

Association for Information Systems

**AIS Electronic Library (AISeL)**

---

ECIS 2022 Research Papers

ECIS 2022 Proceedings

---

6-18-2022

## How Transparency Measures Can Attenuate Initial Failures of Intelligent Decision Support Systems

Christian Matt

*University of Bern*, christian.matt@iwi.unibe.ch

Constantin Schlusche

*University of Bern*, constantin.schlusche@googlemail.com

Follow this and additional works at: [https://aisel.aisnet.org/ecis2022\\_rp](https://aisel.aisnet.org/ecis2022_rp)

---

### Recommended Citation

Matt, Christian and Schlusche, Constantin, "How Transparency Measures Can Attenuate Initial Failures of Intelligent Decision Support Systems" (2022). *ECIS 2022 Research Papers*. 10.

[https://aisel.aisnet.org/ecis2022\\_rp/10](https://aisel.aisnet.org/ecis2022_rp/10)

This material is brought to you by the ECIS 2022 Proceedings at AIS Electronic Library (AISeL). It has been accepted for inclusion in ECIS 2022 Research Papers by an authorized administrator of AIS Electronic Library (AISeL). For more information, please contact [elibrary@aisnet.org](mailto:elibrary@aisnet.org).

# HOW TRANSPARENCY MEASURES CAN ATTENUATE INITIAL FAILURES OF INTELLIGENT DECISION SUPPORT SYSTEMS

*Research Paper*

Christian Matt, University of Bern, Bern, Switzerland, christian.matt@iwi.unibe.ch

Constantin Schlusche, University of Bern, Bern, Switzerland, constantin.schlusche@googlemail.com

## Abstract

*Owing to high functional complexity, trust plays a critical role for the adoption of intelligent decision support systems (DSS). Especially failures in initial usage phases might endanger trust since users are yet to assess the system's capabilities over time. Since such initial failures are unavoidable, it is crucial to understand how providers can inform users about system capabilities to rebuild user trust. Using an online experiment, we evaluate the effects of recurring explanations and initial tutorials as transparency measures on trust. We find that recurring explanations are superior to initial tutorials in establishing trust in intelligent DSS. However, recurring explanations are only as effective as tutorials or the combination of both tutorials and recurring explanations in rebuilding trust after initial failures occurred. Our results provide empirical insights for the design of transparency mechanisms for intelligent DSS, especially those with high underlying algorithmic complexity or potentially high damage.*

*Keywords: intelligent decision support systems, artificial intelligence, transparency measures, initial failures, user trust.*

## 1 Introduction

The rapid progress of Artificial Intelligence (AI) over the past years has enabled new possibilities for intelligent decision support systems (DSS), allowing users to profit from systems that are able to adapt to their behavior over time or that can support them even in complex tasks (Benbya et al. 2021). This can range from voice-based assistants that provide shopping recommendations proactively to virtual health coaches that assist users in constantly increasing their training performance. Intelligent DSS are not fully autonomous, i.e. they assist users with decision-support that they can either accept or reject. With DSS becoming intelligent, we see two distinct but related issues arising: First, users may develop expectations that the systems can possibly not fulfill even in the foreseeable future (Alshurideh et al. 2020; Wang et al. 2019). Failing to meet the high user expectations can become particularly problematic during early usage phases, since an established basis of trust is yet to be built (Brown et al. 2008). Second, users face increasing difficulties in assessing the systems' capabilities, since they often perceive intelligent DSS as "black-boxes" they can hardly understand or control (Bauer et al. 2021).

Despite technological progress, system-inherent failures of intelligent DSS cannot be fully avoided, and especially for self-learning systems, their initial performance is often inferior owing to the time it takes to adapt their behavior to users' needs. Therefore, DSS providers need to know how they should inform users of the intelligent DSS' capabilities to avoid potential user discontinuance after an initial negative experience. DSS providers can employ "transparency measures", i.e., explicit actions to inform users about the capabilities and functioning of the intelligent DSS (Felzmann et al. 2019). Here, for a given level of transparency, the question is in which form transparency should be conveyed. Common transparency measures include tutorials that are provided prior to a usage experience ("initial tutorials")

and explanations that are provided by the system during the process (“recurring explanations”) (Confalonieri et al. 2021). Initial tutorials seek to explain the DSS’ functioning and capabilities prior to the first use, while recurring explanations provide information on the underlying rationale for providing a recommendation for each decision-support process step. As single respectively continuous informative interventions, both transparency measures might often be rather simple to implement. However, providers need to consider that too much information can lead to information overload or alert users of potential issues that they might never encounter (Kim et al. 2019). Therefore, choosing an appropriate level of transparency measures is critical for the successful design and adoption of intelligent DSS.

The previous literature has identified that feedback mechanisms can significantly impact trust in automated systems (Cramer et al. 2008; Gedikli et al. 2014; Wang et al. 2016). It has been found that the more transparent the process or the decision-making procedure is for users, the more likely they are to trust the system, albeit especially the early use phases are critical and can be decisive for further use (Seong and Bisantz 2008). Users exhibit a higher willingness-to-pay for intelligent systems that are more transparent (Peters et al. 2020). While hereby highlighting the importance of transparency in informing users about the capabilities of intelligent systems, the previous literature has not yet investigated and compared different transparency measures and their effects on attenuating failures in the crucial initial usage phase for intelligent DSS, i.e. they system provides a recommendation that turns out to be suboptimal or wrong in the first interaction between the user and the system. To close this gap, we pose the following research question:

*How can transparency measures increase trust in intelligent DSS after initial failures?*

Using a user-friendly online experiment, we investigated how trust in intelligent DSS evolves after the intelligent DSS has failed to provide accurate decision support to users during their initial interaction with the system. We manipulated whether an initial tutorial, recurring explanations or both were used to describe the intelligent DSS’ capabilities. Since research has already discovered that transparency has a positive influence on trust in IT systems (Herlocker et al. 2000; Hoff and Bashir 2015), it was not our goal to assess the general effect of transparency on trust. Instead, we investigated how transparency measures should be integrated in user interactions to rebuilt user trust after their first undesirable experiences with the system. Our work does not only have important implications for the field of decision support in general, it also stimulates future research on user experiences for other types of intelligent systems with potentially high technological complexity. Our systematic investigation is also highly relevant for providers of intelligent DSS who might consider the integration of multiple transparency measures in parallel.

## 2 Theoretical Background

### 2.1 Intelligent Decision Support Systems

While having a similar purpose as conventional DSS, intelligent DSS are additionally supplemented by AI technologies when providing support to users. It needs to be noted here that the foundations of AI and DSS were not always closely aligned with each other. While AI might seek to largely replace humans as decision-makers in the long run, DSS primarily focuses on supporting the decisions of decision-makers. However, due to advancements in technological capabilities, the combination of both paradigms can provide additional benefits to users. In the early days of intelligent DSS, they were often supported by rule-based systems, but neural networks and genetic algorithms are nowadays more frequently used (Belciug and Gorunescu 2020). Typical application fields for intelligent DSS include, among others, the (semi-)automatic classification and analysis of images as well as the so-called expert systems. The latter are intelligent, adaptive computer programs that can generate new knowledge in a specialized field with the help of reasoning mechanisms (Wagner 2017).

Intelligent DSS are designed to help users by continuously adapting their problem-solving process and behavior through mechanisms such as machine learning to better match users’ preferences and needs. Therefore, not just their diffusion in practice but also research activities in this domain have increased

in recent years. However, the continuous adaptation to user needs often implies that users lose some control over the system, and that the system could adapt its behavior without explicit authorization of the user (Holliday et al. 2016). Likewise, this may lead to varying outcomes despite constant system input, making the behavior of intelligent DSS even more difficult to predict. While IT background processes are in general often not transparent to users, it hereby becomes even more difficult to impossible for users of intelligent systems to understand their underlying functioning (Meske et al. 2022; Schmidt et al. 2020). As a result, when interacting with intelligent DSS, users need to build trust in the systems, despite an increased uncertainty about the systems' outcomes and the underlying processes.

