Field Evaluation of a New Railway Dispatching Software

Isabel Schütz
Department of Railway Engineering
TU Darmstadt
Darmstadt, Germany
schuetz@verkehr.tu-darmstadt.de

Anselmo Stelzer
Department of Railway Engineering
TU Darmstadt
Darmstadt, Germany
stelzer@verkehr.tu-darmstadt.de

***Abstract*—In this paper, we present a program which assists in choosing the right evaluation methods. The Test Selection Program contains a four step method for choosing evaluation methods: describing important boundary conditions of the study, defining important evaluation criteria, choosing evaluation methods and deciding on the right order in time. As a practical example, connection dispatching will be introduced and used to illustrate the functionality of this program.**

Keywords-evaluation methods; connection dispatching; usability; railway engineering.

I. INTRODUCTION

The domain of railway engineering is safety-critical, so the main focus is on ensuring safe operations. The secondary focus is on punctuality to increase and maintain customer satisfaction and service quality [1].

When developing safety-critical software, the main focus is mostly on safety. Beyond that, usability nowadays does not play a big role [2].

In the last few years, the focus has been shifting more and more to customer oriented dispatching, for example through good customer information or connection dispatching. Especially in the field of connection dispatching, which we will concentrate on in this paper, suitable software to display the connection status or connection conflicts is not yet known. At the moment, general software is used which poorly supports connection dispatching [5]. To comply with requirements given by contracting traffic authorities, e.g., regional transport authorities, this needs to be changed. Moreover, the importance of intermodal connection dispatching rises and is therefore integrated into traffic contracts, often combined with a contractual penalty [4].

The consideration of intermodal connections increases the dispatchers' workload compared to today's situation. The increasing requirements cannot be met with the contemporary personnel shortage. Furthermore, training dispatchers takes a long time because of the need to gain experience in everyday work. Without well-designed software which supports the dispatcher in doing his job of connection dispatching, this kind of workload cannot be sufficiently dealt with. A clear and efficient way of displaying all connections is important to facilitate handling them simultaneously and also to be able to automate several steps within this process of connection dispatching. This is expected to reduce the dispatcher's workload. For these reasons, a project has been established to support the dispatcher in connection dispatching. Its objective is to support dispatchers by displaying planned connections and connection conflicts together in a concise, newly developed software interface. The project contains a field study for which suitable evaluation methods needed to be found. Since the work of a railway dispatcher is not safety-critical (in contrast to the work of a traffic controller), it offers optimal possibilities to establish usability in the field of Railway Engineering [5]. We will use this project as a practical example to illustrate a new process of choosing the right evaluation methods.

Several other evaluation projects with different partners – mainly expert evaluations and evaluations in a simulation environment – have been conducted [8]. During these projects, the importance of using suitable evaluation methods was pointed out, as was the necessity of reasonable combination in the right order in time.

Since the introduced project's field study will end at the beginning of 2015, in this paper we concentrate on the following issues: How to choose the most suitable evaluation methods and how to combine them reasonably and in the right order in time. For this, in Section II the field of railway dispatching in general, and the process of connection dispatching in particular, are introduced. Then, in Section III, the focus is on the tool we developed to support an optimal choice of suitable evaluation methods. Thereafter, in Section IV, special attention will be paid to the specific example of the field study conducted with dispatchers to show how we chose the evaluation methods and why we combined them in the actual way and sequence we did. We finish our contribution with a conclusion and a description of future work in Section V.

II. RAILWAY DISPATCHING

Dispatchers are well-known in different realms, for example rescue services or logistics. Likewise, in the railway system, we distinguish between different dispatchers. In this paper, we will concentrate on the dispatcher of the train operating company (further referred to as "the dispatcher"). It is his job to coordinate activities, such as passenger information, disposition of rolling stock and personnel, communication, coordination and connection dispatching. He concentrates on maintaining and increasing the customer satisfaction. In doing so, a consideration of the feasibility and the consequences of a performed measure is necessary. The dispatcher – in contrast to a traffic controller – has no

 ISBN: 978-1-61208-382-7

responsibility for safety, as he is not using any interlocking systems [3].

