User-Driven Semantic Classification for the Analysis of Abstract Health and Visualization Tasks

Sabine Theis[(✉)], Peter Rasche, Christina Bröhl, Matthias Wille, and Alexander Mertens

Institute of Industrial Engineering and Ergonomics, RWTH Aachen University, Aachen, Germany
`{s.theis,p.rasche,c.broehl,m.wille, a.mertens}@iaw.rwth-aachen.de`

Abstract. Present article outlines characteristics of a general task analysis in terms of digital health visualization evaluation and design. Furthermore, a number of methodological approaches are discussed. One example, in which a hierarchical structure was empirically built with semantic classification by 98 users, will be discussed together with the expected benefits of its successful implementation with respect to system development and human factors research on health data visualizations. It is concluded that experimental approaches to taxonomy construction offer considerable promise in capturing tasks which are relevant but that further investigation is needed validating and iteratively extending the abstract task structures. We thus recommend based on our experiences to conduct a combination of semantic classification with users and hierarchical task analysis to capture all needed task abstraction levels.

Keywords: Task analysis · Methods · Task abstraction · Generalizability

1 Introduction

The volume of data circulating throughout the world is growing rapidly every year. Also in the healthcare sector, more and more researchers, companies and physicians are working with huge amounts of data. Health data are no longer collected only in practice and in studies, but are also captured by patients themselves – be it via mobile phones and body-worn sensor devices, via mobile apps or with social networks. Large amounts of health data open up promising new perspectives for the research, prevention, diagnosis and treatment of diseases.

In this article we consider interconnected information and communication technology aiding healthcare professionals and patients to manage illnesses and health risks, as well as promoting health and wellbeing as digital health systems. Digital health applications support the user in a various amount of tasks. Most consumer applications serve personal monitoring tasks like monitoring of medication intake [1, 2]

or monitoring of health specific vital parameters and health-related behavior in which visualizations are applied to influence users behavior [3–8].

On the professional side of digital health systems the major part of the applications focusses on communicating medical test results and personal health records [9–12]. Digital health systems for example aggregate medical data from different sources so that data appreciates in value if related to events, or correlated with vital sensor or behavioral data. Automation may support the analysis to some extent, but given the dynamics, flexibility and the creativity of the human brain it can hardly be substituted by machines. To integrate users into data driven health processes data visualization are an effective layer between abstract mathematical concepts and human cognition. Finding the right information, understanding and making sense of in large amounts of structured and unstructured health data is strongly influenced by human capabilities, but deriving generalizable results from ergonomic visualization evaluation requires a set of tasks which are relevant to all or as much digital health applications as possible.

For general research purposes, there is a need for some generalizable knowledge about human activities across applications and over time. But not only research generalizability, also the design of efficient and effective digital health systems requires an analysis of relevant tasks. Thereby it needs to be investigated to which extent a task can be supported by a certain technology when performed by patients and professionals. Together with patients and professionals goals they build the reference against which system functionalities then can be tested. Here, besides a behavioral description of the activity, considerable effort is oriented to the specific case serving development oriented purposes. A combination of the research approach and the development approach is where factors will be measured during a task conducted with a system. When data of a human conducting a task with a certain system is collected, this can also have an emphasis either on research or on development objectives.

We consider the research perspective here as aiming at generalizable output of a controlled laboratory experiment. There are thus different objectives for the tasks resulting from a task analysis.

We differentiate following objectives for the tasks resulting from a task analysis and assign them with different degrees of abstraction, a degree of specificity in terms of users and tasks or domain (Table 1):

Table 1. Task objectives in terms of task abstraction and domain/user specificitie as very high (++), high (+), low (−) or medium (o)

Objectives	Domain specificity	Task abstraction	User specificity
Measure factors depending on a task (performance, human error, strain e.g.)	++	−	+
Design a system that supports a task (user-centered/participatory design)	+	−	++
Measure factors depending on a task supported by a system (usability test, system evaluation, visualization evaluation)	+	+/−	+

A challenge regarding controlled visualization evaluation is the choice of task, which needs to be relevant to real world use of the visualization. Particularly in controlled studies, it can be easy finding a task that is modestly measurable by disregarding its relevance for the application context. This article looks at methods to find tasks which can be used during controlled experiments to measure human factors depending on the task supported by visualization.

