



Since January 2020 Elsevier has created a COVID-19 resource centre with free information in English and Mandarin on the novel coronavirus COVID-19. The COVID-19 resource centre is hosted on Elsevier Connect, the company's public news and information website.

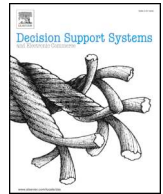
Elsevier hereby grants permission to make all its COVID-19-related research that is available on the COVID-19 resource centre - including this research content - immediately available in PubMed Central and other publicly funded repositories, such as the WHO COVID database with rights for unrestricted research re-use and analyses in any form or by any means with acknowledgement of the original source. These permissions are granted for free by Elsevier for as long as the COVID-19 resource centre remains active.



ELSEVIER

Contents lists available at ScienceDirect

Decision Support Systems

journal homepage: www.elsevier.com/locate/dss

The effect of interactive analytical dashboard features on situation awareness and task performance



Mario Nadj^{a,*}, Alexander Maedche^a, Christian Schieder^b

^a Karlsruhe Institute of Technology (KIT), Kaiserstraße 89-93, 76133 Karlsruhe, Germany

^b Technical University of Applied Sciences Amberg-Weiden, Hetzenrichter Weg 15, 92637 Weiden, i.d.Opf., Germany

ARTICLE INFO

Keywords:

Interactive analytical dashboards
 What-if analysis
 Situation awareness
 Out-of-the-loop syndrome
 Eye-tracking
 Operational decision support systems

ABSTRACT

In recent years, new types of interactive analytical dashboard features have emerged for operational decision support systems (DSS). Analytical components of such features solve optimization problems hidden from the human eye, whereas interactive components involve the individual in the optimization process via graphical user interfaces (GUIs). Despite their expected value for organizations, little is known about the effectiveness of interactive analytical dashboards in operational DSS or their influences on human cognitive abilities. This paper contributes to the closing of this gap by exploring and empirically testing the effects of interactive analytical dashboard features on situation awareness (SA) and task performance in operational DSS. Using the theoretical lens of SA, we develop hypotheses about the effects of a what-if analysis as an interactive analytical dashboard feature on operational decision-makers' SA and task performance. The resulting research model is studied with a laboratory experiment, including eye-tracking data of 83 participants. Our findings show that although a what-if analysis leads to higher task performance, it may also reduce SA, nourishing a potential out-of-the-loop problem. Thus, designers and users of interactive analytical dashboards have to carefully mitigate these effects in the implementation and application of operational DSS. In this article, we translate our findings into implications for designing dashboards within operational DSS to help practitioners in their efforts to address the danger of the out-of-the-loop syndrome.

1. Introduction

In the wake of the COVID-19 pandemic, dashboards that summarize large amounts of information have become a topic of public interest. In particular, an interactive web-dashboard from Johns Hopkins University has proven useful in monitoring the outbreak of the pandemic and keeping billions of people informed about current developments [1,2]. Such dashboards help to illustrate key figures on a screen full of information [3] and provide visual, functional, and/or interactive analytical features to handle chunks of information effectively [2]. Similarly, these tools can also help ensure that operational decision-makers in organizations are not overwhelmed by the avalanche of data when trying to use information effectively [4]. The challenge for contemporary organizations and related research is, therefore, to make the collected data more valuable to operational decision-makers. Dashboards and their underlying features have become an essential approach to addressing this challenge [4]. Negash and Gray [5] describe dashboards as one of the most effective analysis tools. As the interaction between operational decision-makers and decision support

systems (DSS) largely occurs via graphical user interfaces (GUIs) [3], the design and implemented features of a dashboard are of specific importance for the effectiveness of a DSS, particularly at the operational level [4]. Over the years, research has proposed several possible solution types. A first research direction focuses on reporting current operations to inform operational decision-makers [3]. Such dashboards are rather static by nature and do not interact with the user. Static dashboards focus on visual features, which attempt to present information efficiently and effectively to a user [4]. Few [3] reports 13 common mistakes in visual dashboard design (e.g., introducing meaningless variety or encoding quantitative data inaccurately). However, static dashboards appear no longer sufficient to account for the increased need to analyze complex and multidimensional data at the operational level. This issue is a key challenge for the visualizations used in static dashboards, which are often labeled “read-only” [6].

Echoing this concern, a second research direction complements static dashboards with functional features [6]. Functional features refer indirectly to visualization but define what the dashboard is capable of doing [4]. The aim is to create interactive dashboards that involve users

* Corresponding author.

E-mail addresses: mario.nadj@kit.edu (M. Nadj), alexander.maedche@kit.edu (A. Maedche), c.schieder@oth-aw.de (C. Schieder).

<https://doi.org/10.1016/j.dss.2020.113322>

Received 15 August 2019; Received in revised form 12 May 2020; Accepted 12 May 2020

Available online 19 May 2020

0167-9236/ © 2020 Elsevier B.V. All rights reserved.

during data analysis. Scientific articles in this stream suggest that individuals should drill down, roll up, or filter information, or are automatically alerted to business situations [7]. Such involvement appears to support users' comprehension of the complex nature of the data. However, obtaining this understanding comes at the expense of cognitive effort and time required to process the data manually, which might lead to delayed decisions or errors [8].

Against this backdrop, interactive analytical dashboards have emerged in recent years as a third research direction [9]. Such dashboards recognize the limits of low human involvement with static dashboards, the high human involvement with interactive dashboards, and value the introduction of interactive analytical features [10]. The idea is to solve optimization problems automatically via computational approaches in the backend (analytical component) but to (still) involve the individual in the optimization process via the GUI (interactive component) [11]. Implementations range from robust trial-and-error to more advanced approaches such as interactive multi-objective optimization [12]. Yigitbasioglu and Velcu [4] and Pauwels et al. [9] argue that the promise of interactive analytical features such as a what-if analysis is unrealized in current dashboard solutions for operational DSS. A what-if analysis is a trial-and-error method that manipulates a dashboard's underlying optimization model to approximate a real-world problem [13]. However, by introducing an interactive analytical feature, operational decision-makers can interact with the optimization system without knowing the optimization model or understanding how the optimization procedure behind the results operates [14]. Such behavior can harm an individual's comprehension of the current situation, nourishing an out-of-the-loop problem—a situation in which a human being lacks sufficient knowledge of particular issues [15]. The fear is “that manual skills will deteriorate with lack of use” and individuals will no longer be able to work manually when needed ([15], p. 381).

Although what-if analyses have been investigated for DSS, empirical results have shown conflicting results. In particular, opposing patterns for the relationship toward decision performance have been reported in different studies. Benbasat and Dexter [16], as well as Sharda et al. [17], found a positive relationship. However, Fripp [18] and Goslar et al. [19] have demonstrated no statistically significant effect. In turn, others like Davis et al. [13] found that decision-makers performed better without a what-if analysis. With these conflicting results, it is surprising that academic literature gives relatively little attention to (1) how interactive analytical dashboard features should appear or (2) how they influence human beings' cognitive abilities in operational DSS [4]. An adequate level of situation awareness (SA) is emphasized as a promising starting point to examine this unfolding research gap [20]. Human factors research confirms that human beings' SA represents a key enabler for operational decision-making and task performance [21,22]. The importance of SA raises the question of how dashboard features can be designed to influence SA positively [23,24]. The objective of this study is, therefore, to examine the relationship between an interactive analytical dashboard feature, human beings' SA, and task performance in operational DSS. We implement a what-if analysis to study this relationship. Prominent references acknowledge that present dashboards “just perform status reporting, and development of [a] what-if capability would strongly enhance their value” ([9], p. 9). We rely on Endsley's [25] SA model as a theoretical guide for our analysis. It is established in academia and has shown favorable outcomes in different studies [21]. We formulate the following research question: *How does a what-if analysis feature in an interactive analytical dashboard influence operational decision-makers' situation awareness and task performance in operational DSS?*

The primary contribution of this work resides in linking the what-if analysis feature to SA and task performance. A second contribution resides in incorporating an analysis of eye-tracking parameters and the Situation Awareness Global Assessment Technique (SAGAT) in one of the first large-scale lab experiments ($n = 83$) on this topic to offer insights from a holistic viewpoint. A third contribution resides in the

translation of our findings into implications for designing dashboards within operational DSS to help practitioners in their efforts to address the danger of the out-of-the-loop syndrome.

We have chosen production planning as the study context, focusing on advanced planning and scheduling (APS), the synchronization of raw materials and production capacity (cf. section 4.2). First, the amount of APS data sets up severe challenges concerning the display resolution in dashboards [26]. Second, until now, few publications addressed dashboards and their influence in APS and production planning in general. For instance, Wu and Acharya [27] outlined a basic GUI design in the context of metal ingot casting. In addition, another study by Zhang [28] visualized large amounts of managerial data for evaluating its feasibility. There is a need to understand how planners effectively use dashboards because planning is not a static, one-time activity but rather requires continuous adaptation of vast data amounts. Third, this challenge is even more daunting because planners must increasingly make business-critical decisions in less time [26].

The ensuing section introduces the background related to SA theory and dashboards. The section thereafter develops the study's hypotheses. In the fourth section, we illustrate the research method. The data analysis and results are explained in the fifth section. The sixth section discusses the theoretical and practical implications. The final section concludes the study.

2. Background

2.1. Theory of situation awareness

In recent decades, SA has become one of the most discussed concepts in human factors research [24]. Although the original theoretical impetus of SA resided in military aviation, the concept has splashed over almost any area that addresses individuals executing tasks in complex and dynamic systems. We used the model by Endsley [25] as this model is one of the most prominent and widely used perspectives in the scientific community and debate around SA. Its robust and intuitive nature has been confirmed in different domains enabling scholars to measure the concept and derive corresponding requirements for DSS [21].