A. Definition of a Connection

A connection is the possibility for the passengers to change from an arriving feeder train to a departing distributor train within a certain interchange time. Possibly, also connections from and to different modes of transport can be considered and are referred to as "intermodal connections" [3].

B. The Process of Connection Dispatching

To travel from one location to another, interchanges between trains are often necessary in a railway network as dense as in Germany. To change from the feeding train to the distributing train under consideration of the physical distance between the platforms and the local conditions inside the station, a certain interchange time and also an additional buffer time have to factored into the planning process of a travel. Normally, this carefully considered time is enough to successfully switch trains without any further work for the dispatcher.

The dispatcher's work in connection dispatching begins when trains are delayed. In this case, a connecting train might not be reached within the above-mentioned predetermined interchange time. In some cases, trains can wait for each other without generating any issues within the so called waiting-time arrangements which permit the distributor a few minutes delay from the original departure time.

If the remaining time between the arrival of the feeding train and the departure of the collecting train under consideration of the waiting time is not sufficient for the passenger to switch trains, the dispatcher has to perform several activities:

He has to decide whether this connection can be reached or not. This decision is often based on his experience gathered about interchanging passengers or his knowledge of the location which allows him to predict that an interchange can be performed faster than assumed initially, e.g., because of the interchange taking place on the same platform. If the dispatcher decides upon securing the connection, he has to confer with the railway infrastructure company and possibly with other train operating companies whether it is possible to let the distributing train wait longer than the initial waiting time. Both can reject keeping this connection when they fear negative consequences along the following route of the train or further connection conflicts at other stations. In the case of having decided not to keep the connection, the dispatcher has to inform the passengers about alternative connections. In both cases, he has to inform the train drivers, his dispatching colleagues and the travellers on the train [3].

C. The Newly Developed Dispatching Software

Nowadays, the process described in the previous section is mainly done by hand. That is why a prototype software has been developed which is intended to support the dispatchers during this process by displaying connections and connection conflicts [7]. The connections are arranged in a matrix with the feeder train on the y axis and the collecting train on the x axis. Connections are presented within the cells of the matrix. Thus, it is possible to see all important information at a glance. By clicking the cell, more detailed information can be retrieved and further actions can be taken [5].

For the visualisation, each connection is assigned to a category – green, yellow or red – comparable to a traffic light. The coloured cell, which also displays additional information concerning the connection, represents this category. It advises the user on how to proceed with a connection (conflict) [5].

The prototype software developed was initially evaluated by experts. Several Focus Groups were hosted to gain input on the interface and its functionality, but also several cognitive walkthroughs were employed on that account. The experts consisted of special dispatchers, who also conduct trainings, and of usability experts. Subsequently, the prototype software was tested in a first user study within a simulation environment close to reality – the so-called Eisenbahnbetriebsfeld Darmstadt (EBD) – with prospective users to confirm the design and its suitability for use [8].

The EBD is a research facility for railway operations which provides the complete chain of railway operation and dispatching. It embodies a realistic simulation environment where tracks and trains are models, but interlocking technology, dispatchers' software and auxiliary equipment, such as phones and walkie-talkies, are real. In this user study, the EBD was used to prove that this prototype software enables the dispatchers to handle a realistic testing scenario. In this way, the prototype software was tested for the first time under realistic conditions and could indeed prove its benefits [9].

Apart from making sure that this prototype software can be used in the field, further improvement of the user interface was to be achieved before testing in the field – using the feedback of the tests in the EBD. Moreover, one major aim of this test was to ensure that all essential functions have been implemented as well as keeping the amount of errors small to prevent users in the field study from being frustrated. Since field testing is comparably expensive, an assurance was necessary that all results are valid and can be used for the analysis and improvement of the prototype software [8].