2 Abstraction in Visualization Tasks

Abstraction levels of tasks have been discussed within the visualization community. Rind and his colleagues [13] for example describe tasks as different in terms of abstraction, composition and perspective. In order to disambiguate the use of task terminology they also construct a three-dimensional conceptual space of visualization tasks. According to this three-dimensional conceptual space a task has a certain level of abstraction ranging from concrete to abstract. Furthermore, a task disposes of a granularity level when broken down in several sub-tasks. The authors describe on the one hand why a task is done by the objective and on the other hand how it is done by different actions.

Munzner [14] provided a model of nested layers for the design and evaluation of visualizations. At the outer level the domain and problem of interest are defined while during the following step the data and task abstractions for that problem are identified. At the third layer visual encodings and interaction methods for data and task abstractions are developed so that at the innermost level according algorithms will be developed. This model invokes domain problem descriptions or objectives followed by an identification of according data and task abstractions as preparation for generalizable human factors evaluation results.

A task framework with different abstraction levels covering objectives on the why-dimension and actions on the how-dimension was developed by Brehmer and Munzner [15]. Subsequently, Brehmer et al. [16] characterize task sequences related to visualizing dimensionally-reduced data. This time, information from interviews with 24 analysis experts was classed with the multilevel typology of abstract visualization task framework. Even if interview questions and the translation of interview transcriptions into abstract visualization tasks is not described in detail, it becomes clear, that their task analysis considers concrete analytical procedures applied by analysts in different domains. This task analysis focusses on objectives and actions during the handling of a specific data type (Fig. 1).

In addition, Miksch and Aigner [17] described an abstract framework to guide visualization designers in constructing time-based visualizations with the data-user-task design triangle. The authors describe tasks as manifold and depending on the type of data that has to be analyzed. Following the authors, tasks are defined by the questions users want to answer with the help of visual representations. Two task types are described as twofold [18]: elementary tasks address individual data elements like groups of data or individual values. To look up a blood-pressure value at one point in time for example. Then the user has a target and just wants to find it within the data.
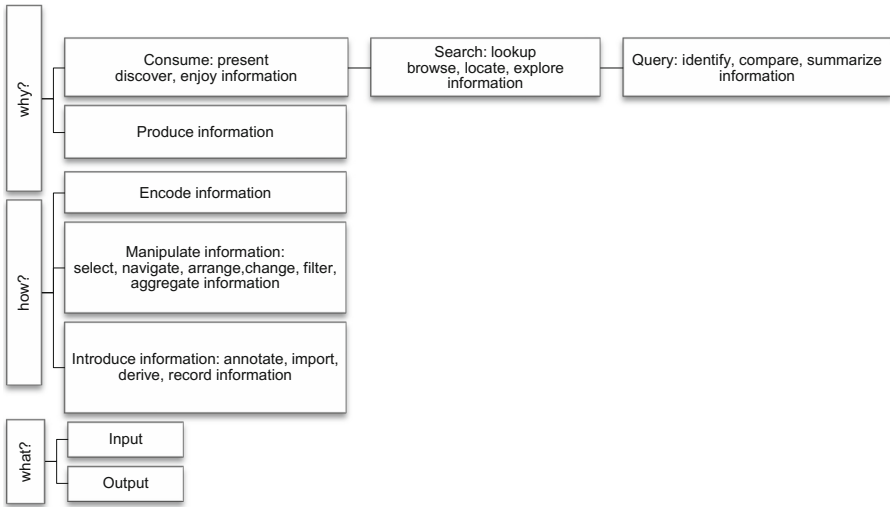
Fig. 1. The multi-level typology of abstract visualization tasks represents the 'why', 'how' and 'what' dimensions of visualization tasks [15]. Task abstraction is made explicit.