SA refers to an individual's knowledge about a specific situation [25]. It arises through his or her interaction with the environment. An adequate level of SA is known to affect subsequent decisions and actions positively. The resulting activities induce changes to the environment. As the environment changes, SA must be updated, which requires a cognitive effort of the respective individual. Due to this interaction, forming and maintaining SA can be difficult to achieve. Endsley [25] describes these difficulties as obstacles such as the out-of-the-loop problem. The resulting degree of SA is thereby often described as “high” or “low” [29]. Thus, we define SA as a quality criterion in terms of completeness and accuracy of the current state of knowledge, stating whether a human has an appropriate SA level.

However, measuring SA refers to a complex proposition. A plethora of approaches exists and academic discourse continues regarding which of those represent the most valid and reliable one to measure SA [29]. In general, research advocates a mixture of approaches, such as (1) freeze-probe or (2) process indices. In addition, most SA studies also measure (3) task performance indicators [29]. In the following, these three measurement approaches will be introduced.

2.1.1. Freeze-probe technique

Freeze-probe techniques involve random stops during task simulation and a set of SA queries concerning the current situation [29]. Participants must answer each query based on their knowledge of the current situation at the point of the stop. During each stop, all displays are blanked. The Situation Awareness Global Assessment Technique (SAGAT) is an established freeze-probe procedure that is in line with the model by Endsley [25]. The SAGAT approach has been specifically

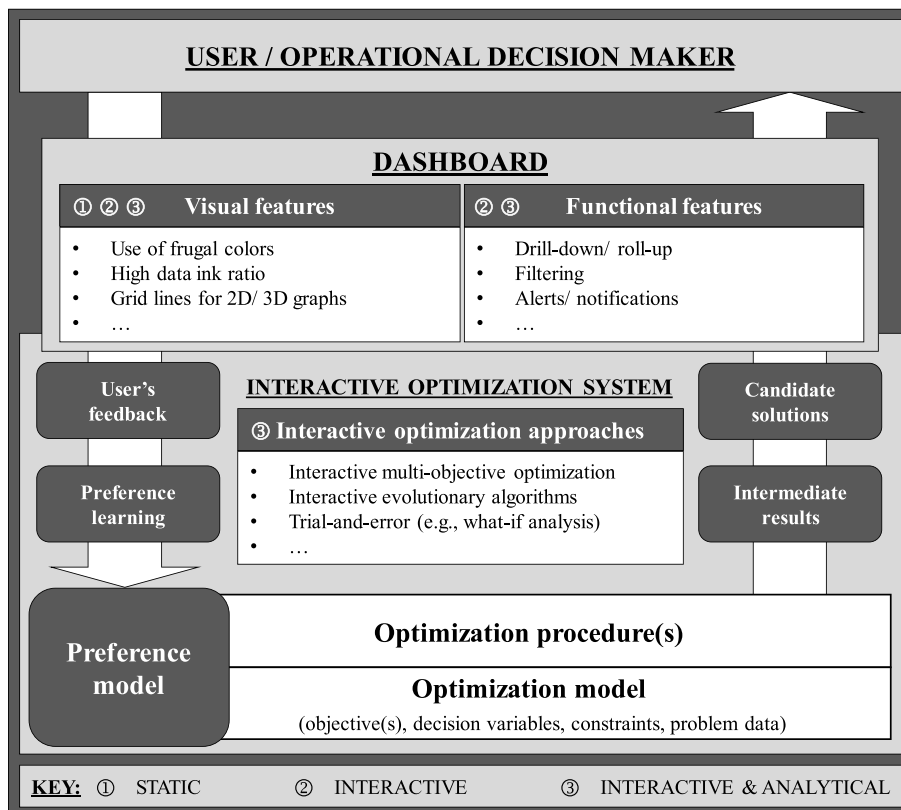


Fig. 1. Components of a Dashboard based on Meignan et al. [11].

created for the military and aviation domain. However, over the years, different SAGAT versions have been created in other fields, making this measurement approach one of the most popular ones to assess SA [29]. The benefits of this approach represent their alleged direct, objective nature, which eliminates the problems associated with assessing SA data post-trial [22]. This approach possesses a high validity, reliability, and user acceptance [29].

2.1.2. Process indices

Process indices assess cognitive processes of human beings to form and maintain SA during task execution [29]. Specifically, eye-tracking supports the definition of proficient process index measures and therefore refers to a further technique to measure SA [29]. Such devices capture fixations of human beings on the elements of the screen to indicate the degree to which such elements have been processed. This operationalization is based on the assumption that a high fixation level engages the participant in information-relevant thinking. Although this assumption represents only a proxy for true SA, it follows the same logic as attempts in other research communities that leverage gaze patterns to deduce attention or assign elaboration to memory or recall [30,31]. However, whereas freeze-probe techniques abound in academic discourse [29], only few investigations rely on process indices (such as eye-tracking) or study their reliability and validity to measure SA.

2.1.3. Task performance indicators

Most SA studies typically include task performance indicators [29]. Depending upon the task context, different performance measures are defined and collected. For instance, “hits” or “mission success” represent suitable measures in the military context. In turn, driving tasks could involve indicators like hazard detection, blocking car detection, or crash avoidance. Task completion, on the other hand, has been used as a measure in the study of user performance under the influence of instant messenger technologies [32]. Due to their non-invasiveness and

simplicity to collect, relying on task performance indicators appears beneficial [29].

2.2. Dashboard components and types

Classical business intelligence and analytics (BIA) refers to a data-centric approach promoting strategic and tactical decisions based on (mostly) retrospective analysis connected to a restricted audience of BIA experts and managers [33]. Nowadays, BIA refers to all methodologies, techniques, and technologies that capture, integrate, analyze, and present data to generate actionable insights for timely decision-making [34]. The transformation of raw data into actionable insights requires three BIA technologies: data warehouses (DWH), analytical tools, and reporting tools [35]. A DWH is a “subject-oriented, integrated, time-variant, nonvolatile collection of data in support of management’s decision-making process” ([36], p. 33). Typically, extract, transform, load (ETL) processes load data into a DWH, before analytical tools produce new insights. On this basis, dashboards are created to inform decision-making and next actions [37]. Hereby, dashboards are usually called the “tip of the iceberg” of BIA systems. In the backend, complex and (often not-well) integrated ETL infrastructure resides, attempting to retrieve data from diverse operational systems. The relevance of these backend systems is indisputable. However, dashboards appear to hold a significant role in this information processing chain because the interaction between operational decision-makers and DSS primarily occurs via GUIs [3]. Thus, the dashboard design is relevant for BIA systems due to its influence on an individual’s efficiency and productivity. However, whereas press articles [38,39] and textbooks [3,40] are abundant, few scientific studies have looked at dashboard features and their consequences for users; thus, limited guidance for scholars and practitioners is currently offered [4,9]. The handful of dashboard articles have studied intentions for their usage [41], case studies of instantiations used in organizations [39], different implementation stages [42],

adoption rates [43], and metrics selection [44]. In the following, we describe the components and types of a dashboard (cf. Fig. 1).

2.2.1. Visual features

In any dashboard type, different visual features are used to present data in a graphical form to reduce the time spent on understanding and perceiving them. Visual features relate to the efficiency and effectiveness of information presentation to users [4]. A good balance between information utility and visual complexity is therefore needed. Visual complexity refers to the level of difficulty in offering a verbal description of an image [45]. Studies show that varying surface styles and increasing numbers and ranges of objects increase visual complexity, whereas repetitive and uniform patterns decrease it [46]. Usually, dashboards leverage colors to differentiate objects from one another. The colors red and yellow, which are more likely to attract attention, are examples for signaling distinction [25]. Using such properties too often or inappropriately may lead to confusion or errors because the user would no longer be able to identify the critical information. Similarly, superfluous information within charts, so-called “chart junk” such as overwhelming 3D objects or non-value-adding frames, can appear impressive but severely distract users and induce undesired “visual clutter”. According to Tufte [47], a high data-ink ratio can reduce this problem. Data-ink ratio is a parameter that defines the relationship of ink used to illustrate data to the overall ink leveraged to represent the chart. A potential strategy refers to the erasing method by eliminating all components in charts with non-data ink. Some chart types promote visual illusions, which can potentially bias decision-making. Gridlines within charts are suggested as a useful technique to overcome such pitfalls [48]. A detailed discussion on visual features can be found in Few [3] and Tufte [47].

2.2.2. Functional features

Functional features indirectly link to visual features and comprise what a dashboard can do [4]. Functional features in the form of point and click interactivity enable the user to drill-down and roll-up. Traditional approaches to point and click interaction refer to merely clicking the mouse on the favored part of a list or chart to receive detailed information. Cleveland [49] initiates brushing, an approach that illustrates details of the data, by placing a mouse pointer over the data display. Such details might show information ranging in size from single values to information from related data points. Filtering is another valuable functional feature. It enables users to not only sort for relevant information but also identify hidden data relationships [7]. Alerts and notifications also describe functional features [7]. They can recommend corrective actions or make the user aware of issues as soon as measures no longer meet critical thresholds. Common influencing variables address type (e.g., warning or announcement), modality (e.g., visual, audio, or haptic), rate of frequency (i.e., how often to alert the user), and timing (i.e., when to induce the alert).