## **2.2 Trust in Decision Support Systems**

Information Systems Research often considers trust as a belief, being represented based on the three dimensions competence, integrity, and benevolence (Xu et al. 2016). While originally used to describe trust in interpersonal relationships, the transfer to IT systems has been confirmed in various setting, also in the context of recommender system adoption (Komiak and Benbasat 2006; Wang and Benbasat 2016). In recent years, IS research has increasingly progressed towards evaluating the determinants of trustworthy system adoption (Wang et al. 2016). The focus on the relationship between humans and systems, combined with the increasing automation of information technology creates parallels to previous works in the fields of ergonomics and human factors, which also focused on trust in relation to increasing automation (Lee and See 2004). Thus, some researchers recommend that trust in IT artifacts should not be evaluated solely based on findings on interpersonal trust, suggesting to focus on human trust in the IT artifact itself (Lankton et al. 2015; McKnight et al. 2011; Söllner et al. 2012). Lankton et al. (2015) show that the degree of anthropomorphism of an IT artifact should be considered when choosing a conceptualization of trust. According to them, the more human a system appears to users, the more appropriate is a measurement of trust based on competence, integrity, and benevolence. Conversely, such an interpretation of trust appears less appropriate as the degree of automation increases and the IT artifact appears more inhuman. The response failures of conversational agents (CA) have a negative impact on users' perceptions of the CA in terms of perceived humanness, familiarity, and service satisfaction (Diederich et al. 2021). However, if robots provide the context of why they might make mistakes, there is an increase in trust in certain behaviors of the agent (Natarajan and Gombolay 2020).

Regardless of the perspective, three main groups of trust determinants have emerged in research on DSS (Wang and Benbasat 2008; Xu et al. 2016). In addition to person-related factors, which include all characteristics, attitudes, and perceptions of a user, situation-related factors represent all external influences that affect the trust relationship between users and DSS. However, the most considerable factors are often system characteristics, which mostly comprises factors that are directly related to the performance of a system (Hancock et al. 2011). Here, first and foremost is often the reliability of a DSS. The more consistent DSS output is perceived, the more likely it is that users build trust (de Visser and Parasuraman 2011). If, on the other hand, DSS act unreliably, such failures negatively impact users' trust. For instance, malfunctions such as total system failures can negatively affect trust (Merritt and Ilgen 2008), while erroneous results or false alarms can be reasons for a loss of trust (Dixon et al. 2007).

Researchers also found that feedback and transparency mechanisms have a significant effect on trust in automated systems (Cramer et al. 2008; Ososky et al. 2014). Through such mechanisms, failures can be justified and underlying processes be explained to users, allowing them to assess the system's reliability. As a consequence, such mechanisms can help prevent user resistance (Madhavan et al. 2006; Weiler et al. 2021). While various works recognize the importance of system-related failures for the further development of trust in IT systems (Manzey et al. 2012; Yang et al. 2016), the effects of failures conducted by intelligent DSS during the initial usage phase remains yet to be explored. During their initial interaction with the system, users need to be able to assess what the system does and how the intelligent DSS performs in terms of reliability. In addition to which information is provided here, we assume that the resulting level of user trust also depends on how and when this information is presented to users. With our research, we therefore focus on transparency of system-related factors in a specific

context, namely the occurrence of failures at a very early stage of usage, where first trust building processes take place.

### **2.3 Intelligent Systems and Transparency**

More transparent decision-making processes can establish user trust in intelligent systems (Seong and Bisantz 2008), but excess levels of transparency can also have contrary effects (Kizilcec 2016). Transparency can be particularly helpful for intelligent systems that are not fully reliable and that make mistakes at least occasionally. Openly revealing the reasons behind failures can help mitigate the loss of trust after their occurrence (Adams et al. 2003; Holliday et al. 2016; Madhavan et al. 2006). Thus, when designing DSS, the primary concern should not only be that the system output is accurate, but also whether those results – both correct and incorrect results – can be understood by users (Gönül et al. 2006). The transparent representation of system output and processes are therefore necessary requirements to achieve a better match between user perceptions and actual system properties (Cramer et al. 2008; Ososky et al. 2014).

Explanations have proven to be a suitable means for implementing more transparency into DSS (Papamichail and French 2003). They allow users to better understand how and why intelligent systems demonstrate a certain behavior or provide particular decisions (Haynes et al. 2009). Explanations can also increase confidence in the behavior of intelligent systems as well as in their outputs (McGuinness and Da Silva 2004; Wright et al. 2020). Crucial aspects to make explanations more effective are often the identification of the right form and content of the explanations (Haynes et al. 2009). Recent field of research focuses on how appropriate visualization techniques can be used to increase the transparency of complex algorithmic processes to establish trust (e.g. Kim et al. 2020). Evaluating different explanation structures and characteristics, Gönül et al. (2006) distinguished between short and long text-based explanations with high and low suggested confidence in the results. Highly confident statements and longer explanations turned out to be the most effective design of explanations in this framework with respect to the acceptance of DSS (Gönül et al. 2006). Kizilcec (2016) found that the level of transparency needs to be balanced, since an excess of information can increase doubt and distrust (Kizilcec 2016). Schmidt et al. (2020) manipulated the support functionality of intelligent systems and confirmed that too much transparency can have undesired effects for providers. In addition, providing users with too much information would also contradict the original aim of DSS to reduce cognitive burdens. On the other hand, oversimplification of the provided information may also lead to distrust (Kulesza et al. 2013).

While it is evident from the previous literature that at an appropriate level of transparency is suitable to establish trust in DSS, it is still not fully clear how different transparency measures should be structured in terms of content and timing. This challenge is even more essential after erroneous decisions have endangered trust in user-system relationships that are yet to be built.

## **3 Hypotheses**

It is well-known that system performance has a significant impact on users' trust in the system (Hoff and Bashir 2015). Likewise, we also know that IT systems that do not fulfill their desired purpose or if they do not achieve the required performance, this can reduce trust and endanger the adoption and continuous use of the systems (Hoehle et al. 2012). We can also expect that erroneous recommendations of an intelligent DSS have a negative impact on trust. However, the initial usage phase of an IT system is distinct since users' trust is often still largely based on personality traits or environmental or institutional factors (Zhou 2011). Given the missing experience so far, it requires further interactions with the intelligent DSS to enable a trust assessment based on system performance. If failures occur right at the beginning of a usage experience, all subsequent experiences are based on this initial disappointment, where users' initial expectations have possibly not been met. Drawing upon expectation-confirmation theory (Brown et al. 2008), expectations towards an intelligent DSS can be formulated as a claim that the system provides helpful and correct recommendations to solve a decision problem effectively and efficiently. According to expectation-confirmation theory, such critical

incidents can play a central role in further system use and the formation of trust in the system (Bhattacharjee and Premkumar 2004; Lankton et al. 2014). While users may or may not be aware that intelligent DSS are often more prone to errors in initial usage phases (owing to missing user information), it is still due to the failure in the initial usage phase, that even low user expectations are not met, failing to trigger satisfaction in response, and thus leading to decreasing trust. Therefore, we suggest:

*H1: Initial failures of intelligent DSS lead to lower user trust.*

While transparency measures represent one way to foster transparency (Gönül et al. 2006), the question of when these measures should be provided during the decision process emerges. Potential options range from before, during, or after the decision support is initiated. Instead of timing, previous works on the design of transparency measures have particularly focused on the extent of the provided content. Several researchers have found that extensive information has a positive effect on trust in IT systems in general, and in DSS in particular (Gönül et al. 2006; Kizilcec 2016). However, as the amount of the provided information increases, users' cognitive burden also increases with potentially unwanted consequences such as information overload or even user resistance (Aljukhadar et al. 2012). One option to prevent this is to present the relevant information in reduced form at different points in time, thereby dividing a potentially large bulk of information into smaller slices, which are presented at individual steps of the decision-support process. Splitting information can help ensure that users still receive all necessary information but without putting their information processing capacity to the test (Aljukhadar et al. 2012).

The previous findings seem applicable to the general characteristics of tutorials and recurring explanations. In essence, both transparency measures seek to reveal more information about the processes and operations of a system to make them more comprehensible to the user, but they differ in the way they seek to achieve this (Seong and Bisantz 2008). Tutorials provide users with information on general DSS functionality and capabilities prior to the first use, and in some cases, they might also be accessed again later. However, such tutorials are usually formulated in a more general way as they need to be applicable to various scenarios. In other words, by providing detailed explanations prior to using a DSS, users can better assess and validate later DSS suggestions, but the information from the tutorial is not directly related to one of the later specific suggestions offered by the DSS. Therefore, users need to incorporate and adopt the rather general information from the tutorial into their later decisions. In contrast, recurring explanations can offer information directly applicable to individual process steps. Users can employ such recurring explanations directly to verify whether an intelligent DSS has correctly identified the right decision problem. Therefore, unless they are overloaded with unnecessary information, recurring explanations can often help improve users' understanding of the system (Chazette and Schneider 2020). Based on the previous findings for both tutorials and recurring explanations, we suggest the following hypothesis:

*H2: Recurring explanations lead to higher user trust compared to initial tutorials when there are no initial failures of intelligent DSS.*

If failures by the intelligent DSS occur, users' awareness of future potential failures might increase in further interactions with the systems, hereby raising the need for ways to validate the system's recommendations. As users' expectations were not met, users tend to evaluate the available information more diligently to find explanations for possible discrepancies between their expectations and the suggested results (Kizilcec 2016). Overall, the initial failure is therefore likely to place more weight on users' consideration of the provided transparency measures. Early research has already shown that dealing with transparency measures implies a cost/benefit trade-off for users (Gregor and Benbasat 1999). The major costs relate to the processing of the provided transparency information, leading to additional cognitive burden on the decision maker. This is also one reason why providers of "faultless" systems often employ only relatively little to no transparency information (Dhaliwal 1993). In addition to this, especially an excess of feedback can lead to lower decision-making performance or make users aware of hypothetical problems (Lam et al. 2011; Lurie and Swaminathan 2009). While providing recurring explanations can therefore be also problematic in terms of raising high awareness, initial tutorials in contrast do not provide any explicit opportunities for users to validate the system's processes

and outcome after the initial failure has occurred. Here, participants can only resort to the rather general information from the tutorial to help them better classify the previous incident as well as the plausibility of the current DSS recommendation. Therefore, initial tutorials do not seem to cater for the rather high need for transparency after an initial failure occurred. Therefore, we hold:

*H3: Recurring explanations lead to higher user trust compared to initial tutorials in case of an initial failure of intelligent DSS.*

## 4 Methods

### 4.1 Experimental Task and Data Collection

For the online experiment, we put substantial efforts in providing participants with an intuitive and user-friendly DSS. The system should exhibit comprehensible, intelligent properties for users by integrating a self-learning component. The intelligent DSS was implemented in the form of an image recognition tool, that supported participants with their task of identifying the number of geometric shapes (triangles) in the overall image they were provided with. One advantage of this implementation is that complex decision situations can be represented by a graphical representation without having to generate more narrative, fictitious scenarios. We opted for such a rather basic implementation also because we did not want to trigger participants into questioning the genuineness of the experiment's purpose and its background story, since this could have raised additional trust concerns.

For supporting participants in identifying the correct number of triangles, a fictitious anthropomorphic intelligent DSS ("Robby") assisted them in evaluating the number of figures by providing participants with a suggested solution. We decided to implement Robby with anthropomorphic properties, since the perception of anthropomorphism can increase the enjoyment of using personal agents (Moussawi et al. 2021). Subjects were informed that the purpose of the study was to test a new image recognition tool. They were asked to provide feedback on the intelligent DSS "Robby" to further improve the system before it would be used in practice. Participants received information about Robby's main functionalities and purpose, as well as information on its average recommendation accuracy (>99%). Participants were first told that the tool was a self-learning algorithm, but they were deceived after the experiment, since the tool only provided predefined recommendations to them. Before starting with their first task, participants received an introduction to the experimental user interface to better familiarize with the test environment. A highly simplified example was used to show what kind of decision problems they could expect and how Robby would assist them.

We conducted a pretest to ensure the tasks had an appropriate level of associated user uncertainty concerning the accurateness of the solutions proposed by the tool. We also wanted to make sure that there were substantial efforts involved for participants to complete the task manually, making them at least consider employing the DSS. Participants' goal was to solve all five decision tasks without failure, and for each correct answer participants received additional tickets for the drawing of a cash prize. We raffled three times 100€, whereas participants could increase their chances of winning by a maximum factor of 5, depending on the number of correct answers. All subjects needed to complete the five decision tasks in identical order. The difficulty of the tasks was kept constant. During each task, the experimental interface pretended that Robby in parallel analyzed the images and generated a specific recommendation for the overall identified number of images. Participants could either adopt or reject the recommendation and enter their own decisions. After each decision problem, subjects were provided with direct feedback on whether their answer was correct. In a survey, we also asked subjects to reflect on their experiences with and perceptions of Robby.

We recruited a total of 520 participants (mostly students) who completed the experiment via several university-related social media channels. The responses of 20 participants turned out to be incomplete or inconsistent, therefore these were excluded from the analysis, resulting in a final sample size of 500 participants, divided into six different treatment groups (Table 1). The average age of the participants was 24.37 years; 65.6% were female (Table 2). We conducted an ANOVA to compare the distribution

of age across the different treatment groups, and we did not find any evidence for unequal distributions ( $F_{2,494} = 0.361$ ,  $p = 0.697$ ). For gender, we used the Pearson chi-squared test for which we did neither obtain any evidence for unequal distributions across groups ( $\chi^2_{2, N=500} = 0.155$ ,  $p = 0.926$ ). Therefore, we assumed a successful random assignment of participants to experimental groups based on gender and age on the different treatment groups.

		Performance	
		Initial Failure	No Failure
Transpar. Measure	Initial Tutorial	78	85
	Recurring Explanations	82	85
	Both	85	85

Table 1. Number of participants per group.

Demographics	% of sample
Male	34.4
Female	65.6
16 – 24	58.2
25 – 34	39.0
35 – 44	1.8
45 – 54	0.6
55 and over	0.4

Table 2. Sample demographics.

## 4.2 Manipulations

We sought to compare the influence of initial failures as well as of different transparency measures. For this purpose, participants were randomly assigned to six groups, that differed based on whether (1) there was an initial failure or not, and (2) the system employed an initial tutorial, recurring explanations, or both (Figure 1). To reduce the complexity of the experiment but also since our focus was on comparing different forms of interventions as well as the combination thereof, we did not use a separate control group. Concerning the implementation of the performance variable, only marginal deviations of the tool's recommendations from the correct solution were chosen to avoid that users could identify system failures right away. In order to analyze the effect of the transparency measures, we designed both the tutorial and the recurring explanations following the typology of Gregor and Benbasat (1999). Regarding a potential failure assessment for users, both transparency measures differed in the level of information specificity that they provided to participants. While tutorials only provided general information on the tool's average accurateness ( $> 99\%$ ), recurring explanations provided an individual confidence level for each decision task (again with values  $> 99\%$ ). To test whether the participants noticed their individual treatments, we included manipulation checks for both system performance as well as the transparency measures. We conducted ANOVAs/t-tests to assess the successful manipulations. Both for performance as well as for transparency measures, all tests indicated significant differences ( $p < 0.01$ ) between groups in the perceptions of the different treatments. Therefore, we assume that the manipulations were effectively implemented and perceived as such.