3   Task Abstraction of Digital Health Tasks

While researchers in the field of data visualization are aware of the importance of abstract tasks for evaluation and while there is a large number of work on how to construct hierarchical structures, the analysis of abstract tasks in the medical field is, if at all, only implicit. Furthermore, classifications are applied to make concepts and their relation clear in order to differentiate ambiguous terms representing the concept of IT supported medical processes [19, 20]. Using them as basis for ergonomic evaluation bears the risk of findings with minor practical relevance. On the other hand, domain specific tasks only serve a differentiation of ambiguous terms representing the concept of IT supported medical processes. Bashshur et al. [20] constructed the taxonomy of telemedicine necessary to that end. They differentiated among other dimensions, user tasks when describing the functionality dimensions *consultation*, *diagnosis*, *monitoring* and *mentoring*. Unfortunately, his research remains vague when it comes to the origin of his classification. So current study takes them as starting point for a user oriented perspective on digital health task and data analysis by having it verified from a domain expert's perspective and extend it if needed (Fig. 2).

4   Task Analysis Methods

As illustrated different abstraction levels, user and domain specificity of tasks are relevant when targeting generalizable results of data visualization evaluation. At the same time, methods to analyze tasks produce results differing in these dimensions. Different methods for analyzing tasks exist. Subsequently common human factors
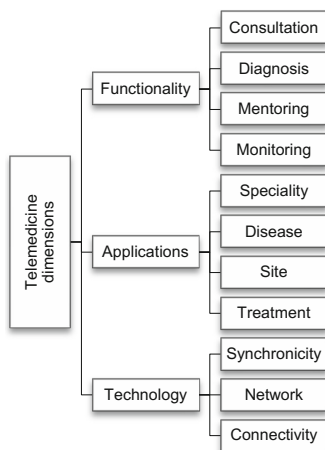
Fig. 2. Taxonomy of telemedicine considers abstract tasks of digital health systems within the functionality dimension. Task abstraction remains implicit here.

methods for task analysis are described: hierarchical task analysis (HTA), cognitive task analysis (CTA) and observation. In addition, a method is described which is less common: semantic classification by users.

4.1   Hierarchical Task Analysis

The idea behind the *Hierarchical Task Analysis* (HTA) is to subdivide tasks performed by humans into sub-tasks in order to create an abstract hierarchical model by clustering. Tasks that bear close resemblance to each other are assigned to one group and must be part of at least one group but may be part of several groups. Groups are then labeled in terms of the work domain or the work functions and can be iteratively refined or regrouped. So an observation of a very specific task with a specific user or user group has to be made resulting of unstructured lists of words describing actions which to organize using notation or diagrams. The task analysis involves users only as observation object, not as integrated into the task analysis process.

4.2   Cognitive Task Analysis

Cognitive Task Analysis (CTA) aims at understanding tasks that require a lot of cognitive activity from the user, such as decision-making, problem-solving, memory, attention and judgement. The CTA methods analyze and represent the cognitive activities users utilize to perform certain tasks. Some of the steps of a cognitive task analysis are: the mapping of the task, identifying the critical decision points, clustering, linking, and prioritizing them, and characterizing the strategies used. For the purpose of CTA various interview and observation procedures are applied in order to capture a description of the knowledge that experts use to perform complex tasks. Complex tasks

are defined as those where performance requires the integrated use of both controlled (conscious, conceptual) and automated (unconscious, procedural or strategic) knowledge to perform tasks that often extend over many hours or days.

4.3 Observation

Participant *observation* is an ethnographic method in which a researcher participates in, observes, and records the everyday activities and cultural aspects of a particular social group. It typically includes research over an extended period of time (rather than a single session) and takes place where people live or work (rather than in a lab). Participant observation involves active engagement in activities in contrast to observation where researchers simply observe without interacting with people. Often this method can be part of a hierarchical task analysis.

