Association for Information Systems

AIS Electronic Library (AISeL)

ICIS 2019 Proceedings

Smart Service Systems and Service Science

Following the Robot? Investigating Users' Utilization of Advice from Robo-Advisors

Christoph Tauchert
Technische Universität Darmstadt, tauchert@is.tu-darmstadt.de

Neda Mesbah
Technische Universität Darmstadt, mesbah@is.tu-darmstadt.de

Following the Robot? Investigating Users' Utilization of Advice from Robo-Advisors

Completed Research Paper

Christoph Tauchert
TU Darmstadt
Hochschulstr. 1
64289 Darmstadt, Germany
tauchert@is.tu-darmstadt.de

Neda Mesbah
TU Darmstadt
Hochschulstr. 1
64289 Darmstadt, Germany
mesbah@is.tu-darmstadt.de

Abstract

Companies are gradually creating new services such as robo-advisors (RA). However, little is known if users actually follow RA advice, how much the fit of RA to task requirements influences the utilization, how users perceive RA characteristics and if the perceived advisor's expertise is influenced by the user's expertise. Drawing on judge-advisor systems (JAS) and task-technology fit (TTF), we conducted an experimental study to measure actual advice-taking behavior in the context of RA. While the perceived advisor's expertise is the most influential factor on task-advisor fit for RA and human advisors, integrity is a significant factor only for human advisors. However, for RA the user's perception of the ability to make decisions efficiently is significant. In our study, users followed RA more than human advisors. Overall, our study connects JAS and TTF to predict advice utilization and supports companies in service promotion.

Keywords: Artificial Intelligence; Robo-Advisor; Advice Taking; Judge-Advisor System; Task-Technology Fit; User's Expertise

Introduction

Current advances in artificial intelligence (AI) are driving companies to develop new services for their customers. Giving an example, in the light of Industry 4.0, manufacturing firms offer their clients the possibility of predictive maintenance or process optimization based on machine data (Rawal 2019). Besides these newly offered services, enterprises are also transforming established services by empowering them using machine learning and make them scalable by taking the human out of the loop. An example of these kinds of changes can be seen in the traditional service sector, such as the legal or financial industry. Typically, legal or financial advisory is done by experts who advise you on how you should act with regard to your specific needs. In recent years, empowered by AI, companies have developed services that give personalized advice using an information system instead of a human expert (HSBC 2017).

In the literature, there is no common definition for AI. Russel and Norvig (2009) define AI based on two dimensions. The first dimension addresses the thought process and behavior, while the second one is concerned with whether success is measured against human performance or against an ideal (rational) performance. The combination of these two dimensions leads to four characteristics which can describe and define an AI-based system: thinking humanly, thinking rationally, acting humanly, or acting rationally (Russel and Norvig 2009). In our context of AI-based advisory, we define AI as a system, which is able to learn, makes rational predictions, and interacts like a human. AI differs significantly from other traditional technologies since AI-based systems do not just follow predefined static rules but have the ability to learn from data (Burrell 2016). Some advantages of AI-based systems are efficiency and scalability (Brundage et al. 2018). In comparison to human advisors, AI-based advisors are not able to explain their recommendation, which is also known as black-box behavior, but due to technological advances, AI-based algorithms can process, utilize, and learn from more information than any human advisor could do in appropriate time because of cognitive constraints (Simon 1972).

A much-discussed example of AI-based advisory, in research and practice, is financial robo-advisory which causes significant changes in the financial industry (Jung, Dorner, Glaser, et al. 2018; Jung and Weinhardt 2018; Sironi 2016). Robo-advisors are automated investment advisory services. Customers are guided through a self-assessment process and are then recommended a target-oriented investment strategy with regard to possible portfolio compositions or estimated stock performances (Jung, Dorner, Glaser, et al. 2018; Sironi 2016; Tertilt and Scholz 2017). If robo-advisors were accepted by users[1], benefits would arise both for providers as well as for users. Due to the simple scalability of advisory services as well as the significant reduction of investment costs, the deployment of robo-advisors is highly attractive for financial service companies like banks (Tertilt and Scholz 2017). By using robo-advisors, users can also reduce their investment costs and perform real-time portfolio surveillance (Tertilt and Scholz 2017).

Assuming robo-advisors can provide good advice, it is not guaranteed that people will necessarily utilize such advice. In the information systems (IS) literature robo-advisors were mostly investigated focusing on the design and architecture of these services as well as related business models (Eickhoff et al. 2017; Jung, Dorner, Weinhardt, et al. 2018; Jung and Weinhardt 2018; Riasanow et al. 2018). Whereas, the exploration of users' perception of robo-advisors was neglected. In the cognitive sciences, the judge-advisor system (JAS) paradigm has been used to investigate the advice taking and giving behavior of people. Although various factors were examined within this research stream, almost exclusively the interaction between human decision makers and human advisors was regarded (Bonaccio and Dalal 2006). In the IS literature, the task-technology fit (TTF) is used to determine how well a technology is suited to assist a person in performing a task (Goodhue and Thompson 1995). However, based on this model, we cannot assess if AI-based advice is utilized differently than human advice. By integrating the TTF model in the JAS context, we want to generate a holistic view to understand the factors leading to advice accepting behavior of AI-based advisory services. Therefore, we examine if users accept a substitution of human financial advisors by robo-advisors and if the investment advice will be at least similarly utilized. This leads us to the following research questions:

RQ 1: *Are there differences in users' advice utilization of robo- and human advice?*

RQ 2: *Is the users' advice utilization affected by the fit of task and advisor as well as how this fit is affected by the advisor's characteristics?*

Since the topic of finance and financial planning concerns the general population (Beketov et al. 2018), it is natural that both experienced and inexperienced individuals might use robo-advisors. Users' perceived expertise was already discussed within the TTF as well as the JAS literature (Harvey and Fischer 1997; Parkes 2013). JAS researchers have shown that experienced decision makers have higher advice utilization when making important decisions (Harvey and Fischer 1997). Furthermore, Parkes (2013) found that users' expertise affects the perception of technology characteristics. Therefore, we explore if users' self-perceived expertise has an impact on the perceived expertise of human and robo-advisors. Consequently, our third research question is:

RQ 3: *Does the users' self-perceived expertise affect the perceived advisor's expertise?*

We are following the call of Rzepka and Berger (2018) to investigate users' interactions with robo-advisors by answering these three research questions. The remainder of this manuscript is structured as follows: To begin with, we provide an overview of the theoretical background related to advice taking and the task-technology fit. Then, we derive hypotheses before describing our online experimental survey study design. After introducing our study sample consisting of 197 participants, we present the collected and analyzed data using group comparison and partial least square (PLS). Thereby, the discussion of findings illustrates contributions to research and practice. Lastly, we conclude the manuscript by summarizing the most important findings as well as pointing out the limitations of our research and proposing specific avenues for future research.

[1] User is defined as user of a robo-advisor and used synonymously to decision maker and judge.