III. Test Selection Program

Considering these above-mentioned preliminary studies, we wanted to make sure that the field study and its results are not distorted by malfunctioning prototype software or by user frustration. Also, crucial for the success of a field study is collecting the right evaluation data by using the most suitable evaluation methods. To ensure the choice of suitable evaluation methods and reasonable combinations thereof in the right order in time, a program was developed during a six-month study [6].

The Test Selection Program is designed to assist in choosing the most suitable evaluation methods given a specific evaluation context. It selects suitable methods based on the criteria entered by the user, employing underlying

 ISBN: 978-1-61208-382-7

filtering algorithms. Its functionality could be proven in several studies and it is integrated in an ongoing updating process.

All evaluation methods are characterised by a descriptive profile. This profile is the basis for the underlying algorithms choosing the most suitable evaluation methods. As can be seen from Figure 1, it is displayed on the left side of the program mainly comprised of toggle buttons, but also of two toggle button groups with a checkbox each. When clicking a toggle button, it changes its colour to red and stays pushed in to indicate a selected criterion, as can be seen in Figure 2. Clicking the button again releases it and thus de-selects the property. The toggle button groups are first greyed out until the respective checkbox has been selected (see Figure 2). These visual cues allow to easily identify the choices the user has taken so far [6].

On the right hand side (see Figure 1), a list is displayed, arranged in a tree structure. Initially, it contains all evaluation methods that are included in the program. As soon as buttons are clicked, the list reduces by the methods not meeting the selected criteria (see Figure 2). So, after having entered all boundary conditions and criteria of the study, the user can instantly see the most suitable evaluation methods [6].

For further technical or functional details about the program see [6].

IV. CHOICE OF EVALUATION METHODS

The choice of evaluation methods based on the Test Selection Program described in Section III can be divided into four steps:

The first step is to describe the important boundary conditions of the evaluation to take place (cf. Section IV.A). The second step is to define important evaluation criteria (cf. Section IV.B). Thirdly, these boundary conditions and criteria are entered into the program and instantly yield an optimal choice of evaluation methods (cf. Section IV.C). Since in most cases there is more than one suitable evaluation method, the user has to decide on how many evaluation methods to use, how to combine them and also which subtype of a method to use, if applicable. In a fourth step, the right order in time has to be chosen (cf. Section IV.D).

During a six-month research study [6], many scientific references concerning existing methods and procedures were analysed, compared and integrated into one process to choose evaluation methods, and a relevant workflow was defined. It embodies a clear structure and helps to avoid missing an important step. Beyond the original definition of the process, it has been and is constantly being refined during other research projects.

In this paper, we will describe this process based on the specific example of the user interface for displaying connections and connection conflicts in the context of connection dispatching (see Section C).

A. *Step 1: Describing Important Boundary Conditions of the Study*

To find the optimal combination of evaluation methods, the boundary conditions of the evaluation need to be described first.

For our study, it can be stated that we are rather late in the development process and therefore have a working prototype software which should be used during the evaluation, and not only a paper prototype or a mock-up. Thus, it is possible to test in the field with real users to get a realistic testing scenario and not in a laboratory or online.

The focusing aspects of the evaluation were worked out with the project partner: design and functionality with special regard to usability and acceptance of the prototype software (and not performance or a comparison of two prototypes, for example) to ensure an optimal support of the dispatchers. The prospective users had an active role in the evaluation, but they were not involved further in the software development process and are therefore not part of the decision making (meaning a more passive contribution).

In summary, these are the important boundary conditions for our study:

- Working prototype software
- Study takes place in the field
- Passive contribution of the user
- Late in the development process
- Focus on design, functionality, usability and acceptance of the prototype software

B. *Step 2: Defining Important Evaluation Criteria*

After this first step of describing important boundary conditions, additional criteria of the evaluation need to be defined.

At the beginning, it was decided that the evaluation methods should generate diverse data to get a comprehensive understanding and a broad spectrum of results. ”Diverse data” refers to subjective and objective data as well as qualitative and quantitative data. The data should be generated either directly during the interaction of the user with the prototype software, or indirectly after having used the prototype software. This depends on the evaluation method and especially on the possibility of integrating such an evaluation method into daily work without distracting the dispatchers too much.

Due to the advanced project progression, evaluation methods with a low to medium expenditure of time and a low to medium effort for analysis and interpretation were needed. Methods with high expenditure of time, high effort for analysis and interpretation were not considered further.

Based on the decision to generate diverse data and because it is not really possible to generate subjective and objective data with the same evaluation method, two different search profiles were created to find all suitable evaluation methods. However, some of the parameters are equal for both profiles:

- High degree of detail
- Low to medium expenditure of time
- Low to medium effort for analysis and interpretation

 ISBN: 978-1-61208-382-7

1) Description of Search Profile 1

The first search profile was supposed to generate subjective, qualitative data and – to make sure that the dispatchers are not distracted during their work – indirect data generation was chosen.

2) Description of Search Profile 2

The second search profile was intended to generate objective and either qualitative or quantitative data gained directly during the usage of the prototype software.

C. Step 3: Choosing evaluation methods with help of the Test Selection Program

All these boundary conditions (see Section A) and criteria (see Section B) are to be entered into the Test Selection Program which was particularly developed to easily choose from many evaluation methods, and which has been described in a previous section.

1) Results for Search Profile 1

For Search Profile 1 (cf. Section B.1) the program delivered the following methods:

- Focus Group
- Interview
    - Half-structured Interview
    - Plus-Minus-Method
- Diary Studies

Because of the great effort to coordinate the interviews and the difficult time availability of the dispatchers being busy in their work, we decided against conducting single interviews. That is why, for the considered project, the methods Diary Studies and a Focus Group were selected.

2) Results for Search Profile 2

For Search Profile 2 (cf. Section B.2), we gained the following methods as results:

- Observation
    - Participatory
    - Non-participatory

We decided to use the Participatory Observation to be fully integrated in the dispatching process and to gain as much objective data as possible. Using this evaluation method, it is possible to interact with the dispatchers and thus prevent them from feeling uncomfortable.

D. Step 4: Deciding on the Right Order in Time

After having chosen the most suitable evaluation methods in Section C, the next step is to decide on the right order in time.

It was decided to start with the Diary Studies, then continue with the Observation and at last, conduct the Focus Group. The reasons for choosing this order were the following: Using the Diary Studies as a first evaluation method allows to ascertain in advance which items to concentrate on in detail during the Observation. This will help the observer to be more focused during the subsequent Observation. Since it was infeasible to discuss remaining questions which arose while reading the Diaries or during the Observation, the Focus Group was chosen as the last evaluation method. With the first two evaluation methods, enough input for the discussion will be gained. Moreover, key items can be discussed in detail and remaining open questions can be clarified. Also, users are able state their opinion and their improvement proposals orally and directly to the persons in charge. With the combination of these three methods in the described sequence, a fully detailed overview of the design and functionality of the prototype software as well as input for improvement can be obtained.

V. CONCLUSION AND FUTURE WORK

Choosing the most suitable evaluation methods, to combine them reasonably and in the right order in time is important for the success of a study since this is the basis for the data generation and thus for undistorted, reliable results. Unfortunately, this preparation is already challenging due to the variety of evaluation methods and the numerous possibilities to combine them as elaborated on in [6]. These encountered challenges are addressed with the help of the presented four-step method (see Section IV) which is integrated in the Test Selection Program described above (see Section III). It aims at supporting the choice and combination of evaluation methods.

For the field study described in this paper (see Section II.C) the software could successfully be used. The next step for the evaluation in the above-mentioned field test is to generate – with the help of the combination of evaluation methods described in the section above – all necessary and also detailed data to improve the prototype dispatching software, to adapt it to the dispatchers' needs in the best way possible such that a better support in connection dispatching is given and thereby the dispatchers' satisfaction can be increased.