4.4 Semantic Classification

During *semantic classification* for analysis of work tasks, each task of the worker is described verbally. Characteristics of words and frequencies found with the sematic differential then lead to the development of a task taxonomy. Consensus judgements of tasks lead to a relevance ranking and structure of the tasks. Semantic classifications are able to support task descriptions and hence the measurement during an evaluation. They perfectly provide conceptual clarity and categorize information for an increased theoretical understanding and predictive accuracy in empirical research. In order to understand the differences to previously mentioned methods we provide following example:

4.5 Example

Given example illustrates how general tasks are identified by digital health systems users with the help of semantic classification. Via online questionnaire professionals in the domain of health and digital health as well as patients

1. rated relevance of given abstract medical tasks [20],
2. rated relevance of given abstract visualization tasks [15, 21]
3. rated abstract visualization tasks relevance for digital health tasks.

The sample of 98 participants consisted of a group of 47 digital health experts and a group of 51 older adults with a mean age of 55.76 representing the patient perspective. Group differences were computed in order to illustrate the extent to which general health and visualization tasks are relevant. The study was able to verify and extend existing abstract visualization and digital health systems. Abstract tasks from the visualization and health domain could be mapped to domain tasks. A chi-square test of independence was performed to examine the relation between relevance frequency counts and user group (older adults, tele medical experts). The relation between these variables was highly significant for mentoring, X^2 (4, N = 67) = 14.14, p = .002**

and monitoring, X^2 (4, N = 70) = 22.13, p = .00**. User rankings and group differences were synthetized in taxonomy. Group differences could be found most often for the abstract visualization tasks. Root node of resulting taxonomy was "digital health systems task". Ranked domain tasks from closed question (1) monitoring, (2) consultation, (3) diagnosis and (4) mentoring built the next level. On the same level tasks resulting from open questions were added according to their code count frequency as siblings to previously mentioned tasks (4) therapy, (5) communication, (6) cooperation, (7) documentation and (8) quality management. Each of these eight main tasks were complemented by (1) data types from open questions, which provide a qualitative description of the data from the user perspective, (2) data types from closed questions and (3) top-ranked on total sample abstract visualization tasks.

5   Discussion

This paper emphasizes abstraction levels of tasks with regard to objectives when using tasks resulting from a detailed task analysis. It was stated that intended generalizability of research results relates to the required abstraction level of an experimental task together with the data types/application domain and user. Based on our experiences from a user study on the classification of abstract visualization and health tasks we can state that relevance of tasks for digital health systems can be easily judged by users. Semantic classification by users provides thus a feasible method to build and rank abstract tasks and to join tasks from different domains (health and visualization). A precise task definition should be given to the participant but it should be avoided to give specific examples as this would prime the user and the level of abstraction from resulting tasks would decrease. While task abstraction is one parameter the user and domain specificity has to be controlled, too. That means a result list of tasks is more generalizable if all important user groups of an application domain agree on the importance of a task. User type variability and sample size increase validity of the results from user driven semantic classifications. The initial goal of deriving tasks which would produce as generalizable visualization evaluation results as possible instead requires further investigation of semantic classifications connecting tasks through different levels of abstraction. In comparison, HTA and CTA are too specific to investigate abstract tasks, but their procedures might be of interest for bridging abstract to specific tasks. One could for example let users subdivide the abstract tasks until they have the granularity required for described evaluation purpose. Further type of presented research type also might utilize a more extensive set of given tasks. One major definitional problem for task analysis and semantic classification is that it is difficult to agree on what a task is. Any set of behavior represented by a textual label can be considered as a task here. A possible solution could be to concentrate on describing tasks just in verbs instead of nouns in order to disconnect them from the application context and user specificities.

