Integrated Design & Process Technology 2000, Dallas, Texas.*

Schroeder, Ulrik & Spannagel, Christian (2003). CleverPHL: Focusing on the learning process by supplementing instructional software with digital recording and evaluation functionality. In Allison Rossett (Ed.), *World Conference on E-Learning in Corp., Govt., Health., & Higher Ed, 3*(1) (pp.2413-2420). Norfolk, VA: AACE.

Schroeder, Ulrik & Spannagel, Christian (2004). The case for action-oriented e-learning. In Lorenzo Cantoni & Catherine McLoughlin (Eds.), *World Conference on Educational Multimedia, Hypermedia and Telecommunications 2004*(1) (pp.2449-2454). Norfolk, VA: AACE.

Spannagel, Christian (2003). Computergestützte Erfassung und Analyse von Benutzerverhalten im Umgang mit Lehr-/Lernsystemen. In Volkhard Nordmeier (Ed.), *Didaktik der Physik. Beiträge zur Frühjahrstagung der DPG – Augsburg 2003* (CD-ROM). Berlin: Lehmanns Media.

Sweeney, Marian; Maguire, Martin & Shackel, Brian (1993). Evaluating user-computer interaction: a framework. *International Journal of Man-Machine-Studies, 38*, 689-711.

Authors

Christian SPANNAGEL, Master of Computer Science (Dipl. Inform.), Ph.D. student at the Pädagogische Hochschule Ludwigsburg, Germany

Research interests: qualitative and quantitative analysis of user behavior, theory-guided development of instructional software, implementation of constructivist and action-oriented models in E-learning scenarios, animated demonstrations in software tutorials

Contact:

Christian Spannagel

Institut für Mathematik und Informatik
Pädagogische Hochschule Ludwigsburg
Reuteallee 46
71634 Ludwigsburg, Germany

E-mail: spannagel@ph-ludwigsburg.de

Dr. Michaela GLÄSER-ZIKUDA, Assistant Professor at the Institute of Educational Science at the Pädagogische Hochschule Ludwigsburg, Germany

Research interests: qualitative and quantitative research methods, learning and instruction (emotions, learning strategies, development of instructional designs, and portfolio assessment)

Contact:

Michaela Gläser-Zikuda

Institut für Erziehungswissenschaft
Pädagogische Hochschule Ludwigsburg
Reuteallee 46
71634 Ludwigsburg, Germany

E-mail: glaeserzikuda@ph-ludwigsburg.de

Prof. Dr. Ulrik SCHROEDER, Professor for computer science and Head of the computer-supported teaching and learning research group in the computer science department at RWTH Aachen University, Germany

Research interests: software engineering methods in instructional design, interactive learning and assessment technology, implementation of action-oriented E-learning

Contact:

Ulrik Schroeder

Lehr- und Forschungsgebiet Informatik IX
RWTH Aachen
Ahornstr. 55
52074 Aachen, Germany

E-mail: schroeder@informatik.rwth-aachen.de

Citation

Spannagel, Christian; Gläser-Zikuda, Michaela & Schroeder, Ulrik (2005). Application of Qualitative Content Analysis in User-Program Interaction Research [39 paragraphs]. *Forum Qualitative Sozialforschung / Forum: Qualitative Social Research*, *6*(2), Art. 29, http://nbn-resolving.de/urn:nbn:de:0114-fqs0502295.

Revised 6/2008