Apart from this, there is ongoing research to further improve the process of choosing the most suitable evaluation methods (see Section IV) and in doing so, also the software supporting this process (see Section III). The latter was well proven first in a simulation environment [8], in the field of expert evaluation [8] and finally also in the here presented field test (cf. Section IV).

Before conducting the aforementioned second turn of evaluations, an investigation concerning participating users has to be done:

Since dispatchers located at two different sites participated in the evaluation, a decision has to be made whether to let

- The same dispatchers or
- A completely new set of dispatchers or
- A combination of previous and new dispatchers or
- All previous dispatchers and additional new dispatchers participate in the second evaluation.

A key question to be addressed is which users to omit, all from one or several from each location. In the case of omitting dispatchers, questions to be addressed are the arrangement of the exclusion, which person should be in charge to decide, how to prevent distortion of results caused by preferring power users or sceptics and decision-making without sufficiently knowing the dispatchers.

These issues should be addressed before starting a second session of field tests for the described project. The results we

 ISBN: 978-1-61208-382-7

could obtain so far already have shown possibilities to further improve user interfaces with respect to usability to grant better support for people working in the field of railway control and dispatching.

Further research also concerns the transferability to other fields of research, e.g., railway traffic controllers, air traffic controllers or even different industries. The conditions on these areas seem similar, but a practical proof of application of the presented tool and method is missing.

REFERENCES

[1] DB Netz AG, “Guideline 408 – Driving Trains and Shunting” (“Richtlinie 408 – Züge fahren und rangieren”). Karlsruhe: DB Kommunikationstechnik GmbH, 2012.

[2] A. Kauppi, “A Human-Computer Interaction Approach to Train Traffic Control”. Uppsala: Department of Information Technology, 2006.

[3] C. Kendziora and A. Böhme, “Guideline 615 01 – Traffic Management Passenger Transport” (“Richtlinie 615 01 – Transportleitung Personenverkehr”). Frankfurt am Main: DB Fernverkehr AG/DB Regio AG, 2010.

[4] Ministerium für Infrastruktur und Raumordnung Land Brandenburg, “Financing Public Transport in the State of Brandenburg” (“ÖPNV-Finanzierung im Land

[5] Brandenburg”). Potsdam: Land Brandenburg, Ministerium für Infrastruktur und Raunordnung, 2007.

[6] T. Schnick, A. Wolters, and A. Stelzer, “Visualisation of Connection Conflicts in Dispatching” (“Visualisierung von Anschlusskonflikten für die Disposition”). Deine Bahn, vol. 41, no. 6, pp. 26-29, June 2013.

[7] I. Schütz, “New user interface concepts for railway management” (“Neue Bedienoberflächenkonzepte für die Betriebsführung von Eisenbahnen”). Darmstadt: Department of Railway Engineering, 2012.

[8] A. Stelzer, A. Oetting, and F. Chu, “Connection Dispatching - an Algorithmic and Visual Support for the Dispatcher”. The 13th World Conference on Transportation Research – Rio de Janeiro 2013, July 15-18, 2013, July 2013, ISBN: 978-85-285-0232-9.

[9] A. Stelzer, I. Schütz, and A. Oetting, “Evaluating Novel User Interfaces in (Safety Critical) Railway Environments”. Human-Computer Interaction. Applications and Services, 16th International Conference, HCI International 2014, Heraklion, Crete, Greece, June 22-27, 2014, vol. 8512, June 2014, pp. 502-512, doi:10.1007/978-3-319-07227-2_48.

[10] C. Streitzig and A. Stelzer, “TU Darmstadt – Research, Training & More Besides”. EURAILmag, issue 26, 2012, pp. 152-159.