2.2.3. Interactive analytical features

Recent developments have emphasized the introduction of interactive analytical features within dashboards [10]. Such features (still) involve the operational decision-maker in the optimization process via the GUI (interactive component). Computational approaches and software tools solve the optimization problem in the backend (analytical component). Any optimization system contains an optimization model and optimization procedures to solve an optimization problem automatically. An optimization model describes an approximation of a real-world problem. Usually, it includes the definition of the objectives, decision variables, and constraints of the optimization problem. The problem data define the values of the parameter of the optimization model when the optimization problem is triggered. In turn, the optimization procedure solves the problem instances directly linked to the optimization problem by issuing either candidate solutions or intermediate results. An intermediate result is not necessarily a solution and

refers to information obtained during this optimization process. Candidate solutions are final solutions to an optimization problem [11]. Interactive optimization systems involve the user at some point in time, allowing him or her to modify the outcomes of the optimization system significantly. These systems leverage user feedback from the presented candidate solutions and intermediate results to adjust either the optimization model or the optimization procedures via a preference model. The preference model contains the preference information obtained from the user feedback [12]. Common examples of a preference model represent weight values in an objective function or heuristic information for the optimization procedure.

The user can refer to different interactive analytical features to offer various forms of feedback to the produced candidate solution or intermediate results. The user feedback can consist of adjustments of parameter values leveraging interactive multi-objective optimization. In an interactive evolutionary algorithm, the user can choose the most promising solution from a set of alternatives. Trial-and-error features are used for both modifying the optimization model and optimization procedures. The first approach corresponds to a what-if analysis in which the user modifies the data, constraints, or objectives of the optimization problem, whereas the optimization system offers a solution for evaluating these adjustments. In the second case, users modify some parameter values in an optimization procedure.

Although not exhaustively implemented, preference learning generalizes the feedback of a user to develop a model of his or her preferences. Such a model can be used to extend the optimization model (e.g., by adding new constraints). However, it is also possible to integrate user feedback into the preference model without preference learning. User values can be directly linked to the parameters of the optimization model. A detailed discussion of interactive analytical features can be found in Meignan et al. [11].

2.2.4. Types of dashboards

By now, numerous types of dashboards have been suggested not only for performance monitoring but also for advanced analytical purposes, incorporating consistency, planning, and communication activities [9]. Next, we consider the dashboard concept as a whole and describe it in terms of three different types.

Static dashboards are reporting tools used to summarize information into digestible graphical forms that offer at-a-glance visibility into business performance. The value resides in their capability to illustrate progressive performance improvements via fitting visual features to users. Depending upon the purpose and context, data updates can occur once a month, week, day, or even in real-time [3]. Static dashboards can fulfill both the urgency of fast-paced environments, offering real-time data support, and tracking of performance metrics against enterprise-wide strategic objectives based on historical data. However, such dashboards do not involve the user in the data visualization process and have issues with handling complex and multidimensional data [6]. Interactive dashboards can be considered a step toward directly involving the user in the analysis process. Interactive dashboards consider not only visual features, but also introduce functional features such as point and click interactivity [4]. These capabilities allow operational decision-makers to enable more elaborate analyses [6]. Addressing the impediments of static dashboards, they are used to establish a better understanding of the complex nature of data, which can also foster decision-making. However, the benefits of interactivity could increase the users' cognitive effort and required manual analysis time, increasing the probability of delayed decisions or (even) errors [8].

In more recent years, research has called for more interactive analytical dashboards to address the limits of low human involvement within static and high human involvement within interactive dashboards [10]. Such dashboard types rely on interactive analytical features [50]. The dashboard literature has confirmed and highlighted the potential value of a what-if analysis feature to quickly access and assess different aspects of the data in current dashboards [9]. Despite the good

prospects, these new types of features have (thus far) not materialized [4]. Potential obstacles concerning users trusting the interactive analytical functionality implicitly appear to remain [14]. Such behavior can induce adverse effects on SA that, in turn, can evolve into an out-of-the-loop problem [15].

3. Hypotheses development

This section describes the research model and the hypothesized effects. We expect different effects of the what-if analysis (in an interactive analytical dashboard) on an individual's SA and task performance. Furthermore, we study the effect of an interactive dashboard (without what-if analysis) to introduce a baseline for the interpretation of our empirical results. Finally, we study the effect of one control variable, an individual's mental workload because it has been acknowledged as a bias in DSS development [51].

3.1. Impact of dashboard design on situation awareness (H1)

Dashboard design can induce significant effects on operational decision-makers' SA because the set of dashboard features available defines the quantity and accuracy of information obtained by a human being [4]. In turn, SA defines the state of knowledge for perceiving and comprehending the current situation and projecting future events [25]. Such information appears to be important because it defines the basis to engage in effortful processing. Thus, the selection of dashboard features should be offered in a way that transmits the information required to the human being in the most effective manner, which (ideally) increases SA. We theorize that such transmission can be affected and optimized by the respective dashboard design and set of features employed. An interactive dashboard design assists operational decision-makers with manual data analysis [4]. The value of such designs stems from their capability to offer means for comparisons, reviewing extensive histories, and assessing performance. In contrast to static dashboards, interactive dashboards can go beyond "what is going on right now," enable filtering options, and drill down into causes of bottlenecks. Such dashboards allow operational decision-makers to move from general-abstract to more concrete-detailed information and thereby identify hidden data relationships from different perspectives [9]. Such data activities enable a detailed review of the current situational context.

Interactive analytical dashboards introduce interactive analytical functionality that facilitates operational decision-makers to simulate trends automatically. What-if analyses allow users to quickly receive system support (even) without the need to review information in a more detailed manner or to understand the overall business situation [4]. Such automated analysis can represent a useful feature to enable individuals immediate assessment of potential areas of improvement, specifically when the dashboard is used as a planning tool [52]. Operational decision-makers can leverage such interactive analytical features to obtain quick feedback on how specific changes in a variable (e.g., order fill rates) influence other values (e.g., profits). This way, they can emphasize the significance of bottleneck components [9].

We expect that the interactive dashboard design will outperform the interactive analytical one in terms of SA. Concerning the interactive analytical dashboard, exceptional planning situations remain. These cannot be handled properly by the optimization model. The real-world problem can involve constraints that are hard to quantify. Furthermore, the operational decision-maker must know several limitations of the quality and efficiency of the optimization procedure [11]. Thus, operational decision-makers (still) play a critical role in the optimization process and are required to be aware of the specifics of the planning situation [53]. However, by increasingly relying on a what-if analysis, planners can be deluded into no longer scrutinizing all objective and relevant information. Shibl et al. [14] showed that trusting DSS represented a success factor; however, the majority of users appeared to

trust the DSS implicitly or even "overtrust" the system. "As an operator's attention is limited, this is an effective coping strategy for dealing with excess demands" ([53], p. 3). The result, however, would be that such human-machine entanglement can lead to intricate experiences such as a loss of SA or (even) deteriorating skills [15].

Research shows that humans who rely on system support have difficulties in comprehending a situation once they recognize that a problem exists. Even when system support can operate correctly most of the time, when it does fail, the ability to restore manual control is crucial [54]. Likewise, a planner can rely on a what-if analysis without understanding the limits of the optimization model or the performance inadequacy behind the results, nourishing an out-of-the-loop problem. Many situations might exist for which neither the optimization model nor the optimization procedures have been implemented, trained, or tested [11]. Such situations require manual intervention. Interactive dashboards offer the possibility for manual data analysis (e.g., via point and click interactivity) to scrutinize all object-relevant information [9]. Typically, such actions are known to keep the human in the loop and aware of the situation [24].

In conclusion, a what-if analysis within an interactive analytical dashboard is an effective coping strategy to quickly show operational decision-makers potential areas of improvement. However, it simultaneously increases the possibility of getting out-of-the-loop. Thus, we theorize that the impact of an interactive dashboard on SA is higher than the impact of the interactive analytical counterpart.

H1. : An interactive analytical dashboard design leads to lower situation awareness than an interactive dashboard design without analytical features.

3.2. Impact of dashboard design on task performance (H2)

The value of dashboards resides in their capability to increase control and handling by the user and thus foster task performance [3]. Task performance relies on the degree to which an operational decision-maker scrutinizes the set of available information within a dashboard. This assumption is based on the importance of human beings' SA when assessing information [24]. However, when the universe of potential solutions is large or time constraints exist, operational decision-makers cannot reasonably assess every possible alternative [31]. They rather create a consideration set, a subgroup of alternatives, which the individual is prepared for and aware of to assess further [55]. Interactive dashboards can guide the planner through the current planning information. Based on the cognitive effort and time constraints of the planner, he or she will create a consideration set from the presented view. In case of breakdowns, interactive dashboards enable human beings to manually review detailed information and analyze their causes (if necessary) [9]. In production planning, these derived insights can be used to adjust the production plan.

Operational decision-makers can also rely on a what-if analysis to adjust their consideration set [10]. They can leverage this functionality to obtain quick feedback for specific relationships of variables (e.g., the effect of order fill rates on profit) [9]. Studies have acknowledged the potential value of such features to outline critical information that illustrates most of the difference [52].

However, we expect that the interactive analytical design will outperform the interactive one in terms of task performance. Interactive dashboards can represent a less effective means to promote task performance. The underlying reasoning is that manual analysis activities (as offered by interactive dashboards) are cognitively demanding and require time. Operational decision-makers must consider information critically and question its relevance before generating a carefully informed (although not necessarily unbiased) judgment [56]. However, due to the limits of the human mind with storing, assembling, or organizing units of information, operational decision-makers might not be able to transfer an (even) increased level of SA into their planning

activities and thus might experience lower task performance [54]. Common obstacles arise from restricted space or natural dissolution of information over time in human beings' working memory. Given abstract information, such dissolution can occur (even) in seconds. Conversely, with interactive analytical dashboards, the cognitive effort and time required are lower because planners can rely on interactive analytical features. For example, a what-if analysis can help to identify defective components rather than scrutinizing all of the information presented on its own [56]. Hence, we expect the following.

H2. : An interactive analytical dashboard design leads to higher task performance than an interactive dashboard design without analytical features.