Theoretical Background

Advice Giving and Taking

Within the cognitive sciences, the phenomenon of people giving and taking advice is investigated under the judge-advisor system (JAS) paradigm (Bonaccio and Dalal 2006). It describes a structured group in which one individual (i.e., the judge or decision maker) holds the sole decision power and seeks advice from one or more advisors (Van Swol 2011). Within this context, various studies have investigated which factors influence the judge's advice utilization, i.e., the extent to which decision makers follow the advice they receive from experts (Bonaccio and Dalal 2006). A robust finding has been *egocentric discounting,* which means that decision makers tend to adjust their initial estimate by just 20% to 30% towards the advisor's suggestion (Harvey and Fischer 1997). In addition, several factors such as trust, competence, distance of advice, power or source of advice have been identified as influencing the advice-taking behavior of decision makers (Bonaccio and Dalal 2006; Schultze et al. 2015; Sniezek and Buckley 1995; Van Swol and Sniezek 2005; White 2005).

One of the most discussed advisor characteristics influencing advice-taking is trust (Jungermann 1999; Van Swol 2011). Trust is "the willingness of a party to be vulnerable to the actions of another party based on the expectation that the other will perform a particular action important to the trustor, irrespective of the ability to monitor or control that other party" (Mayer et al. 1995, p. 712). Since trust is a rather abstract concept, most researchers agree that it has to be studied multi-dimensionally (Komiak and Benbasat 2006; Rousseau et al. 1998). Komiak and Benbasat (2006) categorized trust in three dimensions: (1) cognitive trust in competence, (2) cognitive trust in integrity, and (3) emotional trust. Following their definition, several studies analyzed the impact of an **advisor's competence**, also called **expertise**, on the decision maker's advice utilization. Advisor's competence is defined as the advisor's perceived ability to provide good advice in a specific domain (Mayer et al. 1995). Customers' main concern is whether the advisor has the competence required to provide them with relevant and customized advice (Komiak and Benbasat 2006). Studies have shown that decision maker, who perceive their advisor as competent are more willing to adjust their initial opinion in favor of the advisor's opinion. (Kim et al. 2017; Schultze et al. 2015). **Integrity** is defined as the honesty of the advisor and describes the decision maker's expectation that the advisor acts in his/her interest (McKnight et al. 2002). Consequently, it refers to the extent that the user perceives the advice as objective and unbiased (Komiak and Benbasat 2006). A robo-advisor can be designed in a way that it only recommends products that are most profitable for the service provider who owns the robo-advisor. Such kind of robo-advisor would be considered to have a low integrity. Studies show that the higher the perceived integrity of the advisor is, the more likely it is that the advice will be used (Van Swol 2011). Lastly, **emotional trust** describes the decision maker's feelings of security and comfort about relying on an advisor (Komiak and Benbasat 2006). Similar to the previous dimensions of trust, the stronger the emotional trust in the advisor is, the more likely the judge is to follow the advice (Sniezek and Van Swol 2001).

Besides the advisor's expertise, also the decision maker's expertise was investigated. While generally it is assumed, that the decision maker has a lower competence than the advisor, decision makers with less expertise had higher trust levels and more variability in their trust rating (Sniezek and Van Swol 2001). It has been shown that experienced users tend to follow advice less than inexperienced users (Harvey and Fischer 1997). For incompetent decision makers it is difficult to assess whether a particular advice is good or bad (Ehrlinger et al. 2008), which is a huge challenge in the judge-advisor relationship.

Concluding, in the JAS literature, advisor's and decision maker's expertise, advisor's integrity as well as emotional trust in the advisor have been identified to influence advice utilization. Based on the task-technology fit model, we want to combine these factors. Therefore, we will introduce the TTF model in the next section.

Task Technology Fit

For a technology to be adopted by the user it has to be utilized and it must help the user to achieve his/her goal in a specific task. A well-known theory used for this phenomenon is the task-technology fit (TTF) (Goodhue 1995; Goodhue and Thompson 1995).

The task-technology fit model consists of the components *task*, *technology*, the *fit,* and the *utilization of the information system*. A task describes any action that is carried out to turn inputs into outputs. Relevant task characteristics are those that influence individuals to use or not to use a technology (e.g., difficulty, significance, routineness). A technology is a tool that an individual uses to carry out a task. The fit describes the appropriateness in which a technology helps the individual to succeed in a task. This fit should serve as a predictor for the utilization of an information system since individuals are more likely to use a technology that they perceive to be suitable to assist in solving the task (Goodhue and Thompson 1995). Figure 1 depicts the general idea of the model.

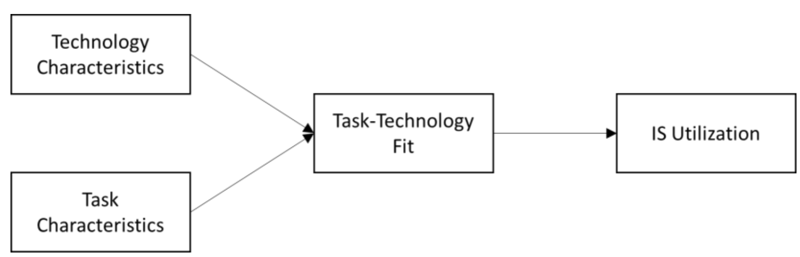

Figure 1. Task-Technology Fit Model
(adapted from Goodhue 1995)

The TTF was already used in various contexts to investigate the success of new technologies including answering managerial questions (Goodhue et al. 2000), online shopping (Klopping and Mckinney 2004), question-answering systems (Robles-Flores and Roussinov 2012) and group support systems (Zigurs et al. 1999; Zigurs and Buckland 1998). However, until now it was not used to evaluate the setting of robo-advisory systems.

Research Model

The purpose of this manuscript is to investigate and compare the behavior of individuals when interacting with robo-advisors and human experts in a financial planning context using the judge-advisor system. Until now, the JAS paradigm was almost exclusively used in a setting where both the judge and the advisor were human beings. However, there is one study which investigated how individuals utilize advice that is deducted from a statistical model (Önkal et al. 2009). Although the advice was presented in the exact same way for the statistical method and the human advisor, the participants discounted the statistical advice more than the same advice from a human expert. While robo-advisors are also mostly based on statistical methods, they have more capabilities. As some studies have shown, they might be perceived differently since human characteristics are perceived in AI-based applications (Rzepka and Berger 2018). Furthermore, compared to human experts, AI algorithms are able to process a vast amount of information in real-time and can incorporate the resulting insights in their advice (Anthes 2017). This implies that robo-advisors must be seen as more than purely statistical tools and this could lead to an increased reliance on robo-advisors due to a perceived superiority:

H1: *Advice from robo-advisors is utilized more than advice from human experts.*

The TTF describes the fit between task characteristics and a technology. By adapting this to the JAS context, the task-advisor fit (TAF) describes the fit between task characteristics and an advisor. Since TTF is a predictor for IS utilization (Goodhue and Thompson 1995), we assumed that TAF would be a predictor for advice utilization:

H2: *A higher task-advisor fit is related to higher advice utilization.*

From the JAS literature, we know that trust is identified as one of the most important factors that lead to advice utilization. Other characteristics such as age (Feng and MacGeorge 2006) and similarity to the decision maker (Gino et al. 2009) were also investigated. Many of these factors are not directly transferable to robo-advisors. Therefore, we focused on the advisor characteristics that can be perceived in a human advisor as well as in a robo-advisor.

To validate the advisor characteristics that were identified through the literature review, we conducted a pre-test among 67 persons. We asked the participants (1) what characteristics they see in a human advisor, (2) what characteristics they associate with a robo-advisor and (3) what differences between those two types of advisors they perceive. The open answers were coded by three IS researchers and as a result, we can confirm the literature-based characteristics but also found that efficiency-enhancing was an often mentioned characteristic, that describes the extent to which an advisor enables efficient decision-making. Therefore, we considered four advisor characteristics: expertise, emotional trust, integrity, and efficiency-enhancing.