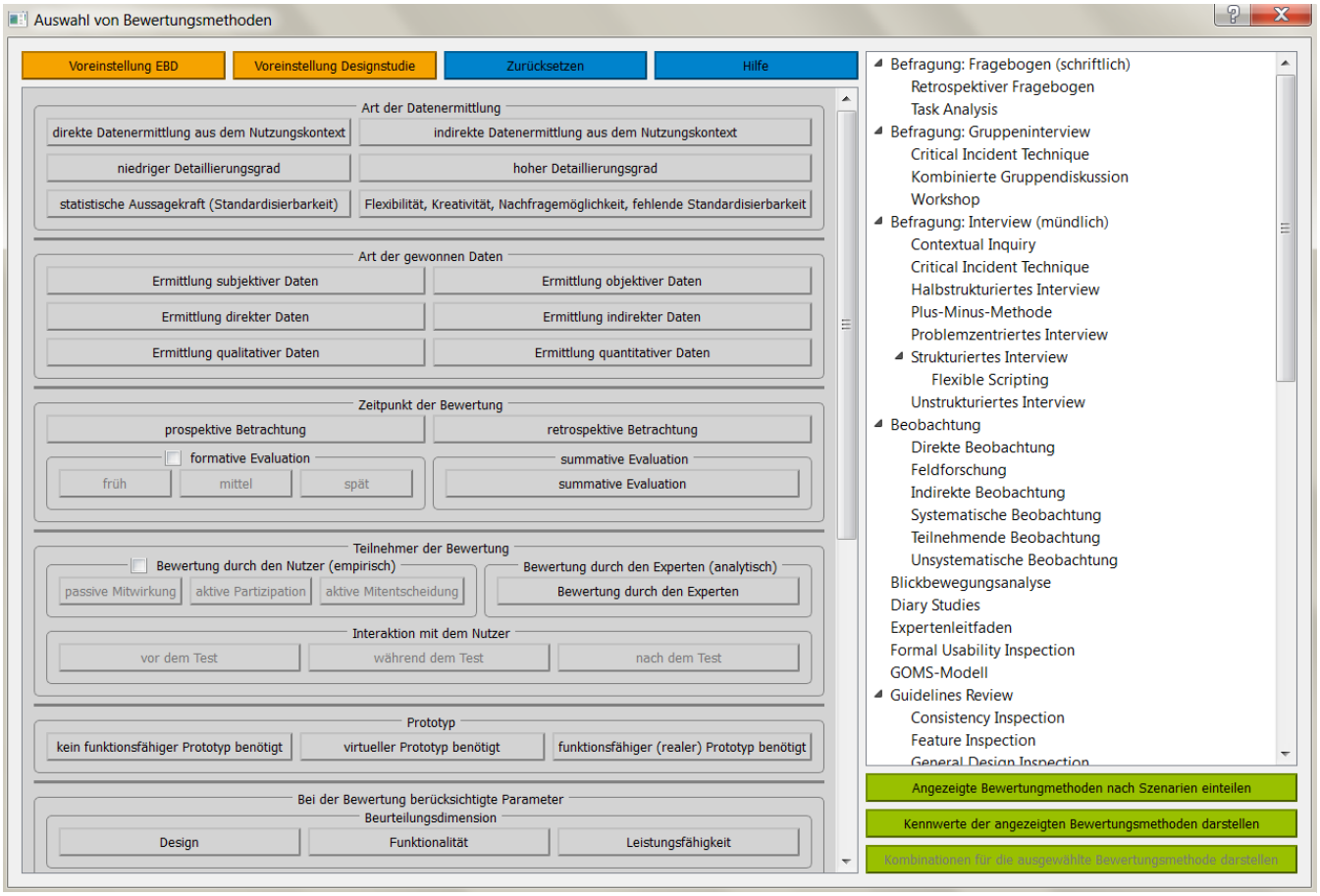

Figure 1. The Test Selection Program. (Source: own representation)