3.3. Association between situation awareness on task performance (H3)

SA is an important antecedent to improving the likelihood to achieve higher task performance. Effective decisions require a good state of knowledge in terms of perception, understanding, and projection of the situation at hand [24]. Individuals experience more successful performance outcomes when they obtain a complete overview and knowledge of the current situation [54]. When issues emerge, they typically do so because some considerations of this overview are incomplete or incorrect [24]. The underlying reasoning is that a better SA can increase the control and handling of the system by the individual and thus contribute to task performance. Studies confirmed this relationship in different domains (e.g., automotive, military, or aviation). A decrease of SA is often correlated with reduced task performance, representing the most critical cause of aviation disasters in a review of over 200 incidents [57]. Hence, we hypothesize the following.

H3. : Situation awareness is associated with higher task performance.

Fig. 2 summarizes the overall research model.

4. Research method

We developed two dashboards. The interactive design assists the user with manual data analysis to filter, roll up, or drill down into the

causes of problems, whereas the interactive analytical one offers a what-if analysis on top. We conducted a laboratory experiment using a single-factor within-subject design to enhance statistical power for each setup and minimize error variance induced by individual differences [58]. We used a common technique to minimize the bias from potential carryover and order effects. We randomized the order of experimental sequences. First, we calculated the number of possible experimental sequences (i.e., 2). Second, we randomly allocated each participant to one of the two sequences. Randomization is a well-known technique to prevent unintentionally confounding the experimental design [59]. Before the data collection, we tested the setup to check on the intelligibility of the experimental tasks and to evaluate the measures. We included feedback rounds with practitioners to ensure closeness to reality.

4.1. Participants

We conducted our laboratory experiment with graduate students who were enrolled in an advanced IS course. In total, 83 graduate IS students (48 males, 35 females) took part. Their ages ranged from 20 to 34 years, with an average age of 24.1 years. We used students as a proxy for dashboard users in APS. We relied on students because similar laboratory experiments have shown that students are an appropriate population not normally biased by real-world experiences [60]. Second, the costs (e.g., financial compensation, physically presence at the campus) to incentivize students to take part in the experiment are relatively low. In turn, one of the main limitations of relying on professionals in experiments is that it is problematic to incentivize them [59]. Third, APS represents a challenging, complex, and cognitive-demanding task. IS students have learned the abilities or knowledge to apply it in complex contexts. Hence, these students seem to fulfill the necessary prerequisite for conducting the experimental task. In this line, we concur with the review by Katok [59] that students can perform as good as professionals do. Bolton et al. [61] confirmed this conclusion in a carefully executed experiment. Fourth, to account for the lack of system experience, we applied different techniques to prepare participants for using the dashboard (e.g., personal support, training tasks, small groups, help button). Because both dashboards participants

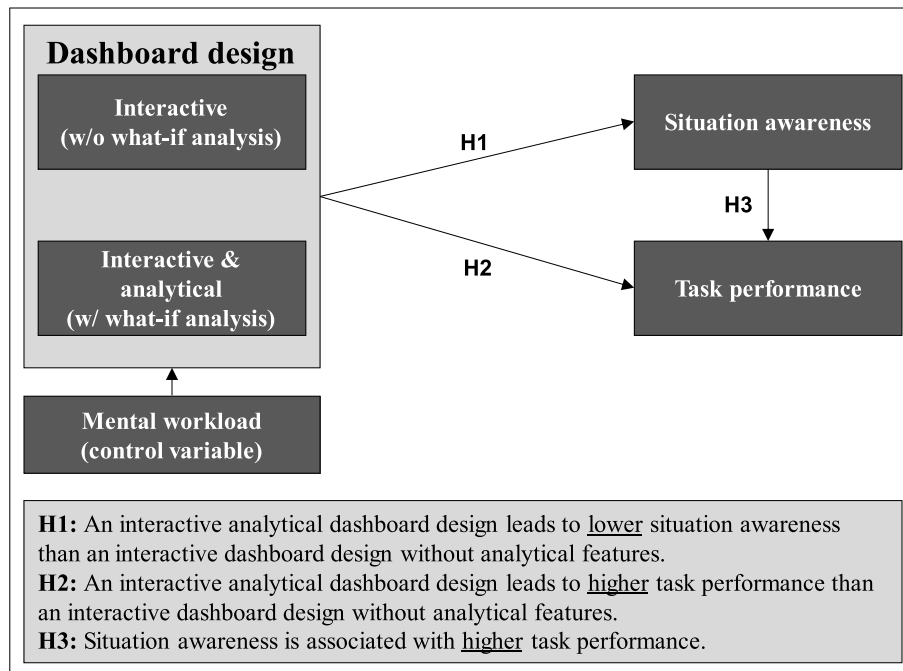


Fig. 2. Research model.

committed 2.5 errors on average, it appears that they had no difficulties in using them. Lastly, we discussed our study with two production planners and followed their experienced advice regarding several aspects, that is, how to assess the degree of SA and performance or how to design both dashboard types and determine appropriate areas of interest (AoI).

4.2. Experimental task

The problem type at hand was an APS task. Participants had to plan the production of bicycles in a fictional manufacturing company. Different components (e.g., bicycle handlebar, bell, lamp, reflectors, tire, chair) were available for each bicycle. Hereby, the amount of available components or which component was required for which bicycle could differ. Per planning week, only a limited amount of components was available. In addition, for each product, the minimal amount of planned customer demands of all planning weeks was offered. Customer demands represented the basis for the calculation of needed components per planning week. In total, four planning weeks had to be planned by the participants so that all constraints were met and the total revenue was maximized. Three parameters were critical to generate an effective production plan: (i) the number of available components, (ii) the number of products to be produced (demand), and (iii) information about which component is linked to which product. Both designs shared the same APS process and objective to create an optimal production plan solving a planning problem within the given constraints. Both plans comprised initial states and planning restrictions. The initial state was characterized by a set of raw data that did not comply with the planning restrictions. For each plan, all decision packages shared the same complexity. The initial raw data for both plans only differed in their conditions, such as component numbers or weekly demand. We introduced constraints to ensure that users could execute the tasks within an appropriate amount of time: component shortage represented the only restriction for planning, whereas human and machine resources were assumed unrestricted. The participants were required to create a production plan by modifying the changeable raw data with (w/) and without (w/o) a what-if analysis.

4.3. Treatment conditions

Fig. 3 shows a screenshot of the interactive dashboard design. The design comprises three elements: i) a toolbar to navigate through different areas of interest at the very top, ii) a navigation bar top right to perform different means of analysis and iii) a content area showing the respective results. To support a holistic view of the as-is planning

situation, the planner obtains access to supply and retail data within sortable tables and bar charts. The sortable tables and bar charts are combined with drill-down, roll-up, and filter functionality to explore the underlying sets of data deeply.

Concerning the what-if analysis, the interactive analytical dashboard design introduces a simulation area capable of visualizing bottleneck components and their criticality for the production plan (cf. Fig. 4). The bar chart uses colorful indications (i.e., green/ red) to assess the effects and criticality of bottleneck components. In this way, it emphasizes immediately whether the current data entries form a valid production plan. In the event of a component shortage or surplus, the chart is colored in red. The difference between needed and possible orders is visualized as a number within the bar chart.

4.4. Experimental procedure

The study lasted for approximately one hour and three minutes and comprised three sections: an introduction (10 min), training (25 min), and the actual testing section (28 min). Upon arrival, participants started by signing the information consent form. The computer lab was prepared with running Tobii pro X2-30 eye-tracker equipment. The eye-tracker recorded the participant's eye movements as X and Y coordinates (in pixels) at intervals of 30 ms, applying dark-pupil infrared technology and video. Participants were introduced to the production-planning problem and the dashboard design by the experiment instructor. They had the opportunity to go through the problem for some minutes to become familiar with the topic. The training section followed the introduction, aiming to practice the generation of a production plan with both alternatives. The training sections comprised small groups of three to five participants to offer participants the possibility to ask questions individually. The instructor and a help menu offered by the application provided guidance and support. Participants could use as much time as they needed; they required approximately 25 min to finish both training activities (one for each design). In the testing section, each participant finished one session for each design. Each session lasted a maximum of 14 min.

4.5. Measurement

4.5.1. Situation awareness

We employed the SAGAT freeze-probe technique to measure SA. In our study, we halted the task four times within each alternative and blanked out the screen. A message on the screen asked the participants to respond to questions related to their perceptions of the situation in order to obtain a periodic SA snapshot. Participants answered three

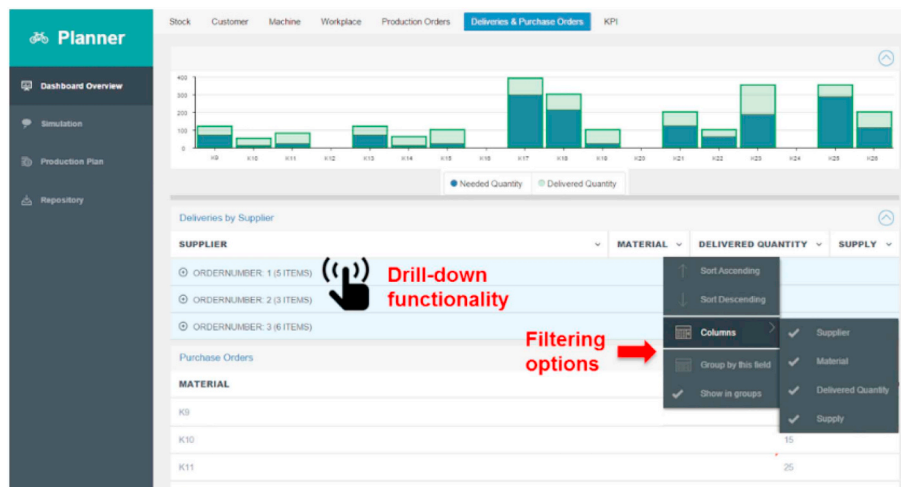


Fig. 3. Screenshot of the interactive dashboard design.

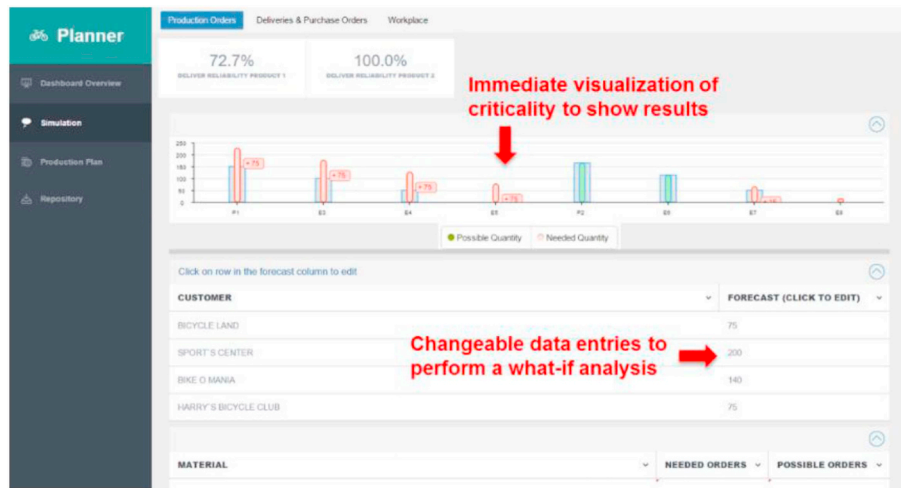


Fig. 4. Screenshot of the interactive analytical dashboard design.

Table 1
SAGAT questionnaire.

1.	What is the current quantity of P1 in the stock?
2.	What is the current size of the purchase order for K26?
3.	What is the current size of the purchase order for K9?
4.	The production of which product is in danger?
5.	Select the machine(s) which have problems:
6.	Enter the current Delivery Reliability:
7.	Enter the supplier which does deliver purchased part K13:
8.	In comparison to the current state, what would happen to Delivery Reliability if machine 1 broke down?
9.	What is the needed amount of purchased part E9 to fulfill all production orders?
10.	What would you do concerning the production orders?
11.	What would you do concerning the purchased parts?
12.	What would you do concerning the workplace capacity?

questions at each step (cf. Table 1).

To evaluate whether participants' replies were correct, we contrasted them with logged data at the moment of the freeze. We analyzed the SA scores as the percentage of correct answers for each dashboard.

Eye-Tracking. Two AoIs for each dashboard feature were defined. An AoI refers to the “area of a display or visual environment that is of interest to the research or design team” ([62], p. 584). We operationalized SA via two eye-tracking measures: (i) participant's fixation duration and (ii) fixation count spent scrutinizing the information of the respective AoI. A fixation occurs if an eye concentrates on a particular point for a specific period in time. Any eye-movement around 2° of visual arc for at least 60 ms in time is called a fixation [62]. Fixation count refers to the total number of fixations counted in an AoI. Fixation duration is defined as the average fixation time spent on an AoI. By studying the fixation duration and fixation count with both designs, we can conclude which alternative drew more attention.

Task Performance. We measured task performance by the number of errors committed for each dashboard alternative. Participants could save their production plan when they felt they had arrived at a suitable solution. If a participant did not successfully manage to save the plan, we took the last version.

Mental Workload. We controlled for individuals' perceived mental workload because it induces a direct effect on their ability to execute tasks. Hence it can influence the effectiveness and efficiency of interactions with computers or GUIs [51]. We have employed the NASA-TLX based on Hart and Staveland [63] to measure perceived mental workload. This index possesses high validity and refers to “a multi-dimensional, self-reported assessment technique that provides an estimation of the overall workload associated with task performance” ([51], p. 72). Table 2 summarizes the corresponding items.

Participants rated each subscale immediately after completing a production plan with the respective dashboard type. We used a Wilcoxon signed-rank test to identify whether there is a significant difference in mental workload for both alternatives. This non-parametric test compares two related samples on non-normally-distributed data (cf. Table 3).

The test showed no statistical significance ($p > .1$), confirming that the perceived mental workload between both designs was comparable. Table 4 summarizes the results.

5. Data analyses and results

5.1. Normality assumption

We used the Shapiro-Wilk statistic to test the data for assumptions of normality. As the p -values were less than the depicted alpha level ($\alpha = .1$), the null hypothesis was rejected (cf. Table 5). Hence, we applied the Wilcoxon signed-rank test statistic. This test is capable of obtaining the non-normally-distributed data of two related samples. This way, we could test our hypothesis in terms of participants' SA (measured by SAGAT score, fixation duration, and fixation count) and task performance (measured by errors committed).

5.2. Impact of dashboard design on situation awareness (H1)

In the following, we assess the descriptive statistics for the constructs employed, compare their statistical significance and calculate the corresponding effect sizes. Our results indicate that the interactive analytical dashboard was able to outperform the interactive design in terms of task performance. Furthermore, the participants' fixation count, duration, and SA scores were significantly higher for the interactive alternative. All participants answered more than 50% of the SAGAT questions correctly (compared with their maximum values) for both alternatives. Thus, both designs appeared to achieve sufficient SA levels. We confirmed hypothesis H1 because the interactive design (Mean = 0.53, Standard Deviation = 0.18) outperformed the counterpart (M = 0.39, SD = 0.16) on the SA scores. The Wilcoxon signed-rank test revealed a significant difference ($p = .000$) with a medium effect size r of 0.306.

The collected fixation duration and fixation count for each alternative were subjected to a Wilcoxon signed-rank test. Participants fixated more on the interactive design (M = 716.34, SD = 314.59) as compared to the interactive analytical screen (M = 335.70; SD = 185.12). The difference between both designs was significant, with a p -value less than .001 with a large effect size r of 0.549.

Table 2
NASA-TLX questionnaire.

Mental demand	How mentally demanding was the task?
Temporal demand	How hurried or rushed was the pace of the task?
Performance	How successful were you in accomplishing what you were asked to do?
Effort	How hard did you have to work to accomplish your level of performance?
Frustration	How insecure, discouraged, irritated, stressed, and annoyed were you?

Table 3
Testing distributions of normality by treatment for mental workload.

Term	Skewness	Kurtosis	Shapiro-Wilk test	
			statistic	p-value
Interactive (w/o what-if analysis)				
Mental demand	-0.653	-0.015	0.953	0.004**
Temporal demand	-0.276	-0.933	0.954	0.005**
Performance	-0.479	-0.685	0.951	0.003**
Effort	-0.157	-0.862	0.960	0.011*
Frustration	-0.605	-0.155	0.953	0.004**
Interactive-analytical (w/ what-if analysis)				
Mental demand	-0.671	0.939	0.939	0.001**
Temporal demand	-0.167	-0.979	0.958	0.008**
Performance	-0.613	-0.184	0.946	0.002**
Effort	-0.046	-1.231	0.938	0.001**
Frustration	-0.579	-0.586	0.931	0.000***

*** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$, † $p < 0.1$, ns = not significant.

Similarly, the interactive design ($M = 208.26$; $SD = 102.29$) outperformed the interactive analytical design ($M = 75.97$; $SD = 49.88$) on average in terms of fixation duration. Analyses indicated that this difference was significant ($p = .000$) with a large effect size r of 0.584. Thus, eye-fixation data also confirmed that the interactive design supported higher levels of participants' SA, compared to the interactive analytical design (H1).

5.3. Impact of dashboard design on task performance (H2)

In turn, participants produced more errors ($M = 2.82$; $SD = 1.08$) within the interactive design when trying to create a production plan. They committed fewer errors ($M = 2.18$; $SD = 1.31$) within the interactive analytical design. In line with our assumptions, this difference was significant between both dashboards ($p = .000$; H2), however only with a small effect size r of 0.295. This seems in line with the descriptive statistics as, on average, both dashboard alternatives produced approximately two to three errors per plan, indicating a low degree of errors committed during planning in general. Both dashboards appeared to achieve appropriate scores of task performance (cf. Table 6).

5.4. Association between situation awareness and task performance (H3)

On average, the correlation analysis by Spearman's Rho showed that participants' fixation duration shared a significantly positive effect on the associated SAGAT scores for both alternatives ($p < .05$). Thus, the

Table 4
Wilcoxon signed-rank test by treatment for mental workload.

NASA-TLX index	Experiment treatment		Wilcoxon signed-rank test		
	Interactive (Mean, SD)	Interactive & analytical (Mean, SD)	U	r	p-value
Mental demand	62.3 (21.7)	59.0 (23.3)	-1.025	0.0796	0.305 (ns)
Temporal demand	51.6 (28.3)	49.1 (27.2)	-0.945	0.0733	0.345 (ns)
Performance	58.4 (26.6)	61.9 (25.5)	-1.248	0.0969	0.212 (ns)
Effort	47.7 (26.9)	43.1 (28.0)	-1.211	0.0940	0.226 (ns)
Frustration	60.4 (22.1)	56.4 (23.8)	-1.485	0.1153	0.137 (ns)

*** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$, † $p < 0.1$, ns = not significant.

Table 5
Testing distributions of normality of constructs by treatment.

Term	Skewness	Kurtosis	Shapiro-Wilk test	
			Statistic	p-value
Interactive (w/o what-if analysis)				
Fixation duration	-0.263	-0.823	0.971	0.058†
Fixation count	-0.508	-0.349	0.964	0.020*
SAGAT	0.369	-1.383	0.821	0.000***
Errors committed	-1.103	1.101	0.821	0.000***
Interactive-analytical (w/ what-if analysis)				
Fixation duration	1.212	2.050	0.920	0.000***
Fixation count	0.770	1.079	0.959	0.010*
SAGAT	0.569	-0.991	0.831	0.000***
Errors committed	-0.444	-0.885	0.876	0.000***

*** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$, † $p < 0.1$, ns = not significant.

more time participants spent on the screen, the higher their SAGAT scores became. The effect size of the interactive design ($r = 0.313$) was larger than for interactive analytical design ($r = 0.296$). Concerning the participant's fixation count, the correlation analysis by Spearman's Rho confirmed a significantly positive correlation of the SAGAT scores for each design alternative, constituting a larger effect size within the interactive analytical design ($r = 0.305$). The interactive design showed only weak significance ($p < .1$). In summary, the correlation analysis for both alternatives indicated that a participant's higher fixation duration (count) led to a greater ability to answer the SAGAT questionnaire correctly (cf. Table 7).

In terms of the relationship between fixation duration (count) and the performance indicator errors committed, we found a significant negative correlation for the interactive analytical design. The interactive analytical design shared a high effect size for both fixation duration ($r = -0.435$) and fixation count ($r = -0.366$). In other words, participants made fewer errors during their planning efforts, the longer they looked at and the more they fixated on the interactive analytical dashboard screen. For the interactive design, participants' fixation duration (count) did not significantly correlate with errors committed.

Concerning SAGAT scores and errors committed, the correlation analysis confirmed the expected negative correlation for each alternative, constituting a larger effect for the interactive analytical design ($r = -0.421$; $p = .004$). In summary, for the SAGAT scores, the results confirmed that a participant's higher SA led to higher task performance with both designs (H3). However, the eye-tracking data confirmed this relationship only for the interactive analytical design (cf. Table 7).

Table 6
Descriptive statistics of constructs by treatment.

Term	Experiment treatment				Wilcoxon signed-rank test		
	Interactive (Mean, SD)		Interactive & analytical (Mean, SD)		Z	r	p-value
SAGAT (in %)	0.53 (0.18)	Min. 0.17 Max. 0.92	0.39 (0.16)	Min. 0.13 Max. 0.75	-3.946	0.306	0.000***
Fixation duration (in msec)	208.26 (102.29)		75.97 (49.88)		-7.528	0.584	0.000***
Fixation count (in no.)	716.34 (314.59)		335.70 (185.12)		-7.074	0.549	0.000***
Errors committed (in no.)	2.82 (1.08)		2.18 (1.31)		-3.804	0.295	0.000***

*** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$, † $p < 0.1$, ns = not significant.

Table 7
Construct correlations by treatment.

Term	[1]	[2]	[3]	[4]
Interactive (w/o what-if analysis)				
[1] Fixation duration	-			
[2] Fixation count	0.812***	-		
[3] SAGAT	0.313*	0.253†	-	
[4] Errors committed	-0.125 (ns)	-0.062 (ns)	-0.311*	-
Interactive & analytical (w/ what-if analysis)				
[1] Fixation duration	-			
[2] Fixation count	0.935***	-		
[3] SAGAT	0.296*	0.305*	-	
[4] Errors committed	-0.435***	-0.366**	-0.421**	-

*** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$, † $p < 0.1$, ns = not significant.

Table 8 summarizes the results of the hypotheses.

6. Discussion

6.1. Theoretical implications

There are four central theoretical implications we want to emphasize. First, our analyses indicated that the interactive analytical dashboard showed significantly higher task performance. This result confirms the value of interactive analytical dashboards and encourages the trend toward introducing interactive analytical features for operational decision-making. We showed that a what-if analysis has the potential to direct participants to the most critical pieces of information such as bottleneck components that require immediate attention. Obstacles to the what-if analysis included a lack of knowledge of operational decision-makers of the current planning situation, which, in turn, resulted in a lower degree of SA (compared with the interactive design). However, our results indicated that if the underlying optimization model and optimization procedure can account for the planning situation (as in our experiment), high-performance outcomes can be possible even in the absence of high participant SA levels. When a planner searches for optimization potential by identifying bottleneck components within production data, the optimization system leads the participant to certain areas of interest. Such focusing, by definition, removes awareness of the ongoing production context, decreasing the overall SA. In conclusion, these cases appear beneficial for task performance as long as the employed optimization model and optimization procedures behind the what-if analysis are capable of addressing the occurring

situation.

Second, our analyses revealed that the introduction of interactive analytical features is not cost-free. Participants who used the interactive dashboard maintained a significantly higher SA when creating a production plan, whereas the interactive analytical design led to a lower degree of SA. Hence, participants who used the interactive analytical variant appeared to follow the system support (through the results proposed by the what-if analysis). Subsequently, they reduced their effort and SA to scrutinize all object-relevant information. These findings illustrate the danger for operational decision-makers of trusting the system results implicitly or (even) of “overtrust”, leading to potentially adverse effects such as an out-of-the-loop problem [53]. With increasing simulation abilities, operational decision-makers can become less aware and slower acting in recognizing problems on their own. Thus, if problems require manual intervention, further time is necessary to comprehend “what is going on”, which sets boundaries on a quick problem resolution. Following this argumentation, eye-tracking studies [64] confirmed that continuous use of simulation reduces individuals’ cognitive skill set because such skills rapidly deteriorate in the absence of practice. These effects can lead to an increasing tool dependence and problems when the underlying optimization model or procedures are not implemented, trained, or tested for the current situation [11].

Third, our correlation analysis showed the link of SA to task performance concerning the SAGAT scores. This finding is in line with extant conceptual arguments as task performance is expected to correlate with SA [54]. Still, in comparison to our Wilcoxon signed-rank tests for both dashboards, these findings might suggest that tension between maintaining high SA and accomplishing high task performance exists. We conclude that high levels of SA are necessary in general but do not suffice for high task performance. Although SA is relevant for decision-making, different influencing variables might be involved in converting SA into successful performance while “it is possible to make wrong decisions with good SA and good decisions with poor SA” ([22], p. 498). For instance, a human with high SA of an error-prone system might not be sufficiently knowledgeable to correct the error or lack the skills needed to trigger that remedy. Due to the good task performance scores for both dashboards, our results, however, do not indicate a problem with regards to the knowledge or skill set of the participants. Thus, future work could delve deeper into other influencing variables that might be involved in converting SA into task performance.

Finally, assessing the SA of operational decision-makers describes a key component to design and evaluate DSS [24]. Comparable with most other ergonomic constructs (e.g., mental workload and human error), a plethora of methods for measurement exists [22]. However, studies

Table 8
Summary of hypotheses.

Hypothesis	Result
H1: An interactive analytical dashboard design leads to <u>lower</u> situation awareness than an interactive dashboard design without analytical features.	Supported
H2: An interactive analytical dashboard design leads to <u>higher</u> task performance than an interactive dashboard design without analytical features.	Supported
H3: Situation awareness is associated with <u>higher</u> task performance.	Supported ^a

^a The SAGAT confirmed the relationship for both designs, whereas the eye-tracking data confirmed this only for the interactive analytical design.

have not agreed upon which of the techniques available represents the most appropriate for assessing SA [65]. This study compared two different measures to assess participants' SA: (i) the established freeze-probe approach (leveraging SAGAT) and (ii) the rather scarcely addressed assessment of SA via eye-tracking [22]. The results indicated that there was a significant correlation between the constructs employed for both methods (i.e., eye-tracking and SAGAT scores). This finding suggests that those approaches were measuring similar aspects of participants' SA during task execution. More interestingly, the analyses showed that only the participants' SAGAT scores derived via the freeze-probe method generated significant correlations with task performance for both dashboard designs. Our eye-tracking data only confirmed the relationship of fixation duration (count) toward errors committed for the interactive analytical dashboard. Hence, the freeze-probe technique seems to represent a stronger predictor of task performance compared with the eye-tracking assessment of SA. This suggests that the rating type in use appears to be a determinant of the respective measure's predictability. In closing, our analyses supported the usage of the freeze-probe technique to assess SA during planning tasks. The findings can be contrasted with previous research that confirmed that SAGAT is one of the most appropriate methods to use when the experimenter knows what SA should comprise beforehand [22].

6.2. Practical implications

Next, we translate our findings into implications for designing dashboards to help practitioners in their efforts to address the danger of the out-of-the-loop syndrome. First, our analyses showed evidence to support design efforts that advocate a higher degree of automation within analytical tools. The underlying reasoning is to provide operational decision-makers relief from their cognitively demanding task by introducing more realistic optimization models and more efficient optimization procedures. This design trend accounts for the challenges decision-makers are facing at the operational level. For instance, they must make increasingly business-critical decisions in a shorter period with exponentially growing amounts of data. In such situations, operational decision-makers experience increasingly high mental workload in their attempt to process data with their own working memory [8] – a rather limited, laggard, and error-prone resource [66]. Over the long term, this design trend drives continuously reducing the human level of authority in the optimization process. The ultimate goal is to diminish the role of human interaction and allow most decision-making procedures to occur automatically in the backend of the system.

However, inherent limitations remain on the integration of optimization systems into DSS such as dashboards. One issue refers to the difficulty of constructing an optimization model that accounts for all aspects of the real problem (i.e., diversity/ number of constraints). In other cases, the problem is not completely specified due to a lack of context knowledge or must be (over-)simplified to be appropriate for the computational optimization approach. Other issues relate to the optimization procedure, which might not be appropriately parameterized before leveraging the system under real circumstances. Furthermore, the performance of the optimization method might not fit the real user requirements. Thus, the importance of having more realistic optimization models and more efficient optimization procedures is indisputable; however, neither of these currently appears to be sufficient [11].

Second, an alternative approach to interpreting our results might be to increase the design efforts to keep the human in the loop of the current situation. Such a trend endorses a higher interaction of operational decision-makers with interactive analytical dashboards. The underlying reasoning is to realize an effective human-computer interaction but simultaneously consider the cognitive abilities of individuals. Our analyses showed that the degree of SA human beings possess defines a relevant factor in such a system design. One promising design attempt might be to introduce a periodic SA audit to increase the

cognitive engagement of an operational decision-maker. Gaze-aware feedback could implement such an audit. This technique has already shown promising improvements in operational decision-makers' visual scanning behavior. Sharma [67] revealed that gaze-aware feedback showed beneficial effects for both a teacher's visual attention allocation and students' gain in knowledge in an e-learning context. However, despite the promising results using eye-tracking data to assess participants' SA, our analysis raised (at least) partial concerns concerning the validity of eye-tracking data to predict the number of errors.

An alternative design attempt is to make information more transparent to operational decision-makers. This information should concern the underlying optimization model (e.g., used problem specifications) and optimization procedures (e.g., parameter setting) behind the interactive analytical dashboard. Research reports that the effect on SA was improved to a substantial degree by the transparency of the optimization system, thus offering understandability of its actions [54]. Such transparency can help individuals to understand the cause of actions and reduce the downside effects of the out-of-the-loop syndrome.

Finally, training should highlight the significance of frequent visual attention scanning to prevent a loss of SA or skills. In other fields, individuals are required to periodically perform in a manual mode to maintain their skills and appropriate SA levels [68]. The measurement of SA is thus key not only to improve SA-related theory but also to advance the design of training and evaluation efforts for practitioners. Hence, scholars demand valid and reliable methods to verify and enhance SA theory. System designers require a means to ensure that SA is increased by new features, GUIs, or training programs [54].

7. Conclusion

Dashboards are important for DSS as they have a significant impact on their effectiveness, particularly at the operational level. Our study objective was to examine the effects of an interactive analytical dashboard feature, a what-if analysis. To date, little is understood about its influence on human cognitive abilities. We argued that designing such a dashboard feature requires a profound understanding of SA because a lack of awareness is known to interfere with human information processing. Further, it entails downstream effects on the human ability to make informed decisions. We created a model that relates a what-if analysis to support SA that, in turn, would positively affect task performance. We conducted one large-scale laboratory experiment, including eye-tracking and SAGAT data, to study this model from a holistic viewpoint.

The significance of our study contributes to DSS literature in several ways. Our results indicate that interactive analytical features induce significant effects on task performance. We showed that the introduction of a what-if analysis could reduce the role of required SA for the operational decision-maker. However, such cases appear only beneficial for task performance as long as the underlying optimization model and optimization procedures behind the dashboard address the occurring situation. We also reported that the implementation of a what-if analysis is not cost-free and can trigger adverse effects such as a loss of SA. This finding confirms the notion of the out-of-the-loop syndrome applying in the APS context and extends it to the dashboard feature level. In addition, our correlation analyses indicated the expected link of SA to task performance. Although multiple studies have shown the relevance of SA for decision-making, we illustrated that low task performance could result from good SA and vice versa. Thus, we believe that other influencing variables could be involved in converting SA into successful performance that require further assessment. However, how SA is assessed is also fundamental because correlations within both dashboard alternatives differed according to the measurements concerning the link of SA to task performance. The SAGAT scores confirmed this relationship for both alternatives, whereas the eye-tracking data showed significant correlations only for the interactive analytical design. We, therefore, provided evidence that using SAGAT seems most

appropriate.

Although our results appear promising, some limitations require discussion. We used graduate students as proxies for planners. Thus, other studies could verify how professionals interact with the proposed designs. We studied the effects of a what-if analysis, which we expected to play a key role in the creation and reduction of SA. Future work might consider further interactive analytical features, such as interactive multi-objective optimization [12]. Other approaches involve the user at different points in time (i.e., user feedback on candidate solutions vs. on intermediate results), leverage the feedback loops with the user for different parts of the preference model (i.e., adjustments of the optimization model vs. optimization procedures), and occasionally include preference learning. It might be interesting to compare individuals' SA and task performance scores obtained with those features.

Declarations of Competing Interest

None.

Funding

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Acknowledgement

We thank Tobias Hochstatter for his support.

References

- [1] E. Dong, H. Du, L. Gardner, An interactive web-based dashboard to track COVID-19 in real time, *Correspondence*. 20 (2020) 533–534.
- [2] K. Gander, Here Are 6 Coronavirus Dashboards Where You Can Track the Spread of COVID-19 Live Online, *Newsweek*, <https://www.newsweek.com/coronavirus-tracking-maps-1491705>, (2020) (accessed May 9, 2020).
- [3] S. Few, *Information Dashboard Design*, O'Reilly Media, Inc., *The Effective Visual Communication of Data*, 2006.
- [4] O.M. Yigitbasioglu, O. Velcu, A review of dashboards in performance management: implications for design and research, *Int. J. Account. Inf. Syst.* 13 (2012) 41–59.
- [5] S. Negash, P. Gray, *Business Intelligence*, F. Burstein (Ed.), *Int. Handbooks Inf. Syst. Handb. Decis. Support Syst.* 2 Heidelberg, Springer, Berlin, 2008, pp. 175–193.
- [6] J. Kohlhammer, D.U. Proff, A. Wiener, *Visual Business Analytics - Effektiver Zugang zu Daten und Informationen*, dpunkt Verlag GmbH, Heidelberg, (2013).
- [7] W.W. Eckerson, *Performance Dashboards: Measuring, Monitoring and Managing Your Business*, John Wiley & Sons, Hoboken, 2006.
- [8] F.J. Lerch, D. Harter, Cognitive support for real-time dynamic decision making, *Inf. Syst. Res.* 12 (2001) 63–82.
- [9] K. Pauwels, T. Ambler, B.H. Clark, P. LaPointe, D. Reibstein, B. Skiera, B. Wierenga, T. Wiesel, Dashboards as a service: why, what, how, and what research is needed? *J. Serv. Res.* 12 (2009) 175–189.
- [10] V.A. Zeithaml, R.N. Bolton, J. Deighton, T.L. Keiningham, K.N. Lemon, J.A. Petersen, Forward-looking focus: can firms have adaptive foresight? *J. Serv. Res.* 9 (2006) 168–183.
- [11] D. Meignan, S. Knust, J.-M. Frayret, G. Pesant, N. Gaud, A review and taxonomy of interactive optimization methods in operations research, *ACM Trans. Interact. Intell. Syst.* 5 (2015) 1–46.
- [12] K. Miettinen, F. Ruiz, A.P. Wierzbicki, Introduction to multiobjective optimization: Interactive approaches, in: J. Branke, K. Deb, K. Miettinen, R. Slowinski (Eds.), *Multiobjective Optim. Lect. Notes Comput. Sci.* Springer, Berlin, Heidelberg, 2008, pp. 27–58.
- [13] F.D. Davis, J.E. Kottemann, W.E. Remus, What-if analysis and the illusion of control, *Proc. Hawaii Int. Conf. Syst. Sci.*, IEEE Comput. Soc. Press, 1991, pp. 452–460.
- [14] R. Shibl, M. Lawley, J. Debusse, Factors influencing decision support system acceptance, *Decis. Support. Syst.* 54 (2013) 953–961.
- [15] M.R. Endsley, E.O. Kiris, The out-of-the-loop performance problem and level of control in automation, *Hum. Factors* 37 (1995) 381–394.
- [16] I. Benbasat, A.S. Dexter, Individual differences in the use of decision support aids, *J. Account. Res.* 20 (1982) 1–11.
- [17] R. Sharda, S.H. Barr, J.C. McDonnell, Decision support system effectiveness: a review and an empirical test, *Manag. Sci.* 34 (1988) 139–159.
- [18] J. Frripp, How effective are models? *Omega*. 13 (1985) 19–28.
- [19] M.D. Goslar, G.I. Green, T.H. Hughes, Decision support systems: an empirical assessment for decision making, *Decis. Sci.* 17 (1986) 79–91.
- [20] D. Manzey, J. Reichenbach, L. Onnasch, Human performance consequences of automated decision aids: the impact of degree of automation and system experience, *J. Cogn. Eng. Decis. Mak.* 6 (2012) 57–87.
- [21] D. Golightly, J.R. Wilson, E. Lowe, S. Sharples, The role of situation awareness for understanding signalling and control in rail operations, *Theor. Issues Ergon. Sci.* 11 (2010) 84–98.
- [22] P.M. Salmon, N.A. Stanton, W.G.H. D. D. Ladva Jenkins, L. Rafferty, M. Young, Measuring Situation Awareness in complex systems: Comparison of measures study, *Int. J. Ind. Ergon.* 39 (2009) 490–500.
- [23] M.L. Resnick, Situation awareness applications to executive dashboard design, *Hum. Factors Ergon. Soc. Annu. Meet. Proc.* 47 (2003) 449–453.
- [24] C.D. Wickens, Situation awareness: review of Mica Endsley's 1995 articles on situation awareness theory and measurement, *Hum. Factors* 50 (2008) 397–403.
- [25] M.R. Endsley, Toward a theory of situation awareness in dynamic systems, *Hum. Factors* 37 (1995) 32–64.
- [26] H. Stadler, Production planning and scheduling, in: H. Stadler, C. Kilger, H. Meyr (Eds.), *Supply Chain Manag. Adv. Plan.*, Springer, Berlin, Heidelberg, 2015, pp. 195–211.
- [27] P.Y. Wu, S. Acharya, Visualizing Capacity and Load: A Production Planning Information System for Metal Ingot Casting, in: *Proc. Conf. Inf. Syst. Appl. Res. CONISAR Proceedings*, Wilmington, 2011, pp. 1–6.
- [28] P. Zhang, Visualizing production planning data, *IEEE Comput. Graph. Appl.* 16 (1996) 7–10.
- [29] P. Salmon, N. Stanton, G. Walker, D. Green, Situation awareness measurement: a review of applicability for C4i environments, *Appl. Ergon.* 37 (2006) 225–238.
- [30] M.G. Calvo, Working memory and inferences: evidence from eye fixations during reading, *Memory*. 9 (2001) 365–381.
- [31] T.O. Meservy, M.L. Jensen, K.J. Fadel, Evaluation of competing candidate solutions in electronic networks of practice, *Inf. Syst. Res.* 25 (2014) 15–34.
- [32] A. Gupta, H. Li, R. Sharda, Should I send this message? Understanding the impact of interruptions, social hierarchy and perceived task complexity on user performance and perceived workload, *Decis. Support. Syst.* 55 (2013) 135–145.
- [33] T. Bucher, A. Gericke, S. Sigg, Process-centric business intelligence, *Bus. Process. Manag. J.* 15 (2009) 408–429.
- [34] H. Chen, R.H.L. Chiang, V.C. Storey, Business intelligence and analytics: from big data to big impact, *MIS Q.* 36 (2012) 1165–1188.
- [35] H.J. Watson, Tutorial: business intelligence - past, present, and future, *Commun. Assoc. Inf. Syst.* 25 (2009) 487–510.
- [36] W.H. Inmon, *Building the Data Warehouse*, John Wiley & Sons, 1996.
- [37] A. Trigo, F. Belfo, R.P. Estébanez, Accounting information systems: the challenge of the real-time reporting, *Procedia Technol.* 16 (2014) 118–127.
- [38] T. Kawamoto, B. Mathers, Key success factors for a performance dashboard, *DM Rev.* 17 (2007) 20–21.
- [39] A. Miller, J. Cioffi, Measuring marketing effectiveness and value: the Unisys marketing dashboard, *J. Advert. Res.* 44 (2004) 237–243.
- [40] N.H. Rasmussen, M. Bansal, C.Y. Chen, *Business Dashboards: A Visual Catalog for Design and Deployment*, John Wiley & Sons, Hoboken, 2009.
- [41] Y. Wind, Marketing as an engine of business growth: a cross-functional perspective, *J. Bus. Res.* 58 (2005) 863–873.
- [42] D. Reibstein, D. Norton, Y. Joshi, P. Farris, Marketing Dashboards: A Decision Support System for Assessing Marketing Productivity, in: *INFORMS Mark. Sci. Conf. INFORMS*, 2005.
- [43] B.H. Clark, A.V. Abela, T. Ambler, Behind the wheel, *Mark. Manag.* 15 (2006) 18–23.
- [44] G.K. DeBusk, R.M. Brown, L.N. Killough, Components and relative weights in utilization of dashboard measurement systems like the balanced scorecard, *Br. Account. Rev.* 35 (2003) 215–231.
- [45] C. Heaps, S. Handel, Similarity and features of natural textures, *J. Exp. Psychol. Hum. Percept. Perform.* 25 (1999) 299–320.
- [46] F. Heylighen, The growth of structural and functional complexity during evolution, in: F. Heylighen, D. Aerts (Eds.), *Evol. Complex*, Dordrecht, 1997, pp. 1–18.
- [47] E.R. Tufté, *The Visual Display of Quantitative Information*, Graphics press, Cheshire, 2006.
- [48] T.S. Amer, S. Ravindran, The effect of visual illusions on the graphical display of information, *J. Inf. Syst.* 24 (2010) 23–42.
- [49] W. Cleveland, *Visualizing Data*, Hobart Press, New Jersey, 1993.
- [50] R. Benayoun, J. de Montgolfier, J. Tergny, O. Laritchev, Linear programming with multiple objective functions: step method (stem), *Math. Program.* 1 (1971) 366–375.
- [51] H.S. Vitense, J.A. Jacko, V.K. Emery, Multimodal feedback: an assessment of performance and mental workload, *Ergonomics*. 46 (2003) 68–87.
- [52] A. Olivia, Gist of the scene, in: L. Itti, G. Rees, J.K. Tsotsos (Eds.), *Encycl. Neurobiol. Atten.*, Elsevier Science, San Diego, 2005, pp. 251–256.
- [53] M.R. Endsley, Automation and situation awareness, in: R. Parasuraman, M. Mouloua (Eds.), *Autom. Hum. Perform. Theory Appl.*, Lawrence Erlbaum Associates, Inc, New Jersey, 1996, pp. 163–181.
- [54] M.R. Endsley, From here to autonomy: lessons learned from human-automation research, *Hum. Factors* 59 (2017) 5–27.
- [55] J.H. Roberts, J.M. Lattin, Development and testing of a model of consideration set composition, *J. Mark. Res.* 28 (1991) 429–440.
- [56] B. Gawronski, L. Creighton, Dual process theories, in: D.E. Carlston (Ed.), *Oxford Handb. Soc. Cogn.* Oxford University Press, New York, 2013, pp. 282–312.
- [57] C.E.J. Hartel, K. Smith, C. Prince, Defining Aircrew Coordination: Searching Mishaps for Meaning, in: *Proc. Int. Symp. Aviat. Psychol.*, Columbus, 1991.
- [58] T. Hill, P. Lewicki, *Statistics: Methods and Applications: A Comprehensive Reference for Science, Industry, and Data Mining*, StatSoft, Tulsa, 2006.
- [59] E. Katok, Using laboratory experiments to build better operations management models, *Technol. Inf. Oper. Manag.* 5 (2011) 1–86.
- [60] J. David Xu, I. Benbasat, R.T. Cenfetelli, The nature and consequences of trade-off

- transparency in the context of recommendation agents, *MIS Q.* 38 (2014) 379–406.
- [61] G.E. Bolton, A. Ockenfels, U.W. Thonemann, Managers and students as news-vendors, *Manag. Sci.* 58 (2012) 2225–2233.
- [62] R.J. Jacob, K. Karn, Eye tracking in human-computer interaction and usability research: Ready to deliver the promises, in: J. Hyona, R. Radach, H. Deubel (Eds.), *Mind's Eye Cogn. Appl. Asp. Eye Mov. Res.*, Elsevier Science, Oxford, 2003, pp. 573–605.
- [63] S.G. Hart, L.E. Staveland, Development of NASA-TLX (task load index): results of empirical and theoretical research, in: P.A. Hancock, N.B.T. A, P. Meshkati (Eds.), *Hum. Ment. Workload*, North Holland Press, Amsterdam, 1988, pp. 139–183.
- [64] N.B. Sarter, R.J. Mumaw, C.D. Wickens, Pilots' monitoring strategies and performance on automated flight decks: an empirical study combining behavioral and eye-tracking data, *Hum. Factors* 49 (2007) 347–357.
- [65] J. Patrick, N. James, A. Ahmed, P. Halliday, Observational assessment of situation awareness, team differences and training implications, *Ergonomics*. 49 (2006) 393–417.
- [66] D.D. Woods, E.S. Patterson, E.M. Roth, Can we ever escape from data overload? A cognitive systems diagnosis, *Cogn. Tech. Work* 4 (2002) 22–36.
- [67] K. Sharma, H.S. Alavi, P. Jermann, P. Dillenbourg, A Gaze-Based Learning Analytics Model: In-Video Visual Feedback to Improve learner's Attention in MOOCs, in: *Proc. ACM Press, New York, Int. Conf. Learn. Anal. Knowl.*, 2016, pp. 417–421.
- [68] E.L. Wiener, Cockpit automation, in: E.L. Wiener, D.C.B.T. HF, A. Nagel (Eds.), *Hum. Factors Aviat, Academic*, San Diego, 1988, pp. 433–461.

Mario Nadj is a post-doctoral researcher at the Karlsruhe Institute of Technology (KIT) in Germany. His research interests include the fields of Business Intelligence & Analytics (BIA), Machine Learning, Physiological Computing, as well as Human-Computer

Interaction. He has three years of industry experience in BIA and Business Process Management solutions from SAP SE. His research has been published in leading international journals and conferences, such as the *Journal of Cognitive Engineering and Decision Making*, *ACM Conference on Human Factors in Computing Systems*, *International Conference on Information Systems*, *European Conference on Information Systems*, and *International Conference on Information Systems and Neuroscience*.

Alexander Maedche is a professor at the Karlsruhe Institute of Technology (KIT) in Germany. His research work is positioned at the intersection of Information Systems (IS) and Human-Computer Interaction. Specifically, he investigates novel concepts of interactive intelligent systems for enterprises and the society. He publishes in leading IS and Computer Science journals such as the *Journal of the Association of Information Systems*, *Decision Support Systems*, *Business & Information Systems Engineering*, *Information & Software Technology*, *Computers & Human Behavior*, *International Journal of Human-Computer Studies*, *IEEE Intelligent Systems*, *Vldb Journal*, and *AI Magazine*.

Christian Schieder is a professor at the Technical University of Applied Sciences Amberg-Weiden in Germany. His research focuses on Business Intelligence & Analytics (BIA), Enterprise Decision Management and Business Rules Technologies. His work has appeared in several conferences and journals including the *HMD - Praxis der Wirtschaftsinformatik*, *Hawaii International Conference on System Sciences*, and *European Conference on Information Systems* among others. Models and techniques developed by him have been incorporated in several dynamic pricing systems. For several years he held the position of a Chief Digital Officer at a large multi-national machinery company where he was responsible for the digitalization strategy and the data-driven solution business.