References

1. Browne, S.H., Behzadi, Y., Littlewort, G.: Let visuals tell the story: medication adherence in patients with type II diabetes captured by a novel ingestion sensor platform. JMIR Mhealth Uhealth **3**(4), e108 (2015)
2. Theng, Y.-L.: Socially-mediated mobile system for tuberculosis public health alert, care and education (MobiTB-ACE). In: Medicine 2.0 Conference (2012)
3. Jones, A.P.: Indoor air quality and health. Atmos. Environ. **33**(28), 4535–4564 (1999)
4. Subramonyam, H.: SIGCHI: magic mirror - embodied interactions for the quantified self. In: Proceedings of the 33rd Annual ACM Conference Extended Abstracts on Human Factors in Computing Systems, pp. 1699–1704. ACM (2015)
5. Brown, B., Chetty, M., Grimes, A., Harmon, E.: Reflecting on health: a system for students to monitor diet and exercise. In: CHI 2006 Extended Abstracts on Human Factors in Computing Systems, pp. 1807–1812. ACM, New York (2006)
6. Ali-Hasan, N., Gavales, D., Peterson, A., Raw, M.: Fitster: social fitness information visualizer. In: CHI 2006 Extended Abstracts on Human Factors in Computing Systems, pp. 1795–1800. ACM, New York (2006)
7. Swan, M.: Sensor mania! The internet of things, wearable computing, objective metrics, and the quantified self 2.0. J. Sens. Actuator Netw. **1**(3), 217–253 (2012)
8. Frost, J., Smith, B.: Visualizing health practice to treat diabetes. In: CHI 2002 Extended Abstracts on Human Factors in Computing Systems, pp. 606–607. ACM (2002)
9. Milash, B., Rose, A., Widoff, S., Shneiderman, B., Plaisant, C.: LifeLines: visualizing personal histories. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, pp. 221–227 (1996)
10. Kim, M.I., Johnson, K.B.: Personal health records: evaluation of functionality and utility. J. Am. Med. Inform. Assoc. **9**(2), 171–180 (2002)
11. Archer, N., Fevrier-Thomas, U., Lokker, C., McKibbon, K.A., Straus, S.E.: Personal health records: a scoping review. J. Am. Med. Inform. Assoc.: JAMIA **18**(4), 515–522 (2011)
12. Frost, J., Massagli, M.: Social uses of personal health information within PatientsLikeMe, an online patient community: what can happen when patients have access to one another's data. J. Med. Internet Res. **10**(3), e15 (2008)
13. Rind, A., Aigner, W., Wagner, M., Miksch, S., Lammarsch, T.: User tasks for evaluation: untangling the terminology throughout visualization design and development. In: Proceedings of the Fifth Workshop on Beyond Time and Errors: Novel Evaluation Methods for Visualization, pp. 9–15. ACM, New York (2014)
14. Munzner, T.: A nested model for visualization design and validation. IEEE Trans. Vis. Comput. Graph. **15**(6), 921–928 (2009)
15. Brehmer, M., Munzner, T.: A multi-level typology of abstract visualization tasks. IEEE Trans. Vis. Comput. Graph. **19**(12), 2376–2385 (2013)
16. Brehmer, M., Sedlmair, M., Ingram, S., Munzner, T.: Visualizing dimensionally-reduced data: interviews with analysts and a characterization of task sequences. Submitted to ACM BELIV Workshop, pp. 1–8 (2014)
17. Miksch, S., Aigner, W.: A matter of time: applying a data–users–tasks design triangle to visual analytics of time-oriented data. Comput. Graph. **38**, 286–290 (2014)
18. Andrienko, N., Andrienko, G.: Exploratory Analysis of Spatial and Temporal Data: A Systematic Approach. Springer Science & Business Media, Berlin (2006)
19. Ingenerf, J.: Telemedicine and terminology: different needs of context information. IEEE Trans. Inf. Technol. Biomed. **3**(2), 92–100 (1999)

20. Bashshur, R., Shannon, G., Krupinski, E., Grigsby, J.: The taxonomy of telemedicine. Telemed. J. e-Health: Off. J. Am. Telemed. Assoc. **17**(6), 484–494 (2011)
21. Shneiderman, B.: The eyes have it: a task by data type taxonomy for information visualizations. In: Proceedings/IEEE Symposium on Visual Languages, September 3–6, 1996, Boulder, Colorado, pp. 336–343. IEEE Computer Society Press, Los Alamitos (1996)