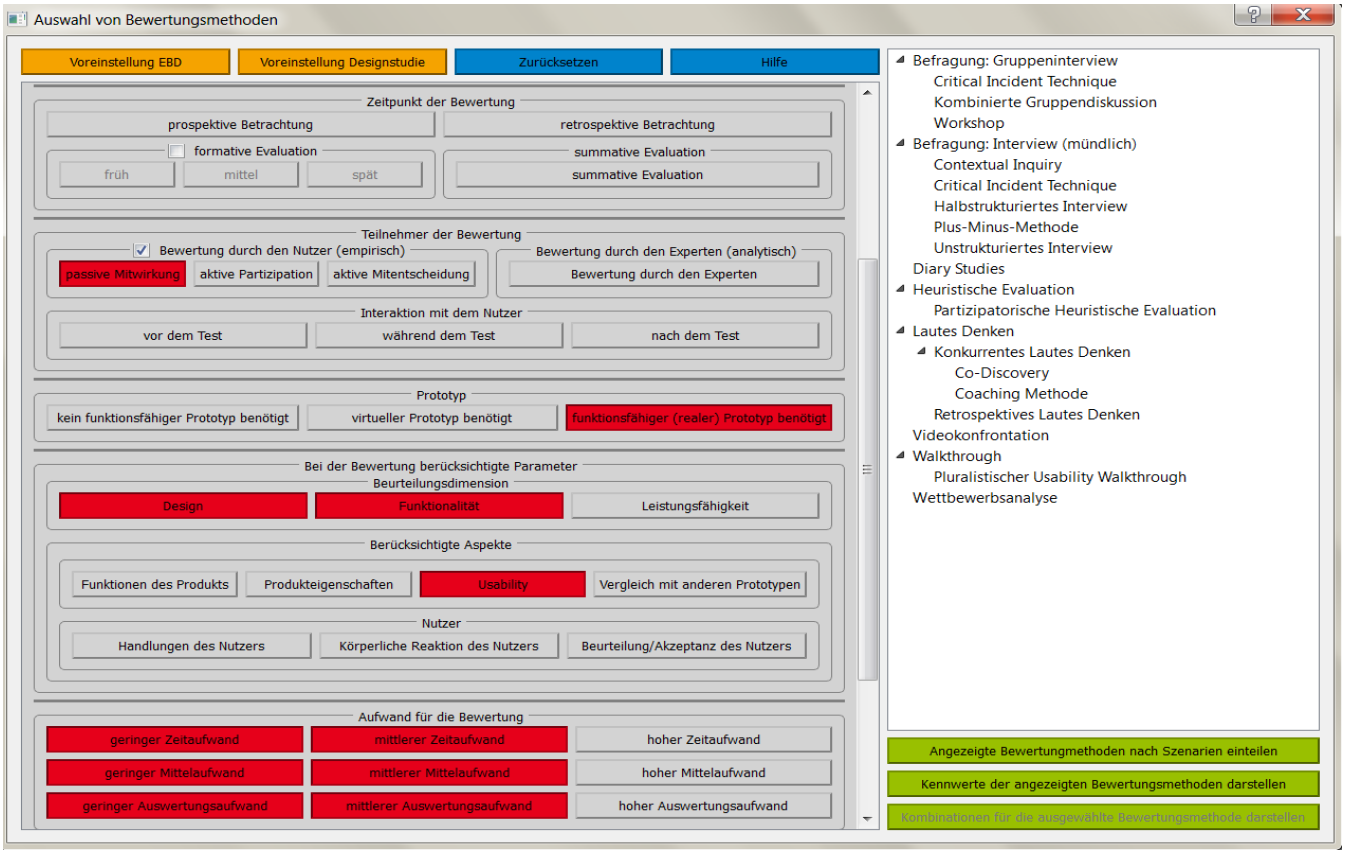

Figure 2. The Test Selection Program. Visible are some selected buttons and an accordingly filtered list of evaluation methods. (Source: own representation)

 ISBN: 978-1-61208-382-7