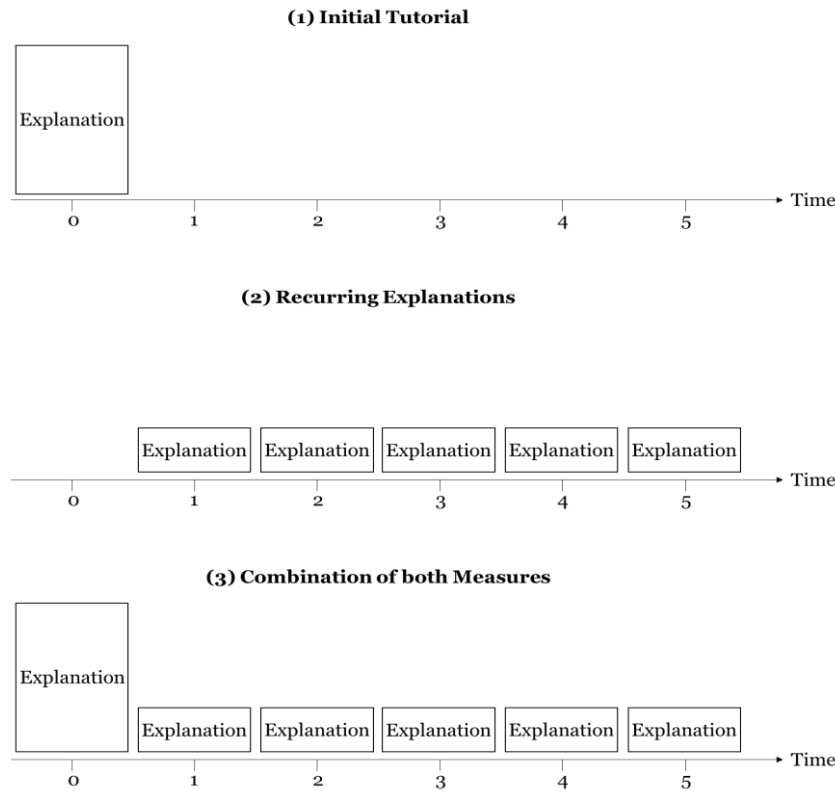


Figure 1. Experimental transparency treatments.

### 4.3 Operationalization and Validation

For the implementation of trust, we followed the well-established dimensions of *capability*, *integrity*, and *benevolence* from Mayer et al. (1995), because all of them seem highly relevant for intelligent DSS. To operationalize these three dimensions, we adopted the scales from Lankton et al. (2015) and captured each of them by three items each. Examples of the adapted questions included "Robby is competent and effective in evaluating images" (ability), "Robby is honest" (integrity), and "Robby acts in my best interest" (benevolence). Previous research found that person-related processes play a significant role in the development of trust, indicating that people have different propensities to trust other people or things (Brown et al. 2004). Furthermore, the more human-like a robot looks, the more human-like characteristics are attributed to the robot (Hegel et al. 2008). To operationalize this propensity, we adopted the scale from Cheung and Lee (2001). Accounting for the well-established relations between the main Technology Acceptance Model (TAM) variables and trust in IT systems (e.g., Agag and El-Masry 2016; Gefen et al. 2003), we adopted the scales for ease of use and perceived usefulness from Koufaris and Hampton-Sosa (2004). We recorded all of these items using a 7-point Likert-type scale, ranging from (1) strongly disagree to (7) strongly agree.

Testing for construct reliability, Cronbach's alpha for all scales were 0.85 or above, indicating that the measures were reliable (Table 3). Since all square roots of the average variance extracted (on the diagonal in Table 3) were larger than the corresponding correlations of the construct with every other construct, requirements for discriminant validity were fulfilled (Fornell and Larcker 1981).

	Variable	Reliability	Min	Max	Mean	Std. dev	1	2	3	4	5	6
1.	Trust	0.977	1.73	7.00	5.26	0.97	0.709					
2.	Propensity to Trust	0.856	1.00	7.00	4.00	1.30	0.004	0.780				
3.	Perceived Usefulness	0.913	1.00	7.00	5.52	1.33	0.573**	0.041	0.856			
4.	Perceived Ease of Use	0.876	1.00	7.00	5.88	1.04	0.519**	-0.020	0.636**	0.807		
5.	Age	n/a	16	59	24.38	4.82	-0.102*	0.060	-0.059	-0.050	n/a	
6.	Gender	n/a	0	1	0.34	0.48	-0.053	0.037	0.049	-0.054	0.091*	n/a
*: p<0.05; **: p<0.01												

Table 3. Descriptive statistics.

## 5 Results

We first used a two-factor ANCOVA to test the influence of system performance (initial failure vs. no initial failure) and transparency measures on users' trust in the DSS. An ANCOVA reduces the unsystematic variability in the design by controlling for individual differences and particular co-variables, thereby providing greater power to detect effects (Grabe and Westley 2003). To test the applicability of the ANCOVA, we conducted a Levene test of error variance. It indicated different error variances for the different groups ( $F_{5, 494} = 2.362$ ;  $p = 0.039$ ). However, such differences are tolerable when all treatment groups are of similar size (Parra-Frutos 2013). In addition, the normality assumption was satisfied for every treatment group, therefore meeting the common assumptions for two-factor ANCOVAs (Davison and Sharma 1994).

Concerning the co-variables, the ANCOVA ( $R^2 = 0.392$ ) revealed no significant impact of individuals' propensity to trust ( $F_{1, 489} = 0.057$ ;  $p = 0.812$ ), age ( $F_{1, 489} = 3.016$ ;  $p = 0.083$ ), and gender ( $F_{1, 489} = 2.253$ ;  $p = 0.134$ ). In contrast, it revealed a significant impact of both perceived usefulness ( $F_{1, 489} = 71.151$ ;  $p < 0.001$ ), and perceived ease of use ( $F_{1, 489} = 28.680$ ;  $p < 0.001$ ) on trust. For the first treatment factor, the data supported the effect of system performance on trust ( $F_{1, 489} = 6.164$ ;  $p = 0.013$ ). The post-hoc test for system performance also confirmed that users have lower trust into the intelligent DSS after an initial failure (Table 4), hence supporting H1.

(I) Manipulation	(J) Manipulation	Mean Difference (I-J)	Std. Error	P-value <sup>b</sup>
Initial Failure	No Failure	-0.168*	0.069	0.015
Tutorial	Explanations	-0.203*	0.084	0.050
Tutorial	Both Measures	-0.116	0.084	0.503
Explanations	Both Measures	0.087	0.083	0.892
Based on estimated marginal means, dependent variable trust				
*: p<0.05; **: p<0.01      b: Adjustments for multiple comparisons: Bonferroni.				

Table 4. Aggregated post-hoc comparisons.

Next, we focused on the second treatment factor: transparency measures. The ANCOVA did not reveal a significant effect of the different transparency measures ( $F_{2, 489} = 2.785$ ;  $p = 0.063$ ). Due to the small margin, we also conducted post-hoc tests for the different measures (Table 3). Based on their estimated marginal means, the difference between the initial tutorial ( $M = 5.150$ ,  $SE = 0.060$ ) and the recurring explanations ( $M = 5.353$ ,  $SE = 0.059$ ) proved to be significant, indicating that recurring explanations can be more effective than tutorials in increasing user trust. Meanwhile, the differences between the

other two types of measures were not significant. We took the opportunity to conduct group-level post-hoc comparisons to analyze whether the significant differences between tutorials and explanations on the aggregate level were also present on an individual treatment level (Table 5). The post-hoc tests for the interaction between system performance and the different transparency measures showed that recurring explanations only lead to a significantly higher level of trust than initial tutorials if there was no initial error. Again, these results confirm that the combination of both measures is not more effective than any of the two measures individually. Figure 2 shows a visual representation of the estimated marginal means. Hence, we found support for H2, but did not find support for H3.

Performance	(I) Measure	(J) Measure	Mean Difference (I-J)	Std. Error	P-value
Initial Failure	Initial Tutorial	Rec. Explanations	-0.123	0.121	0.931
	Initial Tutorial	Both Measures	-0.133	0.120	0.807
	Rec. Explanations	Both Measures	-0.010	0.119	1.000
No Failure	Initial Tutorial	Rec. Explanations	-0.283*	0.117	0.049
	Initial Tutorial	Both Measures	-0.099	0.117	1.000
	Rec. Explanations	Both Measures	0.184	0.117	0.353
Based on estimated marginal means, dependent variable trust					
*: $p < 0.05$ ; **: $p < 0.01$ b: Adjustments for multiple comparisons: Bonferroni.					

Table 5. Group-level post-hoc comparisons.

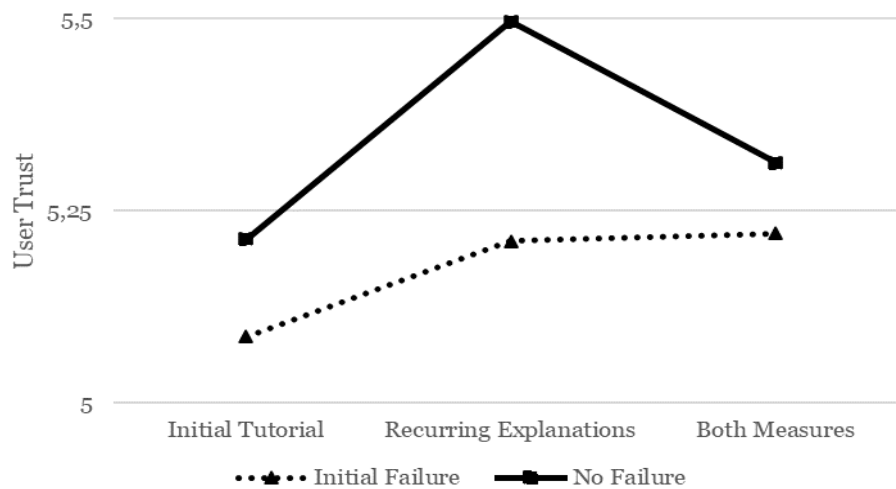


Figure 2. Estimated marginal means.

## 6 Discussion

Again, we start with the co-variables before we discuss the treatment factors and their effects on the dependent variable trust. User-related factors were explicitly included as covariates in our model since the previous literature has highlighted their importance for user trust in the context of DSS (Ezer et al. 2008; Lee 2008; Pak et al. 2012). Contrary to our expectations, these effects turned out to be not significant. Neither trust propensity nor age or gender were identified as significant influences. At least for age, this might be partly explained by the variance in age of our comparably young sample. Even though the average value for the propensity to trust was not particularly high in our research, we see the possibility of an overall higher trust in intelligent systems by digital natives as a possible explanation. In contrast to the aforementioned co-variables, the two factors perceived ease of use and perceived

usefulness were confirmed as relevant factors for intelligent DSS as well. These results are in line with the results of Gefen et al. (2003) and Agag and El-Masry (2016), who analyzed the relation between trust, and technology acceptance factors for other types of DSS.

Incorrect suggestions have a significant negative effect on trust if they occur when the intelligent DSS is used for the first time. This confirms the meta-analysis by Hancock et al. (2011), in which they identified system-related factors as key factors of trust in automated systems. Furthermore, we discovered that for trust as our dependent variable, recurring explanations have a significant advantage over initial tutorials – but only if no initial failure occurs. It is remarkable that the combination of both transparency measures (tutorials + recurring explanations) has no significant advantage over each of the two transparency measures employed individually. One possible explanation for this is that already the detailed presentation of the tutorial in the beginning leads to cognitive overload and therefore restricts participants in assessing further information (Aljukhadar et al. 2012). Although the manipulation checks confirmed that the initial tutorial and also the combination of tutorial and explanations were successfully noticed, we were not able to capture how the information provided was exactly processed by participants. Cognitive overload could also imply that further information would only be processed to a limited extent. It seems a good opportunity for future research to dig deeper into this gap, by assessing users' processing of the different transparency mechanisms using additional tools such as MRI, or eye trackers, or different theoretical lenses such as dual system theory.

Previous research indicated that detailed explanations, e.g., in the form of natural language, positively influence trust in a DSS and its proposed recommendations (Gönül et al. 2006). While additional information seems generally useful as a means of transparency, our results show that in the context of intelligent systems, the type and the timing of presentation matters. Although detailed information about the processes and functioning of the DSS was provided with the help of an initial tutorial, no changes in trust in the system can be inferred from our results. Not only do initial tutorials tend to indicate lower trust levels than recurring explanations, but in combination with recurring explanations they are not better than recurring explanations alone. Our results therefore confirm those of Kizilcec (2016), stating that too much and detailed information can lead to distrust in an intelligent system and doubts towards the system can arise, for the case of intelligent DSS. Another reason, why initial tutorials promote trust to a lesser extent than recurring explanations could be that users do not have any experience with the actual functioning of the intelligent DSS at the time when they receive the initial tutorial. Since the tutorial was displayed before the first interaction with the intelligent DSS, the tutorial's information and explanations about processes and procedures need to be transferred from users to a yet unknown abstract system. In contrast, providing explanations when users directly need the information to assess the overall decision situation, requires less memory capacity as well as a lower ability to abstract.

Recurring explanations, on the other hand, seem to have a positive effect on trust. A possible reason for this advantage over initial tutorials is the higher number of validation possibilities for users to assess the underlying processes and the intelligent DSS' output. Additionally, lower cognitive load on users compared to tutorials, which results from dividing the large extent of information from the tutorials into smaller pieces, could further help users assess the decision situation. We also need to note that transparency measures might be more effective if reasons for failures are made transparent (Adams et al. 2003). For our work, however, explanations for failures were omitted to ensure better comparability of initial tutorials and recurring explanations, as well as higher comparability between the different system performance groups. The present study thus also complements previous research, which investigated how aversion to algorithms can be reduced using modifications (Dietvorst et al. 2018). We show that recurring explanations also seem to have a positive effect on trust, and these seem to reduce aversion to DSS. Table 6 summarizes the findings for the different hypotheses.

Hypothesis		Support
H1	Initial failures of intelligent DSS lead to a decrease in user trust.	Yes
H2	Recurring explanations lead to higher user trust compared to initial tutorials when there are no initial failures of intelligent DSS.	Yes
H3	Recurring explanations lead to higher user trust compared to initial tutorials in case of an initial failure of intelligent DSS.	No

Table 6. Overview of hypotheses.

## 7 Theoretical and Practical Implications

Our results suggest that failures in the initial use of an intelligent DSS have an impact on the further development of trust in the system, and that recurring explanations can more successfully establish trust in the initial phase if the relation between user and system has been free from failures so far. However, they are only as effective as an initial tutorial or the combination of both tutorial and recurring explanations in rebuilding trust after an initial failure has happened. These findings contribute to the rapidly growing body of research on transparency and explainability of intelligent systems (Meske et al. 2022) as well as the research on fairness of AI (Feuerriegel et al. 2020; Weith and Matt 2022). Our results both confirm and extend the results of several previous studies. While Schmidt et al. (2020) directly manipulated the support functionality of intelligent systems, we differentiated the timing and extent of the provision of information about the systems' capabilities. We showed that without changing any system properties, just by configuring how and when information about the intelligent DSS is presented to users, can affect system-user trust relations.

Our work can also stimulate other research on the visualization of neural networks and other complex AI algorithms to users (e.g., Kim et al. 2020), since we indicate that the timing of provider interactions needs to be balanced with users' cognitive capabilities. Future research should analyze whether different forms of visualizations and provisions times might be more suitable for expert or inexperienced users. We confirm that the extent of the transparency measures should be calibrated carefully, since maximal transparency (in our experiment tutorials + recurring explanations) may have no positive or even negative effects, such as additional mistrust (Meske and Bunde 2020). In line with this, we complement previous research that has shown that a recurring information stream increases trust in automated systems (Sanders et al. 2014), or that explanations can increase trust due to higher understandability and predictability of the system (Meske and Bunde 2020). Our findings should also be considered in the context of user expectations and requirements as well as the evaluation and processing of such expectations and requirements.

Our research has direct practical implications, acknowledging that accurate design of transparency mechanisms can be a substantial pillar of firms' strategies to reduce user resistance (blinded for review 2021). Previous findings show that higher transparency can lead to increases in trust in IT systems (Wang et al. 2016). However, especially in critical decision-making situations, additional efforts to establish continuous trust in intelligent DSS may be needed to ensure their continuous use (Hoehle et al. 2012). We have shown that recurring explanations can at least in the short-run build trust to a greater degree than initial tutorials. However, this is only the case if there has been no initial failure, otherwise both strategies have a similar performance. What is at least as important to practitioners is that in all scenarios with and without initial errors, providing both transparency measures did not lead to any additional positive effect. In contrast, the previous literature has already identified that an excess of feedback can negatively impact the quality of decision-making processes (Lam et al. 2011; Lurie and Swaminathan 2009). Accordingly, "the more, the better" is not a suitable target-oriented approach in the context of trust-building transparency measures Kizilcec (2016).

The choice of accurate transparency design should not consider potential effects on trust, but also any implementations costs. Providing users with additional information on how processes and outputs are compiled may not seem to require large economic or technological efforts. However, for particular DSS types and AI algorithms (e.g. deep learning), providing higher transparency remains challenging (Meske

et al. 2022). Therefore, if the implementation of recurring explanations is more costly, opting for an initial tutorial with only rather general information may seem worthwhile. However, the integration of both transparency measures at the same time does not seem advisable. In line with previous research (e.g. Wang et al. 2019), we recommend considering the explanation design in its entirety besides focusing on transparency elements alone.

## **8 Conclusion and Limitations**

Accounting for users' difficulties in assessing the capabilities of intelligent DSS – which are often more error-prone in initial usage phases, where trust has yet to be build – we analyzed which transparency measure can attenuate initial DSS failures. We showed that initial failures decrease user trust. Recurring explanations are only superior to initial tutorials when no initial failure had occurred. Otherwise, both measures are almost equally effective. Users appear to have a certain level of trust in intelligent DSS, which is not significantly impacted by personal characteristics. In contrast, perceived usefulness and ease of use turned out to be relevant predictors in addition to whether initial failures occurred. Instead of considering individual needs, developers of intelligent DSS should therefore primarily focus on essential system characteristics and how they can be communicated to users (Hancock et al. 2011).

Our study is not free from limitations. First, when interpreting the results, it should be considered that a vast majority of younger people between the ages of 16 and 29 participated in the experiment. As digital natives, they are subject to differences in the formation of trust in IT artifacts compared to older generations (Hoffmann et al. 2014; Matt et al. 2019). Hence, we cannot assume that all effects are stable also for other age groups. For instance, it might be plausible that older age groups require more comprehensive information or that they could also react differently to information overload if the information exceeds a certain threshold. It might also be plausible that “digital natives” are more accustomed to more interactive forms of explanations instead of a large tutorial at the beginning of their usage. Therefore, we recommend future research to replicate our study with a representative sample. We also acknowledge that our sample comprised more female than male participants, which may have impacted the results, since previous research has found that gender can have an impact on how suggestions of decision support systems are considered (Djamasbi and Loiacono 2008). However, other research did not find a substantial impact of gender on the adoption of DSS (Shibl et al. 2013), and our data did not show substantial differences between genders either.

Second, there is a wide range of possibilities of how trust can be measured, and we cannot rule out that other theoretical trust lenses would have led to the same results. However, while more technocratic, system-dependent implementations of trust scales are plausible, our experimental implementation was built upon an anthropomorphic personal assistant, who was carefully introduced to participants. Therefore, we explicitly decided to draw from a more interpersonal implementation of trust. We recommend that future research considers personal factors (such as previous trust incidents with other intelligent systems) in a more differentiated way. Here, it would be of interested how the different transparency measures affect a decrease of trust in comparison to a full loss of trust.

Last, owing to the experimental procedure, we were certainly not able to cover the entire wide variety of intelligent DSS, their potential functional scope, the extent of potential failures and their consequences, as well as other factors. For instance, there might be explicit differences between the overall volume of information between initial tutorials and permanent explanations that could explain the superiority of any of the two measures rather than timing or repetition advantage. However, we provided participants with a realistic and easy to use implementation, which was also indicated by the overall high trust scores, even for the treatments with an initial failure. This strengthens our assumption that we provided participants with a suitable decision-making environment. We nevertheless recommend that future research verifies our results in other intelligent DSS contexts and based on field experiments. For this we recommend analyzing the effects of transparency mechanisms over a longer time, and also taking into account that other transparency mechanisms (e.g., external reviews) could be used to establish a high level of trust prior to a first usage experience that could serve as a stronger barrier against the drop in user trust after the initial failure occurs.

## References

- Adams, B. D., Bruyn, L. E., Houde, S., and Iwasa-Madge, K. 2003. "Trust in Automated Systems Literatur Review," Department of National Defence Canada, DRDC Toronto CR-2003-096, <https://cradpdf.drdc-rddc.gc.ca/PDFS/unc17/p520342.pdf>.
- Agag, G., and El-Masry, A. A. 2016. "Understanding Consumer Intention to Participate in Online Travel Community and Effects on Consumer Intention to Purchase Travel Online and Wom: An Integration of Innovation Diffusion Theory and TAM with Trust," *Computers in Human Behavior* (60), 97-111.
- Aljukhadar, M., Senecal, S., and Daoust, C.-E. 2012. "Using Recommendation Agents to Cope with Information Overload," *International Journal of Electronic Commerce* (17:2), 41-70.
- Alshurideh, M., Al Kurdi, B., and Salloum, S. A. 2020. Examining the Main Mobile Learning System Drivers' Effects: A Mix Empirical Examination of Both the Expectation-Confirmation Model (ECM) and the Technology Acceptance Model (TAM), *In: International Conference on Advanced Intelligent Systems and Informatics*, Cairo, Egypt.
- Bauer, K., Hinz, O., van der Aalst, W., and Weinhardt, C. 2021. "Expl(Ai)N It to Me – Explainable Ai and Information Systems Research," *Business & Information Systems Engineering* (63:2), 79-82.
- Belciug, S., and Gorunescu, F. 2020. "A Brief History of Intelligent Decision Support Systems," in *Intelligent Decision Support Systems—a Journey to Smarter Healthcare*. Cham: Springer International Publishing, 57-70.
- Benbya, H., Pachidi, S., and Jarvenpaa, S. 2021. "Special Issue Editorial: Artificial Intelligence in Organizations: Implications for Information Systems Research," *Journal of the Association for Information Systems* (22:2), 281-303.
- Bhattacharjee, A., and Premkumar, G. 2004. "Understanding Changes in Belief and Attitude toward Information Technology Usage: A Theoretical Model and Longitudinal Test," *MIS Quarterly* (28:2), 229-254.
- Brown, H. G., Poole, M. S., and Rodgers, T. L. 2004. "Interpersonal Traits, Complementarity, and Trust in Virtual Collaboration," *Journal of Management Information Systems* (20:4), 115-138.
- Brown, S. A., Venkatesh, V., Kuruzovich, J., and Massey, A. P. 2008. "Expectation Confirmation: An Examination of Three Competing Models," *Organizational Behavior and Human Decision Processes* (105:1), 52-66.
- Chazette, L., and Schneider, K. 2020. "Explainability as a Non-Functional Requirement: Challenges and Recommendations," *Requirements Engineering* (25:4), 493-514.
- Cheung, C. M., and Lee, M. K. 2001. "Trust in Internet Shopping: Instrument Development and Validation through Classical and Modern Approaches," *Journal of Global Information Management* (9:3), 22-35.
- Confalonieri, R., Coba, L., Wagner, B., and Besold, T. R. 2021. "A Historical Perspective of Explainable Artificial Intelligence," *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery* (11:1), e1391.
- Cramer, H., Evers, V., Ramlal, S., van Someren, M., Rutledge, L., Stash, N., Aroyo, L., and Wielinga, B. 2008. "The Effects of Transparency on Trust in and Acceptance of a Content-Based Art Recommender," *User Modeling and User-Adapted Interaction* (18:5), 455-496.
- Davison, M. L., and Sharma, A. R. 1994. "Anova and Ancova of Pre- and Post-Test, Ordinal Data," *Psychometrika* (59:4), 593-600.
- Dhaliwal, J. S. 1993. "An Experimental Investigation of the Use of Explanations Provided by Knowledge-Based Systems." Dissertation, University of British Columbia.
- Diederich, S., Lembcke, T.-B., Brendel, A. B., and Kolbe, L. M. 2021. "Understanding the Impact That Response Failure Has on How Users Perceive Anthropomorphic Conversational Service Agents: Insights from an Online Experiment," *AIS Transactions on Human-Computer Interaction* (13:1), 82-103.

- Dietvorst, B. J., Simmons, J. P., and Massey, C. 2018. "Overcoming Algorithm Aversion: People Will Use Imperfect Algorithms If They Can (Even Slightly) Modify Them," *Management Science* (64:3), 1155-1170.
- Dixon, S. R., Wickens, C. D., and McCarley, J. S. 2007. "On the Independence of Compliance and Reliance: Are Automation False Alarms Worse Than Misses?," *Human Factors* (49:4), 564-572.
- Djamasbi, S., and Loiacono, E. T. 2008. "Do Men and Women Use Feedback Provided by Their Decision Support Systems (Dss) Differently?," *Decision Support Systems* (44:4), 854-869.
- Ezer, N., Fisk, A. D., and Rogers, W. A. 2008. "Age-Related Differences in Reliance Behavior Attributable to Costs within a Human-Decision Aid System," *Human Factors* (50:6), 853-863.
- Felzmann, H., Villaronga, E. F., Lutz, C., and Tamò-Larrieux, A. 2019. "Transparency You Can Trust: Transparency Requirements for Artificial Intelligence between Legal Norms and Contextual Concerns," *Big Data & Society* (6:1), 2053951719860542.
- Feuerriegel, S., Dolata, M., and Schwabe, G. 2020. "Fair Ai: Challenges and Opportunities," *Business & Information Systems Engineering* (62:4), 379-384.
- Fornell, C., and Larcker, D. F. 1981. "Evaluating Structural Equation Models with Unobservable Variables and Measurement Error," *Journal of Marketing Research* (18:1), 39-50.
- Gedikli, F., Jannach, D., and Ge, M. 2014. "How Should I Explain? A Comparison of Different Explanation Types for Recommender Systems," *International Journal of Human-Computer Studies* (72:4), 367-382.
- Gefen, D., Karahanna, E., and Straub, D. 2003. "Trust and TAM in Online Shopping: An Integrated Model," *MIS Quarterly* (27:1), 51-90.
- Gönül, M. S., Önköl, D., and Lawrence, M. 2006. "The Effects of Structural Characteristics of Explanations on Use of a Dss," *Decision Support Systems* (42:3), 1481-1493.
- Grabe, M. E., and Westley, B. H. 2003. "The Controlled Experiment," in *Mass Communication Research and Theory*, G. Stempel, D.H. Weaver and G.C. Wilhoit (eds.). Boston: Allyn&Bacon, 267-298.
- Gregor, S., and Benbasat, I. 1999. "Explanations from Intelligent Systems: Theoretical Foundations and Implications for Practice," *MIS Quarterly* (23:4), 497-530.
- Hancock, P. A., Billings, D. R., Schaefer, K. E., Chen, J. Y. C., de Visser, E. J., and Parasuraman, R. 2011. "A Meta-Analysis of Factors Affecting Trust in Human-Robot Interaction," *Human Factors* (53:5), 517-527.
- Haynes, S. R., Cohen, M. A., and Ritter, F. E. 2009. "Designs for Explaining Intelligent Agents," *International Journal of Human-Computer Studies* (67:1), 90-110.
- Hegel, F., Krach, S., Kircher, T., Wrede, B., and Sagerer, G. 2008. Understanding Social Robots: A User Study on Anthropomorphism, In: *RO-MAN 2008 - The 17th IEEE International Symposium on Robot and Human Interactive Communication*.
- Hoehle, H., Huff, S., and Goode, S. 2012. "The Role of Continuous Trust in Information Systems Continuance," *Journal of Computer Information Systems* (52:4), 1-9.
- Hoff, K. A., and Bashir, M. 2015. "Trust in Automation: Integrating Empirical Evidence on Factors That Influence Trust," *Human Factors* (57:3), 407-434.
- Hoffmann, C. P., Lutz, C., and Meckel, M. 2014. "Digital Natives or Digital Immigrants? The Impact of User Characteristics on Online Trust," *Journal of Management Information Systems* (31:3), 138-171.
- Holliday, D., Wilson, S., and Stumpf, S. 2016. User Trust in Intelligent Systems: A Journey over Time, In: *Proceedings of the 21st International Conference on Intelligent User Interfaces*.
- Kim, B., Park, J., and Suh, J. 2020. "Transparency and Accountability in Ai Decision Support: Explaining and Visualizing Convolutional Neural Networks for Text Information," *Decision Support Systems* (134), 113302.
- Kim, T., Barasz, K., and John, L. K. 2019. "Why Am I Seeing This Ad? The Effect of Ad Transparency on Ad Effectiveness," *Journal of Consumer Research* (45:5), 906-932.

- Kizilcec, R. F. 2016. How Much Information? Effects of Transparency on Trust in an Algorithmic Interface, *In: Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*, San Jose, CA, USA
- Komiak, S. Y. X., and Benbasat, I. 2006. "The Effects of Personalization and Familiarity on Trust and Adoption of Recommendation Agents," *MIS Quarterly* (30:4), 941-960.
- Koufaris, M., and Hampton-Sosa, W. 2004. "The Development of Initial Trust in an Online Company by New Customers," *Information & Management* (41:3), 377-397.
- Kulesza, T., Stumpf, S., Burnett, M., Yang, S., Kwan, I., and Wong, W.-K. 2013. Too Much, Too Little, or Just Right? Ways Explanations Impact End Users' Mental Models, *In: Proceedings of the 2013 IEEE Symposium on Visual Languages and Human-Centric Computing (VL/HCC)*, San Jose, CA, USA
- Lam, C. F., DeRue, D. S., Karam, E. P., and Hollenbeck, J. R. 2011. "The Impact of Feedback Frequency on Learning and Task Performance: Challenging the "More Is Better" Assumption," *Organizational Behavior and Human Decision Processes* (116:2), 217-228.
- Lankton, N., McKnight, D. H., and Thatcher, J. B. 2014. "Incorporating Trust-in-Technology into Expectation Disconfirmation Theory," *The Journal of Strategic Information Systems* (23:2), 128-145.
- Lankton, N., McKnight, D. H., and Tripp, J. 2015. "Technology, Humanness, and Trust: Rethinking Trust in Technology," *Journal of the Association for Information Systems* (16:10), 880-918.
- Lee, E.-J. 2008. "Flattery May Get Computers Somewhere, Sometimes: The Moderating Role of Output Modality, Computer Gender, and User Gender," *International Journal of Human-Computer Studies* (66:11), 789-800.
- Lee, J. D., and See, K. A. 2004. "Trust in Automation: Designing for Appropriate Reliance," *Human Factors* (46:1), 50-80.
- Lurie, N. H., and Swaminathan, J. M. 2009. "Is Timely Information Always Better? The Effect of Feedback Frequency on Decision Making," *Organizational Behavior and Human Decision Processes* (108:2), 315-329.
- Madhavan, P., Wiegmann, D. A., and Lacson, F. C. 2006. "Automation Failures on Tasks Easily Performed by Operators Undermine Trust in Automated Aids," *Human Factors* (48:2), 241-256.
- Manzey, D., Reichenbach, J., and Onnasch, L. 2012. "Human Performance Consequences of Automated Decision Aids: The Impact of Degree of Automation and System Experience," *Journal of Cognitive Engineering and Decision Making* (6:1), 57-87.
- Matt, C., Trenz, M., Cheung, C. M. K., and Turel, O. 2019. "The Digitization of the Individual: Conceptual Foundations and Opportunities for Research," *Electronic Markets* (29:3), 315-322.
- McGuinness, D. L., and Da Silva, P. P. 2004. "Explaining Answers from the Semantic Web: The Inference Web Approach," *Journal of Web Semantics* (1:4), 397-413.
- McKnight, D. H., Carter, M., Thatcher, J. B., and Clay, P. F. 2011. "Trust in a Specific Technology: An Investigation of Its Components and Measures," *ACM Transactions on Management Information Systems* (2:2), 12:11-25.
- Merritt, S. M., and Ilgen, D. R. 2008. "Not All Trust Is Created Equal: Dispositional and History-Based Trust in Human-Automation Interactions," *Human Factors* (50:2), 194-210.
- Meske, C., and Bunde, E. 2020. Transparency and Trust in Human-Ai-Interaction: The Role of Model-Agnostic Explanations in Computer Vision-Based Decision Support, *In: Proceedings of the International Conference of AI in HCI2020*, Copenhagen, Denmark.
- Meske, C., Bunde, E., Schneider, J., and Gersch, M. 2022. "Explainable Artificial Intelligence: Objectives, Stakeholders, and Future Research Opportunities," *Information Systems Management* (39:1), 53-63.
- Moussawi, S., Koufaris, M., and Benbunan-Fich, R. 2021. "How Perceptions of Intelligence and Anthropomorphism Affect Adoption of Personal Intelligent Agents," *Electronic Markets* (31:2), 343-364.
- Natarajan, M., and Gombolay, M. 2020. "Effects of Anthropomorphism and Accountability on Trust in Human Robot Interaction," in: *Proceedings of the 2020 ACM/IEEE International Conference*

- on Human-Robot Interaction. Cambridge, United Kingdom: Association for Computing Machinery.
- Ososky, S., Sanders, T., Jentsch, F., Hancock, P., and Chen, J. Y. C. 2014. Determinants of System Transparency and Its Influence on Trust in and Reliance on Unmanned Robotic Systems, *In: Proceedings of SPIE Defense + Security*, Baltimore, Maryland, USA.
- Pak, R., Fink, N., Price, M., Bass, B., and Sturre, L. 2012. "Decision Support Aids with Anthropomorphic Characteristics Influence Trust and Performance in Younger and Older Adults," *Ergonomics* (55:9), 1059-1072.
- Papamichail, K., and French, S. 2003. "Explaining and Justifying the Advice of a Decision Support System: A Natural Language Generation Approach," *Expert Systems with Applications* (24:1), 35-48.
- Parra-Frutos, I. 2013. "Testing Homogeneity of Variances with Unequal Sample Sizes," *Computational Statistics* (28:3), 1269-1297.
- Peters, F., Pumplun, L., and Buxmann, P. 2020. Opening the Black Box: Consumer's Willingness to Pay for Transparency of Intelligent Systems, *In: Proceedings of the 28th European Conference on Information Systems (ECIS)*, An Online AIS Conference.
- Sanders, T. L., Wixon, T., Schafer, K. E., Chen, J. Y. C., and Hancock, P. A. 2014. The Influence of Modality and Transparency on Trust in Human-Robot Interaction, *In: Proceedings of the 2014 IEEE International Inter-Disciplinary Conference on Cognitive Methods in Situation Awareness and Decision Support*, San Antonio, TX, USA.
- Schmidt, P., Biessmann, F., and Teubner, T. 2020. "Transparency and Trust in Artificial Intelligence Systems," *Journal of Decision Systems* (29:4), 260-278.
- Seong, Y., and Bisantz, A. M. 2008. "The Impact of Cognitive Feedback on Judgment Performance and Trust with Decision Aids," *International Journal of Industrial Ergonomics* (38:7), 608-625.
- Shibl, R., Lawley, M., and Debuse, J. 2013. "Factors Influencing Decision Support System Acceptance," *Decision Support Systems* (54:2), 953-961.
- Söllner, M., Hoffmann, A., Hoffmann, H., Wacker, A., and Leimeister, J. M. 2012. Understanding the Formation of Trust in It Artifacts, *In: Proceedings of the 33rd International Conference on Information Systems (ICIS 2012)*, Orlando, FL, USA.
- Wagner, W. P. 2017. "Trends in Expert System Development: A Longitudinal Content Analysis of over Thirty Years of Expert System Case Studies," *Expert Systems with Applications* (76), 85-96.
- Wang, D., Yang, Q., Abdul, A., and Lim, B. Y. 2019. Designing Theory-Driven User-Centric Explainable Ai, *In: Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems* Glasgow, Scotland UK.
- Wang, W., and Benbasat, I. 2008. "Attributions of Trust in Decision Support Technologies: A Study of Recommendation Agents for E-Commerce," *Journal of Management Information Systems* (24:4), 249-273.
- Wang, W., and Benbasat, I. 2016. "Empirical Assessment of Alternative Designs for Enhancing Different Types of Trusting Beliefs in Online Recommendation Agents," *Journal of Management Information Systems* (33:3), 744-775.
- Wang, W., Qiu, L., Kim, D., and Benbasat, I. 2016. "Effects of Rational and Social Appeals of Online Recommendation Agents on Cognition- and Affect-Based Trust," *Decision Support Systems* (86), 48-60.
- Weiler, S., Matt, C., and Hess, T. 2021. "Immunizing with Information – Inoculation Messages against Conversational Agents' Response Failures," *Electronic Markets* (forthcoming, <https://doi.org/10.1007/s12525-021-00509-9>).
- Weith, H. V. K., and Matt, C. 2022. When Do Customers Perceive Artificial Intelligence as Fair? An Assessment of Ai-Based B2c E-Commerce, *In: HICSS 2022 Proceedings*, Maui, HI, USA.
- Wright, J. L., Chen, J. Y. C., and Lakhmani, S. G. 2020. "Agent Transparency and Reliability in Human–Robot Interaction: The Influence on User Confidence and Perceived Reliability," *IEEE Transactions on Human-Machine Systems* (50:3), 254-263.

- Xu, J. D., Cenfetelli, R. T., and Aquino, K. 2016. "Do Different Kinds of Trust Matter? An Examination of the Three Trusting Beliefs on Satisfaction and Purchase Behavior in the Buyer–Seller Context," *The Journal of Strategic Information Systems* (25:1), 15-31.
- Yang, X. J., Wickens, C. D., and Hölttä-Otto, K. 2016. How Users Adjust Trust in Automation: Contrast Effect and Hindsight Bias, *In: Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, Washington, DC, USA.
- Zhou, T. 2011. "An Empirical Examination of Initial Trust in Mobile Banking," *Internet Research* (21:5), 527-540.