As mentioned before, studies have shown that decision maker, who perceive their advisor as competent are more willing to adjust their initial opinion in favor of the advisor's opinion. (Kim et al. 2017; Schultze et al. 2015). Furthermore, advisors with higher expertise are able to assess the difficulties and challenges of a task better and thus, are more suitable to solve a task successfully. Therefore, we hypothesized:

H3a: *For the robo-advisor, a higher perceived advisor expertise is related to higher task-advisor fit.*

H3b: *For the human advisor, a higher perceived advisor expertise is related to higher task-advisor fit.*

A great advantage of robo-advisors is their ubiquity since they are available for consultation 24/7 other than human financial advisors. Furthermore, they provide advice instantaneously because of their superior data processing capabilities. Therefore, robo-advisors enable users to make investment decisions more efficiently. In the case of the human advisor, efficiency will not be a decisive factor when it comes to whether the advisor is perceived as suitable. Nonetheless, due to the access to the advisor's additional expertise, efficiency in decision-making increases. Thus, leading to the following hypotheses:

H4a*: For the robo-advisor, a higher perceived advisor efficiency-enhancing ability is related to higher task-advisor fit.*

H4b: *For the human advisor, a higher perceived advisor efficiency-enhancing ability is related to higher task-advisor fit.*

From the JAS literature, we know that trust has a major influence on advice utilization (Jungermann 1999; Van Swol 2011). Emotional trust describes the feeling of security and comfort about relying on the advisor (Komiak and Benbasat 2006). Thus, the decision maker perceives the advisor as credible and helpful, leading to a positive influence on TAF:

H5a: *For the robo-advisor, a higher emotional trust is related to higher task-advisor fit.*

H5b: *For the human advisor, a higher emotional trust is related to higher task-advisor fit.*

When interacting with an advisor, decision makers cannot be sure of the advisor's intentions. It is not necessarily clear, whether the advisor is advising in the best interest of the client or if he/she acts for their own personal gains. Especially in the context of financial advice this topic gained some media coverage with advisors maximizing their commission fees and kickbacks. Therefore, decision makers will deem the advisor suitable for the task if they perceive them to have a higher integrity:

H6a*: For the robo-advisor, a higher perceived advisor integrity is related to a higher task-advisor fit.*

H6b*: For the human advisor, a higher perceived advisor integrity is related to a higher task-advisor fit.*

As described before, a main problem of incompetent decision makers is to evaluate whether the received advice is correct and useful (Ehrlinger et al. 2008). Therefore, only if a user has a certain task knowledge, he/she can assess the expertise of an advisor. Thus, we concluded:

H7a: *For the robo-advisor, a higher self-perceived user expertise is related to higher perceived advisor expertise.*

H7b: *For the human advisor, a higher self-perceived user expertise is related to higher perceived advisor expertise.*

Since the focus of our study lies in the perceptual differences of robo- and human advisors, we assessed the model using only one task and did not manipulate any task characteristics. Nonetheless, we measured a set of task characteristics such as significance as well as difficulty (Petter et al. 2013) and we did not find any variation in the task characteristics. Finally, Figure 2 shows the final research model, which builds the foundation of our study.

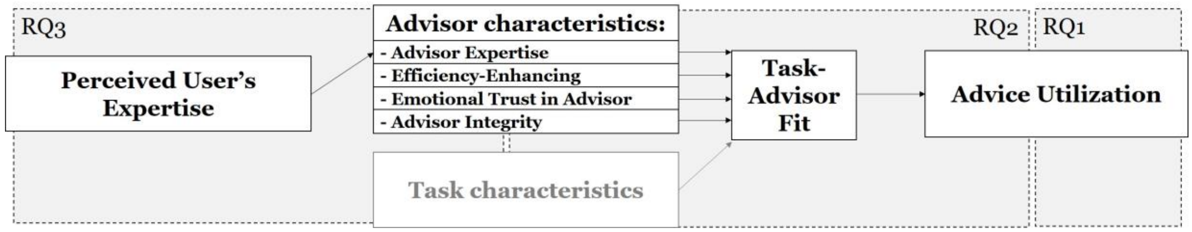

Figure 2. Research Model

Research Method

In order to investigate the differences between the utilization of advice from a robo-advisory system and a human expert, we set up an online experimental survey. Our goal was to measure the participants' actual behavior during the interaction with the advisor instead of their self-reported perception as it was called upon by Rzepka and Berger (2018). We developed an experiment following the approach of many studies in the JAS context (e.g., Gino and Moore 2007). The participants were randomly assigned into two groups whereby one group was instructed that their advice comes from a human expert and the other group was told that the advice is given by a robo-advisor. In order to motivate the participants to reveal their true intentions, they had the chance to win up to 2 Euro during the experiment if they perform well (Camerer and Hogarth 1999).

To acquire a diverse and highly representative (in terms of age, gender, and occupation) sample of internet users for our study, we used a market research company (Lowry et al. 2016). The participants received an incentive of 0.5 Euro from the agency regardless of their performance during the experiment. At the landing page of our study, the participants were instructed that they participate in a scientific study, that their data is stored anonymously and, that, besides the experiment task, there are no right or wrong answers. This was done to counter common methodological biases (Podsakoff et al. 2003).

Experiment Description

To answer the research questions and validate our hypotheses, we decided the context of stock prediction to be a good fit for our survey. While this approach does not reflect the typical interaction process of clients and robo-advisors, it allowed us to apply a widely recognized measure within the JAS research stream (i.e., the weight of advice). Furthermore, this use case deemed to be appropriate for four reasons: (1) Users are familiar with stock prices due to daily news coverage. (2) The prediction of stock prices is a part of robo-advisory services since it is necessary to recommend a good stock portfolio. (3) The prediction of stock prices is not just a knowledge-based task due to the high uncertainty of stock markets (Dzielinski 2012). (4) Finally, it is also very important that advice is reliable as it has a long-term negative effect in the event of failure for users (Lee 2009).

The study was structured as followed: At first, we collected the participants' demographics before having them answering some self-assessment constructs. Then, a description of the experiment scenario was shown to the participants: „Imagine: You want to invest in company shares and must, therefore, forecast the performance of various stocks. Your task is to estimate how a particular stock will perform within a year." Furthermore, they got the information that they will see real historical stock valuation charts from the recent past and that the closer their final estimation is to the real stock valuation the higher their compensation will be. The experiment roughly consisted of three repeating steps per stock:

1. Analyzing the provided stock chart and giving an independent initial estimation.
2. Getting the valuation estimation of an expert, which was either a human or robo-advisor.

3. After receiving the expert's opinion, the participants were free to adjust their initial estimation. They were explicitly told that they could but do not need to change their personal estimation.

After the scenario description, the experiment began and the participants were sequentially shown five charts showing a 3-year historical (*t* to *t+3*) stock performance of enterprises out of five different industries (i.e., aviation, pharmaceutical, automotive, technology, and energy). Additionally, the participants received a small description (one sentence) about the enterprise. We withheld the information of the exact timeframe and the companies' names to avoid that individual experiences were weigh in that might distort advice utilization (Önkal et al. 2009). After each chart, the participants had to guess the stock valuation a year later (*t+4*). After they had estimated the last stock value, the participants were told that they now get professional advice from an expert. The first group was informed that the expert is a professional (human) financial advisor, who had a profound education in finance and is founding advice on his/her experience, current news, and economic developments. The second group was told the advice comes from a robo-advisor, an application based on AI that uses historical stock data, analyzes current news as well as economic developments to generate an advice. The provided advice was the same for both groups and corresponded to the true stock valuation. Afterward, the participants were again shown the charts sequentially with the additional information of their initial estimation as well as the advisor's estimation. The participants were asked to give a final estimation of the expected stock valuation.

Items

To measure the degree of advice utilization we used the *weight of advice* (WOA), which has been used in several studies (e.g., Gino and Moore 2007; Önkal et al. 2009; Sah et al. 2013; Schultze et al. 2015):

WOA = (final estimate – initial estimate) / (advice – initial estimate)

The WOA measures to what extent an individual utilizes an advice in his final estimation by dividing the distance of final and initial estimate by the distance of advice and initial estimate (Yaniv 2004). For rational decision makers the WOA is supposed to be in the range of 0 and 1. 0 meaning that the participant completely ignored the advice and did not adjust his/her initial estimate and 1 implicating that the decision maker completely adopted the advice. Values in-between 0 and 1 indicate partial incorporation of the advice in the final estimate, whereby a value of 0.5 means that a participant has calculated the mean between his/her initial estimate and the advice and weighs his/her opinion just as much as the advisor's. Irrational decision makers can have WOA measure under 0 or over 1, meaning that either moved in the opposite direction of the advice or that he/she even over-utilized the advice. However, these cases occur rarely (Gino and Moore 2007; Harvey and Fischer 1997). We calculated the mean WOA using the five measured WOA values for each participant.

For the evaluation of the constructs, we have used measurements from the established literature. We used the scales of Komiak and Benbasat (2006) to measure emotional trust and cognitive trust in integrity. To measure trust in competence we adapted the scale of McKnight et al. (2002). The perceived efficiency-enhancing ability of the advisor was measured using the construct of Chan et al. (1997), while we used Moore and Benbasat's (1991) scale for the task-advisor fit. Finally, we adopted the item of Radel et al. (2011) to measure user's self-perceived task expertise. All of our items were measured using a 7-point Likert scale ranging from 'strongly disagree' to 'strongly agree' and can be found in the appendix in Table 4. Additionally to the items of our main constructs, we measured tendency towards fantasizing as marker variable to counteract common method bias (Podsakoff et al. 2003) based on the three-item scale of Darrat et al. (2016).

Results

In our study 247 participants took part. We included several checks – manipulation check and rationality check – to guarantee the quality of the study's results (Meade and Craig 2012). During the rationality check, we excluded all participants who had a WOA over 1 or under 0. We excluded 21 participants due to failing the manipulation check. After excluding 29 more participants who failed the rationality check, our sample consisted of 197 responses, which could be used for further analysis. The demography of our sample reflects the typical European internet users quite accurately by age, gender, and employment status (Eurostat 2018). 93 females and 104 males took part in our study with an average age of 38.54 years ranging from 18 to 68 years. 58.4% of our participants were employees and 11.2% students. From our remaining participants, 104 were assigned to the group with the robo-advisor while 93 participants were assigned to the group with the human advisor. In order to compare the behavior of both groups, we first ensured that the groups had perceived task characteristics equally and that user's self-perceived expertise was not significantly different by using an independent t-test.

H1 hypothesized that advice from the robo-advisor would be more utilized. To test H1, we ran an independent t-test of WOA. The result of the t-test ($t(195) = 1.771$, $p = .039$) showed that the advice of the robo-advisor was statistically more utilized ($M = .44$, $SD = .253$) than those of a human advisor ($M = .38$, $SD = .260$). Concluding, H1 is supported.

To evaluate H2 to H7, we analyzed our research model based on a well-establish method (Qureshi and Compeau 2009) by comparing the structural equation model of each group through a variance-based partial least squares multi-group analysis as implemented in SmartPLS (Ringle et al. 2015). We opted for this approach for two main reasons. (1) This approach is well suited for theories in their early stages (Fornell and Bookstein 1982). (2) It is possible to test both the research models and the path differences simultaneously through multi-group analysis (Brook et al. 1995).

Table 1. Item Loadings

Constructs (measured on 7-point scales)	*Items*	*Item Loadings Robo-Advisor*	*Item Loadings Human Advisor*
Advice Utilization (WOA)	WOA1	1.000	1.000
Task-Advisor Fit (TAF)	TAF1	.861	.919
	TAF2	.952	.936
	TAF3	.936	.913
Advisor Expertise (AEX)	AEX1	.928	.944
	AEX2	.961	.960
	AEX3	.930	.957
	AEX4	.872	.927
Advisor Efficient-Enhancing (EFF)	EFF	1.000	1.000
Emotional Trust in Advisor (EMO)	EMO1	.961	.970
	EMO2	.977	.971
	EMO3	.973	.977
Advisor Integrity (INT)	INT1	.906	.881
	INT2	.931	.916
	INT3	.919	.926
User's Self-Perceived Expertise (UEX)	UEX1	.941	.937
	UEX2	.975	.961
	UEX3	.959	.959
	UEX4	.938	.929

By determining convergent validity (statistical similarity of construct items) and discriminant validity (statistical difference of items that measure different constructs) of our research model we validated our measurement model (Hair et al. 2013). We confirmed convergent validity by examining item loadings, Cronbach's α, and composite reliability (CR) as well as the average variance extracted (AVE) by the constructs (Xu et al. 2012). The item loadings are reported in Table 1 and all loadings are above the threshold value of 0.7 (Hair et al. 2013). For each construct the Cronbach's α and composite reliability values achieve the threshold of 0.7 and AVE values threshold of 0.5 (Hair et al. 2011) as can be seen in Table 2.

We assessed the cross loadings as well as the square root of the AVE for each construct model and therefore, we confirmed discriminant validity (Fornell and Larcker 1981). As reported in Table 2, all constructs' square roots of the AVE are higher than their correlation to another construct. The loading of each item is greater to its associated construct than to other constructs, but we do not report the cross loadings due to space limitations.

Table 2. Cronbach's α (Cr. α), Composite Reliability (CR), Average Variance Extracted (AVE) and Construct Correlations
(First Row: Robo-Advisor; Second Row: Human Advisor)

Constructs	*Cr. α*	*CR*	*AVE*	*WOA*	*TAF*	*ACOM*	*EFF*	*EMO*	*INT*	*SCOM*
Advice Utilization (WOA)	1.000 1.000	1.000 1.000	1.000 1.000	**1.000** **1.000**						
Task-Advisor Fit (TAF)	.905 .913	.941 .945	.841 .851	.233 .300	**.917** **.923**					
Advisor Expertise (AEX)	.942 .962	.958 .972	.852 .897	.396 .334	.789 .783	**.923** **.947**				
Advisor Efficiency-Enhancing (EFF)	1.000 1.000	1.000 1.000	1.000 1.000	.173 .255	.724 .737	.723 .768	**1.000** **1.000**			
Emo. Trust in Advisor (EMO)	.969 .971	.980 .981	.941 .946	.421 .435	.686 .663	.790 .703	.724 .711	**.970** **.973**		
Advisor Integrity (INT)	.908 .894	.942 .934	.844 .824	.288 .280	.583 .737	.669 .672	.568 .738	.647 .794	**.918** **.908**	
User's Expertise (UEX)	.967 .962	.976 .972	.909 .896	-.146 .015	.167 .130	.181 .172	.220 .102	.100 .187	.171 .145	**.953** **.947**

Before we test our research model through the multi group analysis, we depict the results of the research model for the full sample by running a bootstrapping with 5,000 re-samples (Davison and Hinkley 1997). As we postulated in H2 higher task-advisor fit relates significantly to a higher advice utilization (ß = .273, p = .000). H2 is supported. We can also find a significant impact of each advisor characteristic on task-advisor fit – for advisor expertise (ß = .480, p = .000), for advisor efficiency-enhancing ability (ß = .253, p = .004) and for advisor integrity (ß = .141, p = .050) – except for emotional trust (ß = .036, p = .732). User's self-perceived expertise has a significant impact on perceived advisor expertise (ß = .175, p = .026). None of our control variables – age, gender, IT background, marker variable for common method bias – changes the significances of our research model or are significant predictors of our dependent variable.

At the beginning of the multi-group analysis, we looked at the model fit of both groups. The model fit SRMR is .056 for the robo-advisor sample and for the human advisor sample .052, which refers to a good model fit since it is under the cut-off value of .08 (Hu and Bentler 1999). Based on our model we are able to explain 5.4% of the variance of the advice utilization and 67.4% of the variance in task-advisor fit in the robo-advisor sample and 9.0% of the variance of the advice utilization and 70.4% of the variance in task-advisor fit in the human advisor sample. The path coefficients, their significance as well as their effect sizes are reported in Table 3.

Likewise to the full sample research model, H2 is supported in the multi-group analysis. As postulated in H3a and H3b, advisor expertise has a significant impact on task-advisor fit. The advisor's efficiency-enhancing ability has a positive significant influence on task-advisor fit for robo-advisors, as assumed in H4a. However, we have to reject H4b since there was no significant influence of the efficiency-enhancing ability on task-advisor fit for human advisors. Contrary to our assumption of H5a and H5b, emotional trust has no significant influence on task-advisor fit for either group. Although advisor's integrity has a significant positive influence on task-advisor fit in the group with the human advisors as postulated in

H6b, it has no significant influence on task-advisor fit in the robo-advisor group against our suggestion of H6a. Finally, we have not observed a significant effect of the user's self-perceived task expertise on advisor expertise. Summarizing the results, we were able to support H2, H3a, H3b, H4a andH6b. All other hypotheses had to be rejected. These findings of our multi-group analysis are visualized in Figure 3.

Table 3. Results of Structural Model Testing and Effect Sizes
*(*** p<0.001; ** p<0.01; * p<0.05; RA=Robo-Advisor, HU=Human Advisor)*

Constructs	*Path Coefficients and p-Values*				*Multi-Group Testing*		*f^2 values*	
	RA	*p*	*HU*	*p*	*Diff.*	*p*	*RA*	*HU*
Task-Advisor Fit → Advice Utilization	.233*	.017	.300*	.011	.067	.673	.057	.099
Advisor Expertise → Task-Advisor Fit	.516***	.000	.461***	.000	.055	.376	.244	.258
Advisor Efficiency-Enhancing → Task-Advisor Fit	.304**	.006	.165	.228	.140	.211	.117	.029
Emotional Trust in Advisor → Task-Advisor Fit	.028	.825	-.057	.726	.085	.343	.001	.003
Advisor Integrity → Task-Advisor Fit	.047	.562	.351**	.002	.305*	.017	.003	.128
User's Expertise → Advisor Expertise	.181	.079	.172	.191	.009	.487	.034	.030

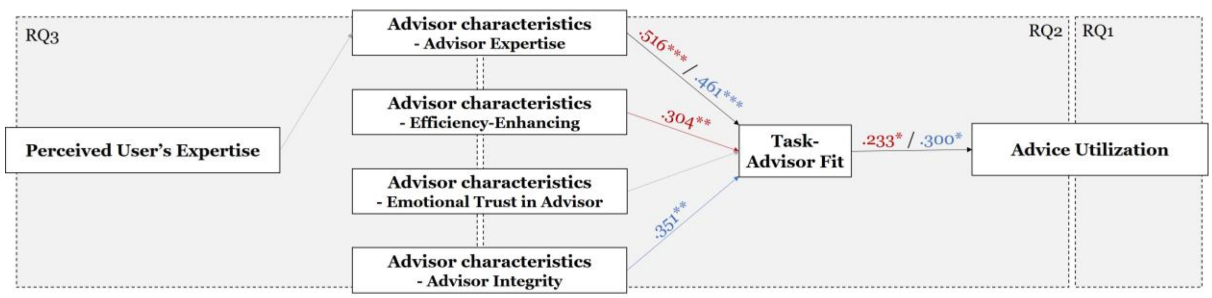

Figure 3. Summary of Structural Model Testing
(Robo-Advisor in Red and Human Advisors in Blue)

Discussion and Contribution

The goal of our research was to investigate (1) whether users utilize advice differently depending on the source of advice (i.e., robo-advisor vs. human expert), (2) if the task-advisor fit affects advice utilization as well as how advisor characteristics influence the task-advisor fit and (3) the influence of users' self-perceived expertise on the perception of the advisor's expertise. To address our research questions, we conducted an experimental study with 197 participants and thereby contributed to the IS advice-taking literature.

Previous studies have shown that the origin of advice can have a significant influence on the user's utilization of advice. It has been shown that advice that is derived from statistical models is discounted more than advice from human experts in a financial setting (Önkal et al. 2009). Other studies that have investigated the perception of 'traditional' computer-generated advice have also found that human advice is trusted more (Wærn and Ramberg 1996). Our experiment's findings showed that the advice of robo-advisors was utilized more than the advice of human experts for the specific setting of stock price predictions. To understand the differences of the findings in our study, we argue that while robo-advisors base their advice mostly on statistical and mathematical calculations one can interact with robo-advisors more naturally due to natural language processing and speech synthesis abilities. Therefore, the advantages of both advisor types are combined. However, this result needs to be validated in further studies and causalities have to be derived.

With regard to the advisor characteristics, we found that in our context different characteristics affect the task-advisor fit for the different advisors. For the robo-advisor, we can see that expertise and efficiency-enhancement are the significant antecedents while for the human advisor, expertise and integrity are contributing to the task-advisor fit. Even though we had to reject H6a (i.e., a positive influence of integrity on TAF) when calculating an independent t-test, we noticed that for robo-advisors a significantly higher integrity is perceived than for human advisors ($m_{AI} = 5.12$, $m_{HU} = 4.09$, $t(195) = 5.74$, $p = .000$). This could be an indication that users suspect the dishonesty of humans but do not believe that robo-advisory services give malicious advice. Nonetheless, by comparing the f^2 values it can be seen that for both cases expertise was the most influential antecedent. Concluding, human and robo-advisors are perceived with different strengths that are influencing the task-advisor fit. Furthermore, it can be seen that the TAF is a predictor for advice utilization and the integration of the TTF model in the JAS paradigm was successful.

Another finding of our study is the influence of the decision maker's self-perceived expertise on the perception of the advisor's expertise. Even though we hypothesized a positive relationship between these two constructs, we do not find a significant effect. This is quite surprising because the decision makers' knowledge in the area of interest should allow them to assess the quality of the given advice better (Ehrlinger et al. 2008). Since our advisor always gave the true estimation, we expected that the competent users would rate the advisor's expertise higher. However, they did not have much information about the advisor and the interaction was different in comparison to a real consultation. Therefore, it might have been difficult to evaluate the advisor's expertise. Additionally, we did not measure the participants' real expertise but rather the self-perceived expertise. This perception could be overestimated. The Dunning-Kruger effect describes that incompetent individuals often do not know that they are incompetent (Kruger and Dunning 1999). So this effect could lead to a false self-assessment of our participants and consequently an overestimation of their expertise.

Besides the theoretical contributions, we also identified various practical implications for professional entities: we found first evidence that AI-based financial advice could be utilized more than human advice by users. This indicates that enterprises can deploy robo-advisors without generally having to fear that customers will reject the suggestions. Furthermore, since the task-advisor fit might be a predictor for the actual advice utilization, organizations can conduct market research surveys to assess the suitability of potential robo-advisory services. Enterprises can leverage our findings about robo-advisor characteristics to adjust service development. They could increase the perceived robo-advisor's expertise, for example, by providing more transparency about the used data or the algorithm so that the assessment of the workflow and performance would be easier. Another option is providing key performance indicators, which enable simpler evaluation of the (historical) performance. Lastly, organizations could emphasize the efficiency of robo-advisors for personal financial planning.

Limitations and Future Research

Naturally, the findings of our study are subject to various limitations. First of all, we selected the setting of robo-advisors as context and used a stock valuation experiment to measure advice utilization. While scenario-based experiments are a common method in IS research (e.g., Önkal et al. 2009; Wang and Benbasat 2007; Ye and Johnson 1995), the findings need to be validated in other robo-advisor tasks such as portfolio composition, especially since stock evaluation is not a typical task during the interaction of financial advisors with customers. Furthermore, since robo-advisors can be used in a variety of domains such as in the legal or insurance industry, one could select a different experimental setting. The investigation of tasks that have different task characteristics (e.g., difficulty, significance, locus of control, (non-)routines) could be very interesting. To give an example, task difficulty has been found to have an impact on advice utilization as well as self-perceived expertise (Ehrlinger et al. 2008; Gino and Moore 2007). If it is possible to predict the perfect advisor characteristics based on the task, promising use cases for AI-based advisors could easily be identified.

Furthermore, we compared the perception of robo-advisors and human ones based on an online experiment. We assumed that participants could put themselves in the situation of a real consultation with a human financial advisor by describing the scenario. It would certainly be useful to validate the findings of our study in a more realistic laboratory experiment where participants would interact with a real human advisor and robo-advisor. The authentic interaction with a human and robo-advisor could lead to different perceptions of advisor characteristics like emotional trust, expertise, or integrity.

There are various other different experimental designs that could also be considered in future works: For example, we did not offer the option to choose between two advisors. It could be interesting to investigate the behavior of users when they have a choice between different advisors. Additionally, our experiment required the advisor to provide a numerical estimation, but there are plenty of other types of advice (e.g., advice for sth., advice against sth., binary advice) that can be studied. Furthermore, user expertise was the sole individual's characteristic that was within the scope of our study. Certainly, various other personality traits may influence the perception of advisor characteristics (e.g., confidence, introversion).

To summarize this section, our task-advisor fit model is a first approach to integrate the TTF model in the JAS to understand users' perceptions of robo-advisors and to evaluate the resulting advice utilization.

Conclusion

Due to current technological developments and advancements in the area of artificial intelligence, AI-based agents are gaining importance in enterprise services rapidly. Such agents can be implemented in a wide variety of fields such as in the healthcare, legal or as in our case the financial industry. The use of robo-advisors is currently gaining momentum, but market shares of such services are still relatively low (Jung and Weinhardt 2018). Therefore, the goal of this manuscript was to investigate users' utilization of advice from robo-advisors. In addition, we wanted to explore if the users' advice utilization is affected by the fit of task and advisor as well as how this fit is affected by the advisor's characteristics. Furthermore, the influence of the user's self-perceived task expertise on the perception of the advisor's expertise was addressed.

By conducting a scenario-based experimental study with 197 participants in a European country, placed in the context of financial advisory, and using performance-based incentives, we were able to measure actual advice utilization. Thus, we were able to show that: (1) Users utilize advice from a robo-advisor differently than advice from a human expert. In our setting the users utilized the advice from robo-advisors more than the advice from human advisors. (2) Users perceive different advisor characteristics for robo- and human advisors. In our experimental setting for the robo-advisor, competence and efficiency were perceived as characteristics that influence the task-advisor fit and for human experts, the significant factors were competence and integrity. (3) The user's self-perceived task expertise has no influence on the perception of the advisor's expertise. Our results help to understand the factors influencing how robo-advisor services are perceived by users and what drives them to utilize the advice from these services. Based on our findings, companies can focus on relevant factors when designing and implementing a robo-advisory service.

Acknowledgments

This research project was funded by the Hessian state ministry "Hessisches Ministerium des Innern und für Sport" in Germany.

Appendix

Table 4. Survey Items

Construct	***Item***		***Adapted from…***
Task-Advisor Fit	TAF1	The expert's[2] advisory service is compatible with all aspects of this task.	(Moore and Benbasat 1991)
	TAF2	The expert's advisory service fits very well with my needs in the task.	
	TAF3	The expert's advisory service fits into my way of decision-making.	
Advisor Expertise	AEX1	The expert is competent and effective in estimating the stock price.	(McKnight et al. 2002)
	AEX2	The expert performs its role of estimation the stock price very well.	
	AEX3	Overall, the expert is a capable and proficient advisor for estimating the stock price.	
	AEX4	In general, the expert is very knowledgeable about the stock price prediction.	
Efficiency-Enhancing	EFF	The expert increases the efficiency of my decision making.	(Chan et al. 1997)
Emotional Trust in Advisor	EMO1	I feel secure about relying on the expert for my decision.	(Komiak and Benbasat 2006)
	EMO2	I feel comfortable about relying on the expert for my decision.	
	EMO3	I feel content about relying on the expert for my decision.	
Advisor Integrity	INT1	The expert provides unbiased recommendations.	(Komiak and Benbasat 2006)
	INT2	The expert is honest.	
	INT3	I consider the expert to be of integrity.	
User's Expertise	UEX1	I feel very competent in the above explained task.	(Radel et al. 2011)
	UEX2	I feel able to meet the challenge of performing well in this task.	
	UEX3	I am able to master this task.	
	UEX4	I am good at doing this task.	

References

Anthes, G. 2017. "Artificial Intelligence Poised to Ride a New Wave," *Communications of the ACM* (60:7), ACM, pp. 19–21.

[2] Depending on the experimental group, the term "expert" is replaced by "human expert" or "robo-advisor" in all items.

Beketov, M., Lehmann, K., and Wittke, M. 2018. "Robo Advisors: Quantitative Methods inside the Robots," *Journal of Asset Management* (19), pp. 363–370.

Bonaccio, S., and Dalal, R. S. 2006. "Advice Taking and Decision-Making: An Integrative Literature Review, and Implications for the Organizational Sciences," *Organizational Behavior and Human Decision Processes* (101:2), pp. 127–151.

Brook, J. S., Whiteman, M., and Cohen, P. 1995. "Stage of Drug Use, Aggression, and Theft/Vandalism: Shared and Unshared Risks," in *Drugs, Crime, and Other Deviant Adaptions: Longitudinal Studies*, H. B. Kaplan (ed.), New York: Springer Science and Business Media, pp. 83–96.

Brundage, M., Avin, S., Clark, J., Toner, H., Eckersley, P., Garfinkel, B., Dafoe, A., Scharre, P., Zeitzoff, T., Filar, B., Anderson, H., Roff, H., Allen, G. C., Steinhardt, J., Flynn, C., HÉigeartaigh, S. Ó., Beard, S., Belfield, H., Farquhar, S., Lyle, C., Crootof, R., Evans, O., Page, M., Bryson, J., Yampolskiy, R., and Amodei, D. 2018. *The Malicious Use of Artificial Intelligence: Forecasting, Prevention, and Mitigation*.

Burrell, J. 2016. "How the Machine 'Thinks': Understanding Opacity in Machine Learning Algorithms," *Big Data & Society* (3:1), SAGE PublicationsSage UK: London, England, p. 205395171562251.

Camerer, C. F., and Hogarth, R. M. 1999. "The Effects of Financial Incentives in Experiments: A Review and Capital-Labor-Production Framework," *Journal of Risk and Uncertainty* (19:1/3), Kluwer Academic Publishers, pp. 7–42.

Chan, Y. E., Huff, S. L., Barclay, D. W., and Copeland, D. G. 1997. "Business Strategic Orientation, Information Systems Strategic Orientation, and Strategic Alignment," *Information Systems Research* (8:2), pp. 125–150.

Darrat, A. A., Darrat, M. A., and Amyx, D. 2016. "How Impulse Buying Influences Compulsive Buying: The Central Role of Consumer Anxiety and Escapism," *Journal of Retailing and Consumer Services* (31), pp. 103–108.

Davison, A. C., and Hinkley, D. V. 1997. *Bootstrap Methods and Their Application*, (10th ed.), New York, NY: Cambridge University Press.

Dzielinski, M. 2012. "Measuring Economic Uncertainty and Its Impact on the Stock Market," *Finance Research Letters* (9:3), Elsevier, pp. 167–175.

Ehrlinger, J., Johnson, K., Banner, M., Dunning, D., and Kruger, J. 2008. "Why the Unskilled Are Unaware: Further Explorations of (Absent) Self-Insight Among the Incompetent.," *Organizational Behavior and Human Decision Processes* (105:1), NIH Public Access, pp. 98–121.

Eickhoff, M., Muntermann, J., and Weinrich, T. 2017. "What Do FinTechs Actually Do? A Taxonomy of FinTech Business Models," in *Thirty Eighth International Conference on Information Systems*, Seoul, pp. 1–19.

Eurostat. 2018. "Internet Access and Use Statistics - Households and Individuals."

Feng, B., and MacGeorge, E. L. 2006. "Predicting Receptiveness to Advice: Characteristics of the Problem, the Advice-Giver, and the Recipient," *Southern Communication Journal* (71:1), Taylor & Francis Group , pp. 67–85.

Fornell, C., and Bookstein, F. L. 1982. "Two Structural Equation Models: LISREL and PLS Applied to Consumer Exit-Voice Theory," *Journal of Marketing Research* (19:4), American Marketing Association, pp. 440–452.

Fornell, C., and Larcker, D. F. 1981. "Evaluating Structural Equation Models with Unobservable Variables and Measurement Error," *Journal of Marketing Research* (18:1), American Marketing Association, pp. 39–50.

Gino, F., and Moore, D. A. 2007. "Effects of Task Difficulty on Use of Advice," *Journal of Behavioral Decision Making* (20:1), pp. 21–35.

Gino, F., Shang, J., and Croson, R. 2009. "The Impact of Information from Similar or Different Advisors on Judgment," *Organizational Behavior and Human Decision Processes* (108:2), Academic Press, pp. 287–302.

Goodhue, D. L. 1995. "Understanding User Evaluations of Information Systems," *Management Science* (41:12), INFORMS , pp. 1827–1844.

Goodhue, D. L., Klein, B. D., and March, S. T. 2000. "User Evaluations of IS as Surrogates for Objective Performance," *Information & Management* (38:2), North-Holland, pp. 87–101.

Goodhue, D. L., and Thompson, R. L. 1995. "Task-Technology Fit and Individual Performance," *MIS Quarterly* (19:2), p. 213.

Hair, J. F., Ringle, C. M., and Sarstedt, M. 2011. "PLS-SEM: Indeed a Silver Bullet," *Journal of Marketing Theory and Practice* (19:2), pp. 139–152.

Hair, J. J. F., Hult, G. T. M., Ringle, C., and Sarstedt, M. 2013. *A Primer on Partial Least Squares Structural Equation Modeling (PLS-SEM)*, Sage Publications.

Harvey, N., and Fischer, I. 1997. "Taking Advice: Accepting Help, Improving Judgment, and Sharing Responsibility," *Organizational Behavior and Human Decision Processes* (70:2), Academic Press, pp. 117–133.

HSBC. 2017. "Trust in Technology."

Hu, L., and Bentler, P. M. 1999. "Cutoff Criteria for Fit Indexes in Covariance Structure Analysis: Conventional Criteria Versus New Alternatives," *Structural Equation Modeling: A Multidisciplinary Journal* (6:1), pp. 1–55.

Jung, D., Dorner, V., Glaser, F., and Morana, S. 2018. "Robo-Advisory: Digitalization and Automation of Financial Advisory," *Business and Information Systems Engineering* (60:1), pp. 81–86.

Jung, D., Dorner, V., Weinhardt, C., and Pusmaz, H. 2018. "Designing a Robo-Advisor for Risk-Averse, Low-Budget Consumers," *Electronic Markets* (28:3), pp. 367–380.

Jung, D., and Weinhardt, C. 2018. "Robo-Advisors and Financial Decision Inertia : How Choice Architecture Helps to Reduce Inertia in Financial Planning Tools," in *Thirty Ninth International Conference on Information Systems*, San Francisco 2018, pp. 1–17.

Jungermann, H. 1999. "Advice Giving and Taking," in *Proceedings of the 32nd Annual Hawaii International Conference on Systems Sciences*, IEEE Comput. Soc.

Kim, H., Benbasat, I., and Cavusoglu, H. 2017. "Online Consumers' Attribution of Inconsistency Between Advice Sources," in *Thirty Eighth International Conference on Information Systems*, Seoul, pp. 1–10.

Klopping, I. M., and Mckinney, E. 2004. "Extending the Technology Acceptance Model Extending the Technology Acceptance Model and the Task and the Task-Technology Fit Model to Technology Fit Model to Consumer E Consumer E-Commerce Commerce," *Information Technology, Learning, and Performance Journal* (Vol. 22).

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), Management Information Systems Research Center, University of Minnesota, pp. 941–960.

Kruger, J., and Dunning, D. 1999. "Unskilled and Unaware of It: How Difficulties in Recognizing One's Own Incompetence Lead to Inflated Self-Assessments.," *Journal of Personality and Social Psychology* (77:6), pp. 1121–34.

Lee, M.-C. C. 2009. "Factors Influencing the Adoption of Internet Banking: An Integration of TAM and TPB with Perceived Risk and Perceived Benefit," *Electronic Commerce Research and Applications* (8:3), Elsevier, pp. 130–141.

Lowry, P. B., D'Arcy, J., Hammer, B., and Moody, G. D. 2016. "'Cargo Cult' Science in Traditional Organization and Information Systems Survey Research: A Case for Using Nontraditional Methods of Data Collection, Including Mechanical Turk and Online Panels," *Journal of Strategic Information Systems* (25:3), pp. 232–240.

Mayer, R. C., Davis, J. H., and Schoorman, F. D. 1995. "An Integrative Model of Organizational Trust," *The Academy of Management Review* (20:3), p. 709.

McKnight, D. H., Choudhury, V., and Kacmar, C. 2002. "Developing and Validating Trust Measures for E-Commerce : An Integrative Typology," *Information Systems Research* (13:3), pp. 334–359.

Meade, A. W., and Craig, S. B. 2012. "Identifying Careless Responses in Survey Data," *Psychological Methods* (17:3), Online First Publication, pp. 437–455.

Moore, G. C., and Benbasat, I. 1991. "Development of Instrument to Measure the Perceptions of Adopting an Information Technology Innovation," *Information Systems Research* (2:3), pp. 192–222.

Önkal, D., Goodwin, P., Thomson, M., Gönül, S., and Pollock, A. 2009. "The Relative Influence of Advice from Human Experts and Statistical Methods," *Journal of Behavioral Decision Making* (22:4), pp. 390–409.

Parkes, A. 2013. "The Effect of Task–Individual–Technology Fit on User Attitude and Performance: An Experimental Investigation," *Decision Support Systems* (54:2), North-Holland, pp. 997–1009.

Petter, S., William, H. D., and McLean, E. R. 2013. "Information System Success: The Quest for Independent Variables," *Journal of MIS* (29:4), pp. 7–62.

Podsakoff, P., MacKenzie, S., Lee, J. Y., and Podsakoff, N. P. 2003. "Common Method Biases in Behavioral Research: A Critical Review of the Literature and Recommended Remedies," *J. Appl. Psychol.* (88:5), p. 879.

Qureshi, I., and Compeau, D. 2009. "Assessing Between-Group Differences in Information Systems Research: A Comparison of Covariance- and Component-Based SEM," *MIS Quarterly* (33:1), pp. 197–214.

Radel, R., Pelletier, L. G., Sarrazin, P., and Milyavskaya, M. 2011. *Restoration Process of the Need for Autonomy: The Early Alarm Stage*.

Rawal, K. 2019. "Servitization – The Future." (https://www.tcs.com/blogs/servitization-the-future, accessed April 26, 2019).

Riasanow, T., Flötgen, R. J., Soto Setzke, D., Böhm, M., and Krcmar, H. 2018. "The Generic Ecosystem and Innovation Patterns of the Digital Transformation in the Financial Industry," *22nd Pacific Asia Conference on Information Systems (PACIS)*.

Ringle, C. M., Wende, S., and Becker, J.-M. 2015. *SmartPLS 3*, Hamburg: SmartPLS GmbH.

Robles-Flores, J. A., and Roussinov, D. 2012. "Examining Question-Answering Technology from the Task Technology Fit Perspective," *Communications of the Association for Information Systems* (30), pp. 439–454.

Rousseau, D. M., Sitkin, S. B., Burt, R. S., and Camerer, C. 1998. "Not so Different after All: A Cross-Discipline View of Trust," *The Academy of Management Review* (23:3), Academy of Management, pp. 393–404.

Russel, S. J., and Norvig, P. 2009. *Artificial Intelligence: A Modern Approach*, (3rd ed.), Upper Saddle River, NJ, USA: Prentice Hall Press.

Rzepka, C., and Berger, B. 2018. "User Interaction with AI-Enabled Systems : A Systematic Review of IS Research," in *Thirty Ninth International Conference on Information Systems*, San Francisco, pp. 1–17.

Sah, S., Moore, D. A., and MacCoun, R. J. 2013. "Cheap Talk and Credibility: The Consequences of Confidence and Accuracy on Advisor Credibility and Persuasiveness," *Organizational Behavior and Human Decision Processes* (121:2), pp. 246–255.

Schultze, T., Rakotoarisoa, A.-F., and Schulz-Hardt, S. 2015. "Effects of Distance between Initial Estimates and Advice on Advice Utilization," *Judgment and Decision Making* (10:2), pp. 144–171.

Simon, H. A. 1972. "Theories of Bounded Rationality," *Decision and Organization* (1:1), pp. 161–176.

Sironi, P. 2016. "The Theory of Innovation: From Robo-Advisors to Goal Based Investing and Gamification," *John Wiley & Sons*.

Sniezek, J. A., and Buckley, T. 1995. "Cueing and Cognitive Conflict in Judge-Advisor Decision Making," *Organizational Behavior and Human Decision Processes* (62:2), Academic Press, pp. 159–174.

Sniezek, J. A., and Van Swol, L. M. 2001. "Trust, Confidence, and Expertise in a Judge-Advisor System," *Organizational Behavior and Human Decision Processes* (84:2), Academic Press, pp. 288–307.

Van Swol, L. M. 2011. "Forecasting Another's Enjoyment versus Giving the Right Answer: Trust, Shared Values, Task Effects, and Confidence in Improving the Acceptance of Advice," *International Journal of Forecasting* (27:1), pp. 103–120.

Van Swol, L. M., and Sniezek, J. A. 2005. "Factors Affecting the Acceptance of Expert Advice," *British Journal of Social Psychology* (44:3), pp. 443–461.

Tertilt, M., and Scholz, P. 2017. "To Advise, or Not to Advise — How Robo-Advisors Evaluate the Risk Preferences of Private Investors," *The Journal of Wealth Management* (Vol. 21), Institutional Investor Journals Umbrella, July 31.

Wærn, Y., and Ramberg, R. 1996. "People's Perception of Human and Computer Advice," *Computers in Human Behavior* (12:1), Pergamon, pp. 17–27.

Wang, W., and Benbasat, I. 2007. "Recommendation Agents for Electronic Commerce: Effects of Explanation Facilities on Trusting Beliefs," in *Journal of Management Information Systems* (Vol. 23), pp. 217–246.

White, T. B. 2005. "Consumer Trust and Advice Acceptance: The Moderating Roles of Benevolence, Expertise, and Negative Emotions," *Journal of Consumer Psychology* (15:2), No longer published by Elsevier, pp. 141–148.

Xu, H., Hock-Hai, T., Tan, B. C. Y., and Agarwal, R. 2012. "Effects of Individual Self-Protection, Industry Self-Regulation, and Government Regulation on Privacy Concerns: A Study of Location-Based Services," *Information Systems Research* (23:4), pp. 1342–1363.

Yaniv, I. 2004. "Receiving Other People's Advice: Influence and Benefit," *Organizational Behavior and Human Decision Processes* (93:1), pp. 1–13.

Ye, L. R., and Johnson, P. E. 1995. "The Impact of Explanation Facilities on User Acceptance of Expert Systems Advice," *MIS Quarterly* (19:2), p. 157.

Zigurs, I., and Buckland, B. K. 1998. "A Theory of Task/Technology Fit and Group Support Systems Effectiveness," *MIS Quarterly* (2:3), pp. 313–334.

Zigurs, I., Buckland, B. K., Connolly, J. R., and Wilson, E. V. 1999. "A Test of Task-Technology Fit Theory for Group Support Systems," *ACM SIGMIS Database* (30:3–4), ACM, pp. 34–50.