GEORG-AUGUST-UNIVERSITÄT
GÖTTINGEN

The Consequences of Disregarding Advice in Judge-Advisor Systems

Dissertation

zur Erlangung des mathematisch-naturwissenschaftlichen Doktorgrades

"Doctor rerum naturalium"

der Georg-August-Universität Göttingen

im Promotionsprogramm Biologie

der Georg-August University School of Science (GAUSS)

vorgelegt von

Christian Treffenstädt,

geb. in Aachen

Göttingen, 2017

Betreuungsausschuss

Prof. Dr. Stefan Schulz-Hardt, Abteilung Wirtschafts- und Sozialpsychologie, Georg-Elias-Müller-Institut für Psychologie

Prof. Dr. Michael Waldmann, Abteilung Kognitionswissenschaft und Entscheidungspsychologie, Georg-Elias-Müller-Institut für Psychologie

Mitglieder der Prüfungskommission

Referent: Prof. Dr. Stefan Schulz-Hardt, Abteilung Wirtschafts- und Sozialpsychologie, Georg-Elias-Müller-Institut für Psychologie

Korreferent: Prof. Dr. Michael Waldmann, Abteilung Kognitionswissenschaft und Entscheidungspsychologie, Georg-Elias-Müller-Institut für Psychologie

Weitere Mitglieder der Prüfungskommission:

Prof. Dr. Margarete Boos, Abteilung Sozial- und Kommunikationspsychologie, Georg-Elias-Müller-Institut für Psychologie

Prof. Dr. York Hagmayer, Abteilung Kognitionswissenschaft und Entscheidungspsychologie, Georg-Elias-Müller-Institut für Psychologie

Prof. Dr. Uwe Mattler, Abteilung Kognitionswissenschaften, Georg-Elias-Müller-Institut für Psychologie

Prof. Dr. Lars Penke, Abteilung Biologische Persönlichkeitspsychologie, Georg-Elias-Müller-Institut für Psychologie

Tag der mündlichen Prüfung: 04.05.2017

Danksagung

An meinem Dissertationsprojekt haben in den vergangenen Jahren viele Menschen direkt oder indirekt mitgewirkt, die mir durch ihre Unterstützung und ihre Freundschaft maßgeblich den Weg bereitet haben. Diesen Menschen möchte ich im Folgenden danken.

Mein besonderer Dank gilt meinem Doktorvater Stefan Schulz-Hardt, der mir im Rahmen meiner Promotion immer wieder die Möglichkeit gegeben hat eigene Ansätze zu verfolgen und mir dabei jederzeit zur Seite stand, sei es mit Anregungen und Ratschlägen, oder mit motivierenden und unterstützenden Worten. Die Arbeit mit Ihm hat von Anfang an meine Begeisterung für psychologische Forschung gefördert und mich damit entscheidend für meinen weiteren Werdegang geprägt.

Des Weiteren möchte ich meinem Zweitgutachter Michael Waldmann dafür danken, dass er sich bereit erklärt hat mein Dissertationsprojekt zu begleiten. Bei Margarete Boos, York Hagmeyer, Uwe Mattler und Lars Penke möchte ich mich für die Bereitschaft bedanken, im Prüfungskomitee meiner Dissertation mitzuwirken.

Meinem Kollegen Thomas Schultze-Gerlach danke ich besonders für seine fortwährende Unterstützung in den letzten Jahren. Sowohl fachlich als auch persönlich hat er mich sehr gefördert und mir immer wieder neue und wichtige Denkanstöße gegeben.

Auch bei allen anderen Kolleginnen und Kollegen in der Abteilung für Wirtschafts- und Sozialpsychologie sowie am gesamten Georg-Elias-Müller-Institut möchte ich mich für die Hilfe und Unterstützung in den letzten Jahren bedanken. Ich schätze mich glücklich, viele von ihnen zu meinen Freunden zu zählen.

Ohne die Arbeit und das Engagement meiner studentischen Hilfskräfte so wie der von mir betreuten Bachelor- und Masterstudierenden wäre dieses Dissertationsprojekt nicht realisierbar gewesen. Sie haben nicht nur bei der Durchführung meiner Experimente, sondern auch bei deren Gestaltung unersetzliche Arbeit geleistet.

Meiner Familie und meinen Freunden danke ich, weil sie mich in der Zeit meiner Promotion fortwährend bestärkt und ermutigt haben. Ihr Zuspruch ist mein wichtigster Antrieb.

Johanna, ich hoffe das sagt alles:

Göttingen, im April 2017 *Christian Treffenstädt*

Table of Contents

1. Introduction 1
- 1.1 The Judge-Advisor System 1
- 1.2 The Advisor Perspective 3
- 1.3 Investigating the Rules Underlying the JAS 4

2. Summary of Manuscript 1:
- "You Better Listen to me" - Consequences of Disregarding Advice in Judge-Advisor Systems 7

3. Summary of Manuscript 2:
- Disregarding Advice in Judge-Advisor Systems With Multiple Advisors 9

4. General Discussion 11
- 4.1 Implications 12
- 4.2 Further Research Questions 14

5. Conclusion 17

6. References 17

Appendix A: Manuscript 1 ("You Better Listen to me" - Consequences of Disregarding Advice in Judge-Advisor Systems)

Appendix B: Manuscript 2 (Disregarding Advice in Judge-Advisor Systems With Multiple Advisors)

Appendix C: Curriculum Vitae

1. Introduction

1.1 The Judge-Advisor System

In many real-world situations, we will not try to find a solution by ourselves when facing a judgment or decision problem. Instead, we ask for the opinions and the advice of others to help us (Heath & Gonzalez, 1995). While we fully remain in control of our final judgments and decisions, advice will often have a significant influence on the decision process and its result. To name only a few examples, people commonly receive purchase advice, financial advice, medical advice, or legal advice, and use it to improve their decisions (Harvey & Fischer, 1997). The integration of advice into judgments and decisions is also a common task for employees with management responsibilities in organizations and in staff work (Brehmer & Hagafors, 1986). As Sniezek and Buckley (1995) argue, all of these advice giving and taking situations involve more than one individual, but they do not classify as conventional group judgments or decisions. The diverging roles of decision makers and their advisors do not compare to the undifferentiated roles in groups of equal partners. Since the vast majority of research on group judgment and decision making relies on the latter type of group (e.g, Kerr & Tindale, 2004; Schultze, Mojzisch, & Schulz-Hardt, 2012; Schulz-Hardt & Mojzisch, 2012; Schulz-Hardt, Brodbeck, Mojzisch, Kerschreiter, & Frey, 2006), this line of research offers only limited insight into the process of integrating advice (Bonaccio & Dalal, 2006; Brehmer & Hagafors, 1986; Sniezek & Buckley, 1995). At the same time, research on individual judgment and decision making does not address interactive decision processes and, therefore, does not contribute much to our understanding of advice giving and taking, either (Sniezek & Buckley, 1995).

Feeling that the process of integrating others' opinions into individual judgments and decisions was not yet adequately represented in judgment and decision making research, Sniezek and Buckley (1995) introduced the *Judge-Advisor System* (JAS). The JAS intends to separate social interaction from advice giving and taking in individual judgments and decisions. This way, it enables researchers to analyze the distinct impact of the advice itself on judgment and decision processes. A prototypical JAS consists of a judge, who is responsible for making the judgment and who receives advice in the process of doing so, and at least one advisor. Both parties typically cooperate on several judgment or decision problems (Bonaccio & Dalal, 2006). For each problem, the judge usually forms an uninfluenced initial opinion before receiving

advice. After receiving advice, the judge can adjust his initial opinion and form a final judgment or decision, which is also the final outcome of the JAS interaction. Being responsible for the integration of advice and for the final outcome, judges assume a higher hierarchical position in the JAS, while advisors assume a lower position. The differentiation of roles is a key aspect of real-world advice giving and taking, which, as argued above, is not adequately represented by either individual or group decision making paradigms (Heath & Gonzalez, 1995; Sniezek & Buckley, 1995).

Subsequent to its introduction in 1995, researchers have investigated various aspects of advice giving and taking within the JAS (for an overview, see Bonaccio & Dalal, 2006). Among these aspects, the judge's amount of advice utilization, also called *weight of advice*, is probably the most investigated aspect of a JAS. In the context of a series of decision tasks, the amount of advice taking is usually measured as the proportion of decisions in which a judge's final opinion matched the advice he received (Bonaccio & Dalal, 2006). Since the precision of this measure is strongly dependent on the number of completed decision tasks, it does usually not allow for a finely gradated assessment of advice taking. Therefore, researchers often prefer to work with numeric judgment tasks when investigating advice utilization in the JAS. When using a numeric judgment task in a single-advisor JAS, the amount of advice taking can be computed by the following formula, which will yield a value between 0 (the judge completely ignores the advice, which equals 0% advice taking) and 1 (the judge chooses the advice to be his final judgment, which equals 100% advice taking) on most occasions (for a detailed explanation, see Harvey & Fischer, 1997):

$$\text{advice taking} = \frac{\text{judge final estimate} - \text{judge initial estimate}}{\text{advisor recommendation} - \text{judge initial estimate}}$$

Regarding judges' advice taking, various phenomena were examined using both judgment and decision tasks, for example accuracy gains due to advice taking (e.g., Yaniv, 2004a, 2004b), ego-centric advice discounting (e.g., Yaniv & Kleinberger, 2000; Yaniv & Milyavsky, 2007), and effects of judges' and advisors' perceived expertise on advice taking (e.g., Harvey & Fischer, 1997; Harvey, Harries, & Fischer, 2000; Sniezek, Schrah, & Dalal, 2004). Other aspects that have been investigated in the JAS are, to name only two examples, shifts in judges' confidence

during the decision process (e.g., Budescu & Rantilla, 2000; Budescu, Rantilla, Yu, & Karelitz, 2003; Heath & Gonzalez, 1995), and effects of advisors' trustworthiness (e.g., Bonaccio & Dalal, 2010; Sniezek & Van Swol, 2001; Van Swol, 2009).

1.2 The Advisor Perspective

As argued before, judges in a JAS assume a superior position because of their higher responsibilities in the decision process. Therefore, it might not be a surprise that all of the above-mentioned research questions focus on the judge rather than the advisor. In fact, in many judge-focused JAS studies, the role of the advisor is reduced to merely a series of computer-generated numbers or pieces of advice from previous experiments, which are presented to the judge (e.g., Harvey & Fischer, 1997; Harvey et al., 2000; Yaniv, 2004a; Yaniv & Kleinberger, 2000; Yaniv & Milyavsky, 2007). As a consequence, Sniezek et al. (2004) criticize the apparent imbalance in JAS research:

> Research that is focused solely on the Judge cannot succeed in fully explicating the dynamics of the Judge–Advisor system. If Advisors are not part of the experimental context, they cannot react to or attempt to influence the behavior of the Judge, and vice versa. (Sniezek et al., 2004, p. 187)

This crucial statement serves as the impetus for the present thesis. In fact, the research I will present in the following is developed upon the idea that advisors' reactions to the behaviors exhibited by judges in a JAS could be of critical importance when trying to understand advice giving and taking. As argued above, advisors in most of the previous JAS studies are not actually present in the experiment. Sniezek et al. (2004) state that the advisors in these studies are not able to react to any behavior the judge exhibits; in most of these studies, they are not even aware of the fact that their judgments might serve as advice at some point. The social context of the JAS is cut back to a minimum which does not allow for any social influence apart from the communication of the advice. By including actual participants as advisors, Sniezek et al. (2004) incorporated the social context into the JAS, which they emphasized to be one of the central features of their study. Interestingly, however, even in their study there was no possible way for advisors to react to the behavior of the judges in the JAS. Advisors were asked to give their advice while the judges formed their initial opinions in a separate room. Afterwards, the written

advice on all tasks was transferred to the judges without any further interaction taking place. This procedure can be observed in other JAS experiments using actual participants as advisors, as well (Sniezek & Buckley, 1995; Sniezek & Van Swol, 2001). To the best of my knowledge, there is currently no study (apart from the research presented in this doctoral thesis) that allows advisors in a JAS to react to the behavior exhibited by a judge.

The question is whether such reactions could be relevant in understanding the cooperation of judges and advisors. Looking at the JAS literature, I found interesting suggestions regarding possible answers to this question. For one, Harvey and Fischer (1997) assume that judges are unwilling to completely ignore advice since advisors could react negatively to this behavior. If they are neglected, they could refrain from offering advice in the future. By taking advice into account to at least a specific token amount, judges could ensure advisors' lasting willingness to cooperate. Additionally, Soll and Larrick (2009) state that a social norm might have pushed advisors to take low-quality advice from a socially regarded expert. Since social norms comprise mutual expectations, by following a social norm, a judge conforms to the expectations of the advisor. In contrast, not following a social norm could lead to punishment and to the end of the cooperation within the JAS (c.f. Fehr & Gintis, 2007). Sniezek et al. (2004) make another suggestion. They argue that judges and advisors enter a psychological contract when cooperating in a JAS. A psychological contract is a set of implicit rules which define expectations and obligations in a social interaction (Rousseau, 1989). According to Sniezek et al. (2004), a central component of the psychological contract in a JAS is a social-exchange agreement, defining the mutual benefits and obligations of both judges and advisors. Goldsmith and Fitch (1997) argue that advice is offered as a form of help and social support, for which advisors, in turn, expect their advice to be followed. Like the violation of social norms, not obeying the rules of social-exchange, in other words the terms of the psychological contract, is associated with a broad range of negative reactions from exchange partners (Adams, 1965).

1.3 Investigating the Rules Underlying the JAS

The above-mentioned suggestions unanimously point to the existence of an implicit set of rules that guide the cooperation of judges and advisors in a JAS. Also, in all of these instances, rule violations are, at least in theory, associated with negative consequences and even punishment from cooperators. In essence, if a judge behaves inadequately according to JAS

rules, it seems that repercussions can be expected. At this point, advisors' reactions to judge behavior become an important factor in understanding the social context underlying a JAS. I think that investigating the assumption of the described negative reactions is important for at least three reasons. First, I share the opinion of Sniezek et al. (2004) that advisors need to come more into focus in JAS research, and that they also need to play an active role in JAS studies. By examining the advisor perspective in the JAS, we might be able to gain insights into its underlying dynamics which we could not gain by focusing solely on the judge perspective. For example, we might be able to demonstrate that advisors, in fact, react negatively to the violation of JAS rules in a way that is consistent with reactions commonly observed to accompany contract violations. This would represent evidence for psychological contracts underlying JAS settings, but from a completely new angle. Secondly, while I agree with Sniezek et al. (2004) that varying specific aspects of the psychological contract in a JAS can help us to better understand advice giving and taking processes, this is not the only viable way to gain new insights. By observing advisors' reactions to different sorts of inadequate judge behavior, or to the inadequate behavior of different sorts of judges, we might learn much about the terms and conditions of advice giving and taking. As Rousseau (1989) states, "the workings of psychological contract are perhaps best understood by examining what happens when a psychological contract is violated" (p. 128). Thirdly, when assessing the rules underlying a JAS only from a judge's perspective, one does not account for the possibility that advisors might have different perceptions of these rules. Focusing exclusively on the judge seems to imply that there should be a corresponding set of rules for the advisor, and that both judges and advisors have compatible perceptions of the mutual benefits and obligations in the JAS. However, this is not necessarily true, due to the subjective nature of implicit rules (c.f. Rousseau, 1989). On the contrary, judges and advisors might often have diverging perceptions of a psychological contract in a JAS because, for example, they do not have access to the same information (Sniezek et al., 2004). Therefore, it might be very important to separately assess the rules of a JAS from the perspective of the judges vs. the perspective of the advisors.

Previous findings in judge-advisor research already point to a mismatch of judges' and advisors' subjective perceptions of the rules underlying the social exchange: One of the most robust findings in JAS research is the general tendency of judges to discount advice in favor of their own opinion (Bonaccio & Dalal, 2006). For example, egocentric advice discounting was

observed for both novice and expert advice (although in different degrees), regardless of whether the judge was a novice or an expert himself, and even in spite of judges receiving performance feedback (Harvey & Fischer, 1997; Yaniv, 2004a; Yaniv & Kleinberger, 2000). Also, advice discounting was observed when integrating judgmental opinions with superior statistical forecasts, even when outcome feedback was provided (Lim & O'Connor, 1995). Furthermore, egocentric discounting is not limited to the single-advisor JAS, but also occurs in multiple-advisor settings where judges ignore those pieces of advice that are most distant to their own opinions (Yaniv & Milyavsky, 2007). There are two explanations that could account for the occurrence of such egocentric advice discounting.[1] One is the self/other effect, which assumes that judges prefer their own opinions over an advice because they have better access to the reasons underlying their own opinion than to those of the advice they receive (Yaniv, 2004a, 2004b; Yaniv & Kleinberger, 2000; Yaniv & Milyavsky, 2007). The second explanation assumes that judges are subject to egocentric bias, falsely believing in the superiority of their own opinion (Bonaccio & Dalal, 2006; Harvey & Harries, 2004; J. I. Krueger, 2003).

The findings on egocentric advice discounting show that judges have difficulties to place sufficient weight on advice they receive. At the same time, both Harvey and Fischer (1997) as well as Soll and Larrick (2009) suggest that the implicit rules underlying a JAS contain prescriptions about sufficient amounts of advice taking for different situations. This makes sense because in a classical JAS setting with no social interaction allowed (besides the mere communication of advice), advice taking is the most evident - and often even the only - possibility for a judge to convey gratitude and to obey the psychological contract. Consistent with this, Goldsmith and Fitch (1997) believe that sufficient advice taking is part of the mutual expectations of judges and advisors, which are included in their psychological contract. This

---

[1] In earlier research, an additional explanation for the occurrence of egocentric advice discounting was offered. According to the anchoring-and-adjustment account (Bolger & Wright, 2011; Lim & O'Connor, 1995), judges' own opinions might serve as an anchor in the subsequent decision process. Being anchored on their initial opinion, judges subsequently do not sufficiently adjust towards an advice. However, studies that allow a comparison between the different accounts have found more evidence in favor of the other two explanations (e.g., Clement & Krueger, 2000; Harvey & Harries, 2004).

contradiction of judges' tendency to discount advice and advisors' assumed expectations of sufficient advice taking has additionally motivated the research I present in the following sections.

In summary, three key considerations stand out. First, previous JAS research has neglected the advisor perspective and did not account for advisors' reactions to judge behavior. Secondly, it can be assumed that a set of implicit rules, stemming from mutual expectations and obligations, is underlying the interaction of judges and advisors in a JAS. Thirdly, to learn more about the underlying rules of a JAS, we should observe advisors' reactions to judges neglecting their advice. As a consequence of these considerations, the research in the present doctoral thesis was devised and conducted. The experiments presented in the following sections allowed us to examine the advisor perspective in a JAS. Specifically, by investigating advisors' reactions to low amounts of advice taking, in other words their reactions to judges disregarding their advice, we hoped to learn more about the rules underlying a JAS from a new angle. In my opinion, understanding the rules that guide both judges' and advisors' behavior from their respective perspective is an important step towards fully understanding the dynamics and the social context that are both part of any JAS.

2. Summary of Manuscript 1: "You Better Listen to me" - Consequences of Disregarding Advice in Judge-Advisor Systems

In our first series of experiments, the main goal was to establish and to precisely assess negative consequences following low amounts of advice taking. We observed advisors in a classical single-advisor JAS and confronted them with blatant amounts of either advice utilization or advice discounting. By comparing these extremes, we hoped to create a strong contrast between reciprocity (following the received advice) and a rule violation (disregarding the received advice), to clearly demonstrate the existence of the assumed negative consequences. As a second goal of our experiments, we tested whether perceived expertise differences between judge and advisor would have a moderating effect on the emergence or the intensity of advisors' negative reactions. When trying to establish the terms and conditions underlying the cooperation within a JAS, the variation of perceived expertise of both judge and advisor is a reasonable choice. Not only do we know from previous JAS research that expertise differences actually have an influence on judges' advice taking. Rather, Sniezek and colleagues (2004) also explicitly

name relative expertise to be a concern of the psychological contract. Following their argumentation, relative expertise is a defining component of both contributions and outcomes that are mutually expected in a JAS. This suggests that advice neglect could be justified on a rational basis if an advisor is much less expert than the judge at a given decision problem. In contrast, low amounts of advice taking might not be permissible, particularly when an advisor is much more expert than the judge. In sum, if relative expertise differences can influence the psychological contract and, hence, advisors' expectations regarding the JAS, we would expect them to also have a moderating effect on advisor's negative reactions to their advice being disregarded. Additionally, we wanted to establish our findings on different types of tasks and with different measures, to get an accurate understanding of both the extent and the prevalence of any negative consequences we might observe.

Regarding both goals described above, we observed a very specific and at the same time very consistent pattern of reactions. Following advice neglect, advisors were less motivated to continue working with their current judge and more often desired to switch partners before further cooperating in a JAS. At the same time, advisors' effort in giving advice and their actual advice accuracy did not depend on the amount of advice giving the judge exhibited. In other words, while advisors' willingness to cooperate in the future was diminished if their advice was neglected by the judge, the ongoing cooperation was not damaged at all. As we argue in the manuscript, it stands to reason that this specificity of negative reactions is related to the rules underlying the JAS. Advisors perceived the judge's behavior to be unfair, which is a crucial measure of how satisfied parties to a psychological contract are with the fulfillment of its terms and conditions (Rousseau, 1989). Also, advisors, who were confronted with low amounts of advice taking, recognized the diminished influence their advices had on the final outcome of the JAS. In this situation, motivational theories (c.f., Vroom, 1964) would predict that advisors lower their effort to give good advice. Notwithstanding these circumstances, advisors maintained the same level of cooperation throughout all of the ongoing interaction in the JAS. In our opinion, a set of strong and binding rules, probably included in a psychological contract underlying the JAS, could best explain these findings. Furthermore, both the reported perception of unfairness and the increased desire to switch partners (thereby leaving the cooperation) are known consequences of violating the rules underlying social exchange relationships (Adams, 1965; Rousseau, 1989). Ultimately, since we were able to observe negative consequences

regarding future cooperation, we think Harvey and Fischer (1997) were right to assume that judges might want to avoid ignoring an advisor to maintain his willingness to cooperate.

As for perceived expertise differences, we did not find moderating effects of relative expertise on the negative consequences for future cooperation (and also for the lack of negative consequences in the ongoing cooperation), unless participants were provided with an artificially emphasized feedback on the warranted amount of advice taking. This finding has two important implications. First, while perceived expertise differences moderate judges' advice taking behavior (Harvey & Fischer, 1997), they do not seem to have an effect on advisors' reactions towards the same behavior. When their advice was disregarded, advisors were less motivated to continue working with a judge and more often desired to switch partners even though low amounts of advice taking were justified. This could indicate differences in the way judges and advisors perceive the rules regarding advice taking in a JAS. Therefore, our findings speak to the importance of a separate investigation of the advisor perspective in JAS research. Secondly, the observed absence of moderating effects has an important practical implication. As we argue in the manuscript, judges should not expect advisors to perceive and react to expertise differences the same way they obviously do themselves. Without further communication or explanation to an advisor, expertise differences do not serve as an acceptable reason to adjust advice taking in the way that we simulated in our experiments. Hence, by disregarding advice in line with their own perception of the psychological contract, judges might provoke negative reactions without being aware of any rule violation.

3. Summary of Manuscript 2: Disregarding Advice in Judge-Advisor Systems With Multiple Advisors

After establishing a specific pattern of negative reactions to low amounts of advice taking in our first series of experiments, we assumed that this pattern was produced by a set of implicit rules which guide the interaction of judges and advisors in a JAS. In a next step, we wanted to better understand these rules. We were especially interested in the extent to which the ongoing interaction between a judge and an advisor in a JAS is protected from the negative consequences of insufficient advice taking, that is, the fact that judges did not reduce their effort although they felt being treated unfairly. Basically, we could think of two ways in which rules underlying the JAS could produce the pattern of negative consequences that we observed. On the one hand,

there could be a universal rule (e.g., not to deliberately disappoint others' expectations in an ongoing social exchange) preventing the ongoing cooperation in a JAS from any damage of the participating parties. If this was true, the same pattern of negative consequences should be observable in other JAS settings and under a broad variety of conditions. On the other hand, the ongoing cooperation could have been spared from damage due to a specific characteristic of the single-advisor JAS: When judges disregarded advice in our first series of experiments, they automatically favored their own initial opinions. As we know, judges assume a higher hierarchical position than advisors in a JAS because of their responsibility for information integration as well as for the final outcome. We think that the importance of the judge's role might influence advisors' perceptions regarding the terms of their psychological contract. Specifically, judges' higher importance might, in the eyes of the advisor, "allow" them to place more weight on their own opinions. If this is true, then we should expect more severe negative reactions following low amounts of advice taking in a situation where the judge does not underweight advice in relation to his or her own (more important) opinion, but rather in relation to the (equally important) opinion of another advisor. If, however, we replicate the same negative reactions as in our previous experiments in such a situation, the notion of judges favoring a more important opinion could not serve as a reason for advisors to accept advice neglect in the ongoing cooperation. Rather, we would believe that a universal rule of social exchange protects the ongoing cooperation in a JAS. To test our assumptions, we decided to conduct additional experiments with a different JAS setting. In a multiple-advisor JAS, we observed advisors whose advice was disregarded in favor of a second advisor.

With respect to the question of to what extent the ongoing cooperation in a JAS is protected by its underlying rules, the results of our multiple-advisor JAS experiments indicate a general protection of the ongoing cooperation. Remarkably, we were able to fully replicate the pattern of results from the first manuscript, although the situation we simulated in our multiple-advisor setting differs substantially from our single-advisor setting. As the most important difference, the second advisor assumed the same role our participants did. Besides the manipulation of relative expertise (which had no effect), there was nothing that would have justified a differential treatment of the two advisors. Nonetheless, we did not observe any repercussions of disregarding advice in the ongoing cooperation, even when our participants perceived to be the vastly more competent of the two advisors. At the same time, we did observe

the same negative reactions regarding future cooperation as in our first set of studies. Advisors were less motivated to cooperate, and they more often wished to switch partners for future interactions if their advice was disregarded. Hence, both the absence of negative consequences for the ongoing cooperation as well as the presence of negative consequences for future cooperation extend from the single-advisor JAS to the multiple-advisor JAS. It seems that judges have to be aware of the possible repercussions of disregarding advice in very different JAS settings.

As mentioned before, perceived expertise differences, in accordance with our previous experiments, did not moderate any of the negative reactions (or absence of particular negative reactions) we observed, similar to what we already found in the single-advisor JAS. Neither the emergence nor the extent of negative reactions to advice neglect depended on whether advisors perceived to have superior or inferior expertise compared to the second advisor. In contrast, we know from previous JAS research that judges in a two-advisor JAS use indicators of advice quality (e.g., advisor confidence, Sniezek & Buckley, 1995) to decide which advice to follow. Hence, we can expect that, from a judge's perspective, differences in expertise justify differential weighting in a multiple-advisor JAS. As a consequence, it stands to reason that judge's and advisor's perceptions regarding adequate advice taking may very well differ in a multiple-advisor setting. This result, once again, demonstrates the importance of investigating the advisor perspective in JAS research.

4. General Discussion

The research presented in the two manuscripts summarized above is the first to directly measure advisor behavior in a JAS. By observing advisors' reactions towards low vs. high amounts of advice taking, we were able to demonstrate that disregarding advice can potentially hurt the cooperation between judges and advisors. Six experiments using different types of tasks, JAS settings, and measures provided a clear and consistent picture of the consequences of this advice neglect: Compared to high advice taking conditions, advisors in low advice taking conditions perceived the judge to be more unfair, reported lower motivation to cooperate further, and were more willing to end the cooperation by switching partners. However, neither advisors' effort in advice giving nor the accuracy of their advice was affected by advice neglect. We think that this pattern of reactions is best explained by implicit rules underlying advice taking and

giving in the JAS. In the following sections, I will review the implications of our findings and discuss current limitations as well as possible questions for future research.

4.1 Implications

In our experiments, advisors showed the same pattern of negative reactions to advice neglect in varying situations, indicating a shared set of rules regarding sufficient amounts of advice taking in the JAS. Therefore, our results concur with the assumption that advisors' reactions are guided by a psychological contract which defines these implicit rules and shapes advisors' expectations in the social exchange with a judge (Sniezek et al., 2004). Supporting the assumptions of Harvey and Fischer (1997), advisors' decisions about whether or not they wanted to continue cooperating with a judge depended on the weight that was previously placed on their advice. This leads me to believe that in a decision process which incorporates advisors' reactions to advice taking, meeting advisors' expectations could be a key concern of decision makers if they wish to avoid negative reactions.

Of course, our experiments have put advisors in the particular situation of the JAS which bears several distinct characteristics. First of all, communication within the classical JAS is restricted to the exchange of advice and, in our case, judges' initial and final opinions. Therefore, we currently do not know whether actual judges might have the means to avoid negative reactions to advice neglect in a less restricted scenario. For example, judges might verbally convey gratitude and respect while neglecting advice which could mitigate advisors' negative reactions (Goldsmith & Fitch, 1997). Alternatively, judges could try to explain their reasons to disregard advice to the advisor in question. Secondly, judges and advisors in our experiments were not acquainted with each other, as it is usually the case in JAS research. Advisors might react differently to being neglected from a judge who is personally connected to them. Thirdly, although only under highly artificial circumstances, we have found evidence of possible moderating effects on advisors' negative reactions (Experiment 4 in our first manuscript) that could be more effective in a different setting than the classical JAS. For example, explicit communication about expertise differences or confidence might help advisors to realize the justification of disregarding advice. However, none of these restrictions challenge the principal significance of advisors' expectations in advice giving and taking. They simply point to the

possibility that judges in real-world situations might have other options to avoid the repercussions of disappointing existing expectations.

On a theoretical level, our results suggest that advisors' expectations might be an important source of normative influence that has yet not been accounted for in JAS research. Advisors might not only shape decisions by providing advice, but also by the amount of advice taking they demand in exchange for their help. As argued above, it stands to reason that the risk of damaging the future cooperation with an advisor pushes judges' to place higher weights on advice than they would have in the absence of this risk. However, this type of influence was not represented in previous JAS research. By isolating judge and advisor, advisors' influence on the decision making process was reduced to the informational influence of the advice. While this approach offers the chance to investigate the pure influence of the advice itself, we can only draw limited conclusions regarding real-world advice taking and giving. Therefore, I agree with Sniezek et al. (2004) that the social context, in particular advisors' reactions to judge behavior, has to be included in JAS studies to comprehensively account for normative influence on advice giving and taking.

On a practical level, since our findings point to the importance of advisors' expectations for the future cooperation in JAS, a problem has to be considered, that was previously less salient in the investigation of the JAS:

> Because advice may be an expression of help and caring, advice recipients may feel pressure to follow advice in order to not disrespect the advice giver or appear ungrateful for his or her concern. However, if an advice recipient follows another's advice, he or she may lose autonomy and risk losing the other's respect for his or her competence to act independently. (Goldsmith & Fitch, 1997, p. 468)

The dilemma described by Goldsmith and Fitch points to a fundamental challenge that judges face in the course of advice giving and taking, namely integrating the individual positions of the advisor and the advice taker into a joint outcome that satisfies both sides. In almost all previous JAS studies, judges did not have to fear disappointing the advisor and, therefore, did not face a real challenge with regard to mutual satisfaction (which would allow them to freely ignore advice). If however, judges were directly confronted with an advisor in the JAS interaction, the advisor's expectations and the risk of damaging future cooperation through advice neglect would

become much more evident. As in the described dilemma, judges might feel pressed to place higher weights on advice in such a situation. Confirming this assumption, Schultze, Mojzisch, Herrmann, & Schulz-Hardt (2017) demonstrated that judges show higher amounts of advice taking in the presence of an actual advisor. Our research adds to these findings by showing - from an advisor's perspective - that neglecting advice, indeed, bears the risk of damaging future cooperation in a JAS.

4.2 Limitations and Further Research Questions

We succeeded in demonstrating that disregarding advice can have negative consequences for the cooperation in a JAS. However, our experiments did not allow for a more detailed examination of the relationship between advice neglect and negative consequences. After establishing the existence of negative reactions to advice neglect, it would be a logical next step to conduct a close investigation of their extent as well as their boundary conditions. For example, observing advisors' reactions to other variations of advice taking than those we implemented in our current experiments would allow for a more detailed examination. In this regard, I see two important research questions.

On the one hand we could implement different degrees of constant amounts of advice taking. Varying the weight of advice in several steps (as compared to the two steps in our present experiments) would allow us to infer, for example, the minimum amount of advice taking that advisors perceive to be sufficient in a given JAS setting and task. Also, we could gain insight into the nature of the relationship between the amount of advice taking and the magnitude of negative consequences. In particular, it would be important to find out whether there is a threshold amount of advice taking that judges are not supposed to fall short of (a token amount, c.f., Harvey & Fischer, 1997), or whether the relationship is linear, with varying degrees of negative consequences depending on the extent of the perceived neglect of advice. In the latter case, sufficient advice taking could refer to the lowest amount of advice taking that does not entail any negative consequences.

On the other hand, advisors' reactions to varying amounts of advice taking within a series of judgment or decision tasks could also offer valuable insights into the boundary conditions of negative reactions to advice neglect. Soll and Larrick (2009) argue that in a series of judge-advisor interactions, the average amount of advice taking does not necessarily reflect constant

advice taking for all trials. Instead, judges might adapt their advice taking depending on their perception of the judgment or decision task at hand. On a trial-by-trial basis, they could choose to ignore advice, they could completely follow advice, they could average their own opinion and a received piece of advice, or they could place different weights on their own opinion and the advice. Hence, it stands to reason that the average amount of advice taking represents a more global assessment of a judge's advice taking for a series of interactions, while, actually, judges might choose different weights in each trial. This raises the question of whether advisors' expectations regarding sufficient advice taking relate to a global level of advice taking or to each individual decision. I think this question has major implications for judges in the JAS, since a global expectation of sufficient advice taking would allow judges to ignore specific pieces of advice by placing more weight on the remaining advice. For example, judges might avoid negative reactions by placing more weight on advice they believe to be of high quality while ignoring advice they perceive to be of low quality.

In addition to further varying judges' advice taking as an approach to subsequent research on the consequences of advice neglect, we could also take a closer look at advisors themselves. First, our experiments only examined the consequences of disregarding advice for the cooperation in a JAS. However, disregarding advice might have an additional impact on advisors that we did not yet take into account. In addition to conveying disrespect and ingratitude (Goldsmith & Fitch, 1997), advice neglect might also indicate low advice quality. Perceiving their advices to be neglected because of low quality could have negative effects on advisors' own assessment of their expertise (which we probably did not observe in our experiments because of our strong expertise manipulations) as well as on their self-esteem. Assessing such additional consequences would allow us to achieve a more comprehensive understanding of the effects of disregarding advice.

Secondly, we have suggested that the observed negative reactions to advice neglect might be a result of the failure to meet advisors' expectations. We believe that these expectations part of a broader psychological contract underlying the JAS. So far, however, we only used an indirect approach to infer the terms of this contract. By directly asking advisors about their expectations and their explicit understanding of the rules underlying a JAS interaction, we might be able to get a better idea of the advisor's' expectations's perspective regarding the cooperation in a JAS. However, since a psychological contract is defined as an implicit set of rules, it is not clear how

much of the contract is explicitly perceived by an advisor. Exploring the perceptions of the psychological contract might help us to better understand how such a contract guides advisors' behavior.

Another question regarding the boundary conditions of negative consequences was brought up earlier, as I suggested that advisors might react differently to being neglected by a judge with whom they are personally acquainted. On the one hand, advisors might feel less offended to be neglected by a friend or acquaintance. By reacting less negatively, advisors could avoid damaging their personal relationship with the judge. They might also trust a friend to neglect their advice only for very good reasons, even if those reasons are not disclosed to them. On the other hand, being neglected by a friend might be seen as an even more severe offense than being neglected by someone the advisor does not know. In this case, the negative consequences of advice neglect by an acquainted judge might be aggravated. By comparing acquainted and unacquainted JAS groups, we might determine the effect of acquaintance on advice giving and taking in a JAS.

The perhaps most important future research question is whether our results would replicate in a real interacting JAS. This idea brings up what I consider to be one of the most important limitations to JAS research so far. As I argued before, the inhibition of social interaction in the JAS precludes judges from many options to avoid negative consequences of disregarding advice. By investigating how judges use verbal communication to ensure cooperation within the JAS, we would make an important step towards understanding real-world advice taking and giving. In terms of ecological validity, it is inevitable that we conduct experiments which allow for more interaction and less restricted communication between judge and advisor. While it would be clearly important to determine whether judges use verbal communication to ensure advisors' willingness to cooperate further, looking at the interaction from an advisors' perspective would help us to understand how advisors are affected by such communication. For example it might be interesting to investigate how expressions of respect and gratitude vs. justifications for advice neglect influence advisors' evaluation of the social exchange.

Investigating advice giving and taking in an interacting JAS would also allow us to examine whether and how advisors use verbal communication to ensure the fulfillment of their expectations. Just as judges might verbally convey gratitude and respect or provide reasons for

neglecting advice, advisors might explicitly demand a judge to show a sufficient amount of advice taking, or they could argue for the quality of their advice. Hence, it will be important to investigate both the judge and the advisor perspective in an interacting JAS. In sum, the interactive JAS setting could help us to answer various important questions about real-world advice taking and giving.

5. Conclusion

Research on the JAS offers important insights into advice giving and taking in judgment and decision making processes. While previous JAS research has almost exclusively investigated the judge ' perspective, the present thesis advances this research by shifting the focus onto the advisor. By examining the consequences of disregarding advice in a JAS, new insights into the implicit rules underlying the interaction of judge and advisor were gained. This work presents the first step towards an promising new research direction.

6. References

Adams, J. S. (1965). Inequity in social exchange. *Advances in Experimental Social Psychology*, *2*(267-299).

Bolger, F., & Wright, G. (2011). Improving the delphi process: Lessons from social psychological research. *Technological Forecasting and Social Change*, *78*(9), 1500–1513.

Bonaccio, S., & 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*, 127–151.

Bonaccio, S., & Dalal, R. S. (2010). Evaluating advisors: A policy-capturing study under conditions of complete and missing information. *Journal of Behavioral Decision Making*, *23*(3), 227–249. https://doi.org/10.1002/bdm.649

Brehmer, B., & Hagafors, R. (1986). Use of experts in complex decision making: A paradigm for the study of staff work. *Organizational Behavior and Human Decision Processes*, *38*(2), 181–195. https://doi.org/10.1016/0749-5978(86)90015-4

Budescu, D. V., & Rantilla, A. K. (2000). Confidence in aggregation of expert opinions. *Acta Psychologica*, *104*(3), 371–398.

Budescu, D. V., Rantilla, A. K., Yu, H.-T., & Karelitz, T. M. (2003). The effects of asymmetry among advisors on the aggregation of their opinions. *Organizational Behavior and Human Decision Processes*, *90*(1), 178–194. https://doi.org/10.1016/S0749-5978(02)00516-2

Clement, R. W., & Krueger, J. (2000). The primacy of self-referent information in perceptions of social consensus. *British Journal of Social Psychology*, *39*(2), 279–299.

Fehr, E., & Gintis, H. (2007). Human motivation and social cooperation: Experimental and analytical foundations. *Annual Review of Sociology*, *33*(1), 43–64. https://doi.org/10.1146/annurev.soc.33.040406.131812

Goldsmith, D. J., & Fitch, K. (1997). The normative context of advice as social support. *Human Communication Research*, *23*(4), 454–476. https://doi.org/10.1111/j.1468-2958.1997.tb00406.x

Harvey, N., & Fischer, I. (1997). Taking advice: Accepting help, improving judgment, and sharing responsibility. *Organizational Behavior and Human Decision Processes*, *70*(2), 117–133. https://doi.org/10.1006/obhd.1997.2697

Harvey, N., & Harries, C. (2004). Effects of judges' forecasting on their later combination of forecasts for the same outcomes. *International Journal of Forecasting*, *20*(3), 391–409. https://doi.org/10.1016/j.ijforecast.2003.09.012

Harvey, N., Harries, C., & Fischer, I., I. (2000). Using advice and assessing its quality. *Organizational Behavior and Human Decision Processes*, *81*(2), 252–273. https://doi.org/10.1006/obhd.1999.2874

Heath, C., & Gonzalez, R. (1995). Interaction with others increases decision confidence but not decision quality: Evidence against information collection views of interactive decision making. *Organizational Behavior and Human Decision Processes*, *61*(3), 305–326. https://doi.org/10.1006/obhd.1995.1024

Kerr, N. L., & Tindale, R. S. (2004). Group performance and decision making. *Annual Review of Psychology*, *55*, 623–655. https://doi.org/10.1146/annurev.psych.55.090902.142009

Krueger, J. I. (2003). Return of the ego–self-referent information as a filter for social prediction: Comment on karniol (2003). *Psychological Review*, *110*(3), 585–90; discussion 595–600.

Lim, J. S., & O'Connor, M. (1995). Judgemental adjustment of initial forecasts: Its effectiveness and biases. *Journal of Behavioral Decision Making*, *8*(3), 149–168. https://doi.org/10.1002/bdm.3960080302

Rousseau, D. M. (1989). Psychological and implied contracts in organizations. *Employee Responsibilities and Rights Journal*, *2*(2), 121–139. https://doi.org/10.1007/BF01384942

Schultze, T., Mojzisch, A., & Schulz-Hardt, S. (2012). Why groups perform better than individuals at quantitative judgment tasks: Group-to-individual transfer as an alternative to differential weighting. *Organizational Behavior and Human Decision Processes*, *118*(1), 24–36. https://doi.org/10.1016/j.obhdp.2011.12.006

Schultze, T., Mojzisch, A., Herrmann, S., & Schulz-Hardt, S. (2017). *Advice taking as social exchange: How social interaction between judge and advisor affects advice taking*. Manuscript in preparation.

Schulz-Hardt, S., & Mojzisch, A. (2012). How to achieve synergy in group decision making: Lessons to be learned from the hidden profile paradigm. *European Review of Social Psychology*, *23*(1), 305–343. https://doi.org/10.1080/10463283.2012.744440

Schulz-Hardt, S., Brodbeck, F. C., Mojzisch, A., Kerschreiter, R., & Frey, D. (2006). Group decision making in hidden profile situations: Dissent as a facilitator for decision quality. *Journal of Personality and Social Psychology*, *91*(6), 1080–1093. https://doi.org/10.1037/0022-3514.91.6.1080

Sniezek, J. A., & Buckley, T. (1995). Cueing and cognitive conflict in Judge-Advisor decision making. *Organizational Behavior and Human Decision Processes*, *62*(2), 159–174. https://doi.org/10.1006/obhd.1995.1040

Sniezek, J. A., & Van Swol, L. M. (2001). Trust, confidence, and expertise in a Judge-Advisor system. *Organizational Behavior and Human Decision Processes*, *84*(2), 288–307. https://doi.org/10.1006/obhd.2000.2926

Sniezek, J. A., Schrah, G. E., & Dalal, R. S. (2004). Improving judgement with prepaid expert advice. *Journal of Behavioral Decision Making*, *17*(3), 173–190. https://doi.org/10.1002/bdm.468

Soll, J. B., & Larrick, R. P. (2009). Strategies for revising judgment: How (and how well) people use others' opinions. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, *35*(3), 780–805. https://doi.org/10.1037/a0015145

Van Swol, L. M. (2009). The effects of confidence and advisor motives on advice utilization. *Communication Research*. https://doi.org/10.1177/0093650209346803

Vroom, V. (1964). *Motivation and work*. New York: Wiley.

Yaniv, I. (2004a). Receiving other people's advice: Influence and benefit. *Organizational Behavior and Human Decision Processes*.

Yaniv, I. (2004b). The benefit of additional opinions. *Current Directions in Psychological Science*, *13*(2), 75–78. https://doi.org/10.1111/j.0963-7214.2004.00278.x

Yaniv, I., & Kleinberger, E. (2000). Advice taking in decision making: Egocentric discounting and reputation formation. *Organizational Behavior and Human Decision Processes*, *83*(2), 260–281. https://doi.org/10.1006/obhd.2000.2909

Yaniv, I., & Milyavsky, M. (2007). Using advice from multiple sources to revise and improve judgments. *Organizational Behavior and Human Decision Processes*, *103*(1), 104–120. https://doi.org/10.1016/j.obhdp.2006.05.006

Appendix A

Manuscript 1

"You Better Listen to me" -
Consequences of Disregarding Advice in Judge-Advisor Systems

"You Better Listen to me" - Consequences of Disregarding Advice in Judge-Advisor Systems

Christian Treffenstädt, Thomas Schultze, and Stefan Schulz-Hardt

University of Goettingen

Abstract

Research on advice giving and taking in the judge-advisor system (JAS) has focused on judges' attitudes and behavior. Advisors have not received much attention. Therefore, little is known about whether and how advisors' expectations and perceptions of the interaction with a judge influence decision making in the JAS. The present study tested the assumption that advice neglect can cause negative reactions from advisors which harm their ongoing as well as the future cooperation with a judge. In a series of four experiments, we confronted advisors with either high or low amounts of advice taking. Our findings show that advisors do not take kindly to the disregard of their advice. Advisors in low (as compared to high) advice taking conditions perceived judges to be more unfair, were less motivated to cooperate with a judge in the future, and more often expressed a desire to switch partners, even if disregarding their advice was justified. However, advisors did not lower their effort in the ongoing cooperation. Our results support the assumption of a strong and binding psychological contract to guide advisors' behavior in a JAS.

Keywords: judgment, decision making, advice giving, judge-advisor system

"You Better Listen to me" - Consequences of Disregarding Advice in Judge-Advisor Systems

The judge-advisor system (JAS; Sniezek & Buckley, 1995) allows researchers to explore judgment and decision processes incorporating advice, which is a critical source of social influence in judgment and decision making. In a prototypical JAS - consisting of a judge and one (or more) advisors - judges can form and express an uninfluenced initial opinion regarding a given judgment or decision problem. In a second step, they receive their advisor(s)' opinion(s). Subsequently, judges can revise their initial opinion and form a final judgment or decision. The amount of judges' advice utilization can be measured, for example, by comparing the difference between their initial and final opinions in relation to the received advice (Harvey & Fischer, 1997; Yaniv & Kleinberger, 2000).

Past research using the JAS has focused almost exclusively on the judge in a judge-advisor system. To name a few examples, judges' adjustments of their initial opinions towards advice (e.g., Harvey & Fischer, 1997; Sniezek & Buckley, 1995), accuracy gains due to such revisions (e.g., Soll & Larrick, 2009), or confidence shifts after receiving advice (e.g., Schultze, Rakotoarisoa, & Schulz-Hardt, 2015) were examined. In contrast, the role of the advisor is usually reduced to a mere influencing factor (Bonaccio & Dalal, 2006), and most JAS studies make use of computer-generated or pre-recorded advice. While this approach undoubtedly facilitated the investigation of judges' behavior in the JAS, it had, at least one, clear disadvantage as well: Currently, little is known about the advisor's perspective in the social context underlying a JAS (Sniezek, Schrah, & Dalal, 2004). In our opinion, taking the perspective of an advisor is a crucial step towards a better understanding of advice giving and taking in decision processes. For example, previous findings in JAS research suggest that judges adjust their advice taking due to

advisors' possible reactions (Harvey & Fischer, 1997; Soll & Larrick, 2009). Unfortunately, advisors do not get a chance to react to a judge's behavior, if they play no active part in an experiment (Sniezek et al., 2004). As a consequence, previous JAS studies were not suited to further investigate whether and how advisors' expectations shape advice giving and taking in a JAS.

To take account of this overlooked issue, we decided to conduct a series of studies investigating advisors' reactions to different forms of advice taking behavior. In a series of experiments, we compare advisors' reactions to low amounts of advice taking with their reactions to high amounts of advice taking. By shifting the focus of our investigation onto the advisor, we hope to gain valuable insights regarding the social rules and dynamics underlying a JAS.

Disregarding Advice as a Rule Violation in Social Exchange

We decided to start our investigation of the advisor perspective in a JAS with the most robust finding regarding judges' behavior, namely egocentric advice discounting (Bonaccio & Dalal, 2006). Multiple studies show that judges improperly discount advice they have been offered in favour of their own initial opinion (e.g., Harvey & Fischer, 1997; Harvey & Harries, 2004; Yaniv, 2004; Yaniv & Kleinberger, 2000). Whereas previous research has focused on explanations and moderating factors for this underweighting of advice, the question of how advisors react to such neglect of advice received no attention.

According to social exchange theory (SET; Emerson, 1976), giving advice is regarded as an *act of benefit* in a reciprocal social exchange. This exchange is not regulated through negotiated agreements, but requires the advisor to trust that the person receiving advice will reciprocate accordingly (Molm, Takahashi, & Peterson, 2000). As a consequence, the success of reciprocal exchange relationships is particularly dependent on the premise that all participating

parties comply with reciprocity rules (Cropanzano & Mitchell, 2005). Applying this premise to a JAS, it can be concluded that advisors expect adequate compensation when giving advice to a judge. Since communication in a prototypical JAS setting is limited to the exchange of judgments and decisions, following advice is the only possible way for a judge to compensate an advisor. Therefore, low amounts of advice taking in a JAS should, in fact, constitute a violation of reciprocity rules.

Several findings correspond to the idea of considering advice giving and taking as a reciprocal social exchange and the resulting implications for the JAS. First, Goldsmith and Fitch (1997) argue that advisors offer both informational and emotional support through their advice. In exchange, advice recipients are expected to show gratitude and respect for the advisor's knowledge by following his advice. By disregarding advice, recipients risk showing both disrespect and a lack of gratitude. Secondly, Sniezek and colleagues (2004) suggest that participants in a JAS are bound to a psychological contract that guides their behavior. In particular, this contract includes a social exchange agreement which specifies mutual obligations and the benefits to expect for both judge and advisor. Failure to comply with a psychological contract usually evokes serious negative consequences (Rousseau, 1989). Thirdly, Harvey and Fischer (1997) propose that advisors might perceive disregarding advice as a socially unacceptable rejection. Therefore, judges typically avoid to completely ignore advice and, instead, adhere to a minimum level of advice taking, termed *token amount*. Additionally, Soll and Larrick (2009) support the idea that judges conform to social norms which prohibit disregarding advice under certain circumstances.

In summary, it stands to reason that advisors react negatively to a low amount of advice taking, if it is perceived as a violation of social exchange rules. From the advisor's perspective,

judges should weight their advice sufficiently high to avoid negative consequences. In contrast, however, the earlier-mentioned findings regarding egocentric advice discounting suggest that judges commonly exhibit low amounts of advice taking. We believe that this contradiction of previously observed judge behavior and the assumed consequences of disregarding advice might point to an important but yet unaddressed component of advice giving and taking. According to Goldsmith and Fitch (1997), advisors potentially face a dilemma when deciding about whether and how to follow advice. On the one hand, they strive for autonomy and want to make their own decisions but, on the other hand, they want to retain a good relationship with the advisor. By investigating the advisor's perspective, we particularly hope to gain valuable information regarding this conflict of interest.

Social exchange theory predicts negative behavioral as well as motivational consequences following assumed rule violations. Adams (1965) argues that failure to reciprocate in a social exchange will result in dissatisfaction, perceptions of unfairness, a lowered willingness to cooperate, and ultimately the urge to quit the exchange relationship for the wronged party. Additionally, by lowering their effort, disadvantaged participants might try to reestablish equity in the exchange situation. We propose that these consequences also apply to rule violations in a JAS:

Hypothesis 1: Compared to high amounts of advice taking, low amounts of advice taking in a JAS will lead advisors to react negatively. Possible negative reactions include perceptions of unfairness, lowering the effort in advice giving, and disengaging from future cooperation.[1]

The Role of Expertise Differences

While we have argued that advisors could react negatively if their advice is not sufficiently taken into account, this reaction need not always be the case. Instead, advisors might feel that, under certain circumstances, low amounts of advice taking are justified and, thus, do not constitute a rule violation. One factor that might moderate advisors' negative reactions to low amounts of advice taking is the relative expertise of judge and advisor. To a (rational) judge, the value of advice primarily depends on whether advice taking will increase decision quality or not. The less competent an advisor is in relation to the judge, the less benefit can be expected from following his or her advice. Furthermore, when judges are considerably more competent than their advisors, normative models aiming to maximize judgmental accuracy explicitly require judges to either show low amounts of advice taking (Bednarik & Schultze, 2015), or to ignore advice altogether (Soll & Larrick, 2009). There is ample evidence that judges understand the

---

[1] Advisors could also impose informal sanctions, like putting pressure on the judge or expressing disapproval and dissatisfaction regarding the cooperation. However, for the present series of experiments investigating the advisor perspective, we decided to rely on a classical JAS experimental setting with no communication between judges and advisors besides exchanging mere decisions or estimates. This allows us to precisely determine and control the judge-advisor interaction.

implications of expertise differences and adjust their advice utilization accordingly (e.g., Harvey & Fischer, 1997; Soll & Larrick, 2009; Soll & Mannes, 2011; Yaniv & Kleinberger, 2000). Hence, it is plausible to assume that advisors understand these implications as well. If so, they might react much less negatively to low amounts of advice taking when the judge is more competent at the task and their advice is, evidently, of limited use.

Hypothesis 2: Perceived expertise differences moderate the extent of negative reactions to low amounts of advice taking. The negative reactions to disregarding advice are ameliorated when the judge is more competent relative to the advisor.

General Method

Experimental Setting and Design

In the following, we present a series of four experiments, which were conducted in order to test the assumption of advisors reacting negatively to low amounts of advice taking, and to examine how perceived expertise differences might interact with such negative reactions. In each experiment, participants took the role of an advisor in a computer-mediated judge-advisor dyad. We led them to believe that they would cooperate with another participant sitting in an adjacent

room while, in fact, they interacted with a computer simulated partner.[2] This simulated judge showed either high or low amounts of advice taking. This way, we were able to precisely manipulate the judge's advice taking and, as a consequence, to ensure a strong and consistent advice taking manipulation for all participants. To manipulate perceived relative expertise differences, we provided our participants with bogus feedback which either suggested an expertise difference in favor of the judge or an expertise difference in favor of the advisor. Hence, our experiments are based on a 2 (amount of advice taking: high vs. low) x 2 (relative expertise: judge more competent vs. advisor more competent) between-subjects design.

Procedure

The experimental procedure was essentially the same in all four experiments. We invited participants in groups of four to five to ensure the credibility of the alleged formation of judge-advisor dyads. If an odd number of participants showed up to an experimental session, the experimenter informed them that one participant would be randomly selected to work on an alternative experiment. After giving participants some time to be acquainted, we led them into

[2] We usually avoid deception whenever it is possible. However, for the purposes of the present experiments, we did not find a feasible alternative setup for the experiments that would work without deception. A test of our hypotheses requires stable amounts of advice taking and constant expertise differences between judges and advisors to be orthogonally manipulated. This means judges need to show either very high or very low amounts of advice taking independent of the expertise differences in the JAS. This could hardly ever be achieved with real interacting JAS dyads.

separate rooms with prepared work places, and they started to work on the computer experiment. In the instruction phase, participants were presented with a bogus randomization procedure to determine their role in the upcoming judge advisor system. At this point, all participants were informed that they had been assigned the role of an advisor in the JAS, while their alleged partner was chosen to take on the role of a judge. After the instruction phase, participants completed a training phase to familiarize themselves with the task they had to advise on in the following JAS interaction. In the experimental phase, participants started each trial by acquiring cues that allowed them to offer accurate advice. The amount of acquired cues in each trial was determined by participants' performance in an effort task. Since a higher amount of cues supposedly led to more accurate advice, this task allowed us to assess participants' effort in giving advice. Following cue acquisition, participants gave advice which allegedly was presented to the judge. After a brief waiting period in which the judge appeared to contemplate the advice, participants received feedback which allowed them to infer the amount of advice taking the judge had exhibited in the trial (by showing participants their own advice as well as the jugde's initial and final judgments). After a series of 20 trials, the experimental phase ended. Subsequently, participants were asked to answer several additional questions regarding the interaction with the judge. Finally, participants were fully debriefed and were paid a fixed fee of €10, before leaving the laboratory.

Dependent Variables

Manipulation checks. To verify the success of our advice taking manipulation, participants were asked to indicate to what extent (ranging from 0 to 100 percent) their partner had weighted their advice during the JAS interaction. Additionally, participants were asked to

rate how much the judge's final judgments depended on their advice on a seven-point Likert scale (1 - *not at all*, 7 - *very much*).

As a manipulation control for our performance feedback manipulation, we asked participants to recall the feedback they had received after the training phase, and to indicate whether they had outperformed their partner or had been outperformed by their partner. Furthermore, participants were asked to rate their own performance as well as their partner's performance on a seven-point Likert scale (1 - *very bad*, 7 - *very good*).

Effort in advice giving. Participants' performance in the effort task served as our primary indicator of the amount of effort they invested into giving accurate advice. As a secondary measure of effort in advice giving, participants were asked to rate how much effort they had invested in cue acquisition on a seven-point Likert scale (1 - *no effort at all*, 7 - *very much effort*) in the final questionnaire.

Advice quality. In Experiments 1, 3, and 4, we determined advice quality as the mean absolute percentage error (MAPE) or the mean absolute error (MAE) of participants' advice in the main experimental phase. Small MAPE values indicate high accuracy and a high quality of advice, while higher MAPE values indicate low accuracy and advice quality. For the decision task in Experiment 2, advice quality was measured as the number of correct pieces of advice which advisors offered the test phase. Since participants could intentionally give inaccurate advice as a reaction to the experimental manipulation, it is important to measure advice quality separately from our measures of effort.

Willingness to cooperate further. After completing 20 rounds of JAS interactions, participants were given the chance to express a wish to either keep or switch their current partner before allegedly continuing to work on further judgment or decision tasks. Participants' desire to

switch partners served as our primary measure for their willingness to cooperate further with the judge. As a second measure, we asked participants to rate how motivated they felt to cooperate further with their current partner on a seven-point Likert scale (1 - *not motivated at all*, 7 - *very motivated*). Afterwards, we informed our participants that they had already finished the test phase, and that there would be no further JAS interactions.

Perceived fairness. Participants were asked to rate the fairness of their partner's behavior on a seven-point Likert scale (1 - *not fair at all*, 7 - *very fair*). Fairness ratings are an important indicator of participants' satisfaction with the fulfillment of a psychological contract (1989). Hence, high fairness ratings in our experiments mean that a judges' behavior matched participants' expectations, while low fairness ratings mean participants expected a different treatment in the social exchange.

Additional Variables. We included additional measures in our experiments for explorative purposes only (a complete list of all additional measures is included in the Appendix). While we will not report analyses of these measures in the present manuscript, the full datasets as well as all materials can be requested at any time from the first author.

Experimental Objectives

In Experiment 1, we assessed negative consequences following low amounts of advice taking in a multiple-cue temperature judgment task. These temperature judgments were combined with a physical effort task for cue acquisition. In our second experiment, we wanted to replicate the findings of Experiment 1 with a decision task instead of a judgment task. To account for the possibility that our findings regarding participants' effort in advice giving were specific to our physical effort task, we conducted Experiment 3 with a dexterity task instead. Also, we used a second multiple-cue judgment task to test whether our pattern of results would

replicate for tasks with lower difficulty. In Experiment 4, we presented participants with direct feedback regarding optimal amounts of advice taking, which allowed for a test of possible explanations for our findings regarding Hypothesis 2.

The descriptions of the four experiments included in the following sections provide additional information on the different judgment and decision tasks as well as the effort tasks we implemented. We also describe any occurring differences between experimental procedures.

Experiment 1

Method

Participants and design. Participants were recruited using the Online Recruitment System for Economic Experiments (ORSEE; Greiner, 2015). We excluded 17 out of the 99 original participants because technical errors prevented them from completing the experiment. The remaining 82 participants were, on average, 22.45 (SD = 4.48) years old. 52 participants (63%) were female. One participant did not provide a response regarding gender.

Task and apparatus. In Experiment 1, participants worked on a numerical multiple-cue judgment task, namely estimating the daily mean temperatures of unknown cities anywhere in the world in degrees Celsius. They received up to five cues from which they could infer the target values. These cues were the respective city's altitude above sea level, position in longitude, position in latitude, monthly mean humidity, and the monthly precipitation.

Through a pilot study conducted online, we tested whether participants understood that knowledge of more cues would imply more accurate temperature judgments. In this pilot study, 383 participants worked on a series of the described temperature judgment tasks, with the amount of available cues being varied between participants (0 vs. 1 vs. 2 vs. 3 vs. 4 vs. 5). As an

indicator of the perceived quality of their own judgments, participants' mean ratings of confidence were compared between groups in a one-way ANOVA. The analysis revealed a significant effect of cue amount on participants' confidence in the accuracy of their own judgments and, therefore, their perceived judgment quality, $F(5, 377) = 3.60$, $p = .003$, $\eta_p^2 = .05$, with higher cue amounts leading to higher confidence. However, we did not find a significant relation between the amount of available cues and participants' actual judgment accuracy, as measured by mean absolute percentage error (MAPE) values, $r(381) = -.01$, $p = .909$. Hence, while our participants understood that judgment accuracy depended on the number of cues, they were not able to generate more accurate judgments when they had more cues available to them. This was, however, of no consequence for our experiments since only participants' perception of the cue-accuracy relationship was an important antecedent for our experimental approach.

In the training phase at the beginning of the experiment, all five cues were available to the participants. In the test phase, however, participants had to acquire cues by pressing a hand muscle trainer (HMT) for as long as possible. Depending on the participant's gender, the hand muscle trainer's pressure resistance was adjusted to account for differences in hand muscle strength. We modified HMTs to allow exact measurement of pressing durations and their immediate evaluation within our computer experiment. Depending on pressing durations, participants received zero to five of the cues randomly selected from the list described above.

HMT pressing durations served as a measure of effort when obtaining information about the unknown cities. Physical effort tasks comparable to HMT pressing are common measures of task-related motivation (e.g., Hertel, Kerr, & Messé, 2000; Kerr, 1983), and a second pilot study confirms the validity of measuring this motivation through HMT pressing durations. In this second pilot study, we compared pressing durations of participants with a performance-linked

participation fee to those of participants with a fixed participation fee. A *t*-test showed that pressing durations were higher for the performance-linked fee group (M = 85.61, SD = 81.31) than for the fixed fee group (M = 53.82, SD = 38.53), $t(63) = 2.00$, $p = .049$, $d = 0.50$. Hence, the effort exerted when pressing the HMT seems to reflect the differences in task-related motivation induced by a monetary incentive.

Procedure. In the instruction phase at the beginning of the computer experiment, we explained participants how to use the HMT. Subsequently, a baseline for hand muscle strength was determined through two HMT pressing trials, where participants learned they would receive a bonus of €1 for good performance. After these initial HMT measurements, participants were informed about the judgment task, and the five types of cues to the daily mean temperature of a given city were explained. The direction of the relationship between each type of cue and the mean temperature was outlined. For example, participants learned that a higher monthly precipitation is usually associated with higher mean temperatures. Afterwards, participants completed the training phase.

The training phase was followed by a bogus performance feedback through which we manipulated relative expertise in the JAS. The feedback was presented as a ranking of both judge and advisor in relation to 100 participants who had allegedly completed the same temperature estimation tasks in an earlier study. Participants in conditions with more competent advisors were told that they had ranked 24^{th} while the judge had only ranked 85^{th}. Hence, they had outperformed the judge by far. In conditions with more competent judges, these rankings were reversed.

After receiving the bogus feedback, participants were informed about the acquisition of cues by pressing the HMT in the following temperature judgment trials. They were told that they

would unlock more cues the longer they pressed the HMT. HMT pressing durations were used as a physical indicator of effort exerted for the advice giving. Participants were also told that their participation fees would not depend on the quality of their advice. Hence, there was no financial disadvantage in exerting less effort in advice giving.

In the test phase, participants started each trial by completing a HMT pressing phase, which was followed by the judgment task. The pressing duration determined the number of available cues in the respective trial. The time required to achieve a certain number of cues was tailored to participants' individual capabilities measured via their baseline HMT performance. Thresholds for the different number of available cues were 35% (one cue), 55% (two cues), 80% (three cues), 110% (four cues) and 150% (five cues) of the baseline performance. After participants stopped pressing the HMT, they learned how many cues (and which ones) they would receive. Obviously, we chose to set very high hurdles for participants to unlock particularly the fourth and the fifth cue, to prevent any ceiling effects in cue acquisition. Based on the presented cues, participants were asked to give their best estimate, which their partner would then receive as advice. As described above, we made sure that participants understood the link between pressing durations, the amount of unlocked cues, and the potential accuracy of their advice. Differences in advice giving effort should, therefore, be observable through differences in pressing durations.

Advice giving was followed by the simulated advice taking phase. At the end of each trial, participants were given a final overview on the judge's initial judgment, their advice, and the final judgment. Based on this information, they could infer how the judge had weighted their advice. in high advice taking conditions, simulated judges weighted advice by 60 to 80 percent, with an average of about 70 percent. in low advice taking conditions, the weight placed on the

advice varied between 0 and 20 percent, with an average of 10 percent. These levels of advice taking correspond to the upper and lower end of advice taking observed in previous research (Harvey & Fischer, 1997; Yaniv & Kleinberger, 2000).

Results

Detailed results for all analyses of variance and covariance which are reported in this section can be found in Table 1.

Manipulation checks. We first tested whether the manipulation of our simulated judge's advice taking was successful. Indeed, participants in high advice taking conditions perceived their partner to place much higher weight on their advice than participants in low advice taking conditions (M = 79.71, SD = 11.8 vs. M = 11.59, SD = 8.53), $t(80) = 29.96$, $p < .001$, $d = 6.62$. As we can see, participants' average assessment of advice taking was very accurate (actual simulated advice taking values were 70% vs. 10%).

The analysis of participants' ratings of how much the final judgments depended on their advice revealed a significant effect of advice taking, $F(1, 78) = 188.97$, $p < .001$, $\eta_p^2 = .71$. Advisors in low advice taking conditions perceived less influence (M = 2.39, SD = 1.00) on the final judgments than advisors in high advice taking conditions (M = 5.32, SD = 0.91). There were no main or interactive effects of relative expertise, both $Fs < 1$, $ps \geq .738$, $\eta_p^2 s \leq .01$.

Regarding the manipulation of relative expertise, 87 percent of our participants correctly recalled the ranking difference they were presented with in the bogus feedback. Participants recalled the correct bogus feedback vastly above chance level, $\chi^2(1, N = 82) = 44.57$, $p < .001$.

We also calculated the difference between our participants' performance ratings for themselves and for their respective partner. Positive scores indicate a performance difference in favor of the partner, and negative scores indicate a performance difference in favor of the

participant. A *t*-test revealed that participants who received favorable performance feedback (advisor ranked better than judge) perceived a significantly smaller performance difference in favor of the judge (M = 0.27, SD = 1.48) than did participants who received unfavorable performance feedback (M = 1.85, SD = 1.44), $t(80) = 4.91$, $p < .001$, $d = 1.08$.[3]

Effort in advice giving. We analyzed HMT pressing durations in a 2 (amount of advice taking: high vs. low) x 2 (relative expertise: judge more competent vs. advisor more competent) ANCOVA with the mean of both baseline measurements as a covariate to control for interpersonal differences in hand muscle strength. Participants' effort ratings were analyzed in a comparable 2 x 2 ANOVA. In both analyses, we did not find any significant effects of advice taking, relative expertise, or their interaction, all $Fs < 1$, $ps \geq .647$ $\eta_p^2 s \leq .01$.

Advice quality. Using a 2 x 2 ANOVA of advice quality (MAPE scores), we found no effects for the amount advice taking, the relative expertise condition, or their interaction, all $Fs < 1$, $ps \geq .658$ $\eta_p^2 s \leq .01$.

Willingness to cooperate further. To test for differences in participants' willingness to cooperate further with the judge, we performed a binary logistic regression analysis with desire to switch partners as criterion and advice taking, relative expertise, and their interaction as predictors. Our results showed a significant effect of advice taking, B = 1.22 (SE = 0.47), OR = 3.37, $OR\ 95\%\text{-}CI$ = [1.46, 10.60], p = .010. Disregarding advice led to a higher rate of

---

[3] Interestingly, we did not find a perceived performance difference in favor of our participants even after receiving favorable performance feedback. This might be ascribed to high task difficulty. Nonetheless, participants did perceive performance differences differently depending on relative expertise condition, which speaks to the success of our expertise manipulation.

participants wishing to switch partners (39% in low advice taking conditions vs. 12% in high advice taking conditions). There were no significant effects of relative expertise, *B* = -0.83. (*SE* = 0.47), *OR* = 0.43, *OR 95%-CI* = [0.14, 1.02], *p* = .078, or the interaction term, *B* = -0.43. (*SE* = 0.67), *OR* = 0.65, *OR 95%-CI* = [0.13, 2.20], *p* = .520.

We found similar results for participants' ratings regarding their motivation to collaborate further with their current partner. A significant effect of advice taking shows that participants in low advice taking conditions are less motivated to continue working with their current partner (*M* = 2.51, *SD* = 1.34) than participants in high advice taking conditions (*M* = 3.56, *SD* = 1.69), *F*(1, 78) = 9.87, *p* = .002, η_p^2 = 0.11. There was neither an effect of relative expertise, nor an interaction, both *F*s < 1, $ps \geq .355$ $\eta_p^2 s \leq .01$.

Perceived fairness. Participants' fairness ratings regarding their partner's advice taking showed a significant effect of advice taking, *F*(1, 78) = 103.49, *p* < .001, η_p^2 = .57. Participants generally rated their respective partner's behavior as being more fair in high advice taking conditions (*M* = 6.29, *SD* = 1.10) as compared to low advice taking conditions (*M* = 3.41, *SD* = 1.55). Relative expertise did not have a significant effect, *F*(1, 78) = 2.20, *p* = .142, η_p^2 = .03, but there was an interaction of advice taking and relative expertise, *F*(1, 78) = 8.51, *p* = .005, η_p^2 = .04.

An analysis of the simple effects within the two advice taking conditions showed significant differences only in low advice taking conditions (judge more competent *M* = 4.05, *SD* = 1.54 vs. advisor more competent *M* = 2.81, SD = 1.33), *t*(39) = 2.77, *p* = .009, *d* = 0.86, but not within high advice taking conditions (judge more competent *M* = 6.10, *SD* = 1.41 vs. advisor more competent *M* = 6.50, SD = 0.61), *t*(39) = 1.18, *p* = .244. In accordance with our suggestions regarding perceived expertise differences, when advice taking was low, participants perceived a

significant difference in situations with a more competent judge than in situations where they were more competent themselves. However, when advice taking was high, participants rated their partner's behavior to be very fair irrespective of any expertise differences.

Table 1

Summary of computed ANOVAs in Experiment 1

Source	*df*	*F*	*p*	η_p^2
DV 1: HMT Pressing Durations				
Baseline Duration	1	41.98	< .001	.35
Advice Taking (AT)	1	< 1	.780	.00
Feedback (FB)	1	< 1	.922	.00
AT x FB	1	< 1	.850	.00
Residuals	77			
DV 2: Effort Ratings				
Advice Taking (AT)	1	< 1	.647	.00
Feedback (FB)	1	< 1	.939	.00
AT x FB	1	< 1	.964	.00
Residuals	78			
DV 3: Advice Quality				
Advice Taking (AT)	1	< 1	.892	.00
Feedback (FB)	1	< 1	.950	.00
AT x FB	1	< 1	.658	.00
Residuals	78			
DV 4: Motivation Ratings				
Advice Taking (AT)	1	9.87	.002	.11
Feedback (FB)	1	< 1	.355	.01
AT x FB	1	< 1	.672	.00
Residuals	78			
DV 5: Fairness Ratings				
Advice Taking (AT)	1	103.49	< .001	.57
Feedback (FB)	1	2.20	.142	.03
AT x FB	1	8.52	.005	.10
Residuals	78			
DV 6: Dependency Ratings				
Advice Taking (AT)	1	188.97	< .001	.71
Feedback (FB)	1	< 1	.738	.00
AT x FB	1	< 1	.973	.00
Residuals	78			

Discussion

The results of Experiment 1 provide first evidence that, as we predicted, there are negative consequences of low amounts of advice taking in a JAS. However, contrary to our expectations, negative consequences occurred only on particular measures. The willingness to cooperate in the future differed between our two advice taking conditions, observable both in the willingness to switch partners, and in motivation to cooperate in the future. Also, participants' fairness ratings differed significantly depending on advice taking condition. In contrast, there were no effects of advice taking on participants' effort in advice giving, or on advice quality. Hence, it seems that advisors' negative reactions were restricted to a future cooperation, whereas the ongoing interaction was not affected by the experience of one's advice being neglected. If these results prove to be robust, this would raise the question which forces prevent advisors to retaliate against low amounts of advice taking in the ongoing cooperation.

Another important observation in Experiment 1 is the lack of influence of relative expertise on the observed outcomes (the only instance where participants seemed to consider relative expertise differences at all was the rating of the judge's fairness within low advice taking conditions). In other words, neither the negative effects of advice neglect on advisors' willingness to cooperate in the future, nor the absence of such effects on their effort in the ongoing interaction depended, by any means, on whether or not the judge had good reasons to do so, namely because the advisor allegedly lacked competence in comparison with the judge. Once again, if this finding proves to be robust, it raises interesting questions with regard to why advisors seem to ignore this fact.

A limitation of our first experiment could be that the average number of cues that the participants acquired was relatively low (M = 1.56, SD = 1.05). It seems that our baseline hand

muscle strength measurements produced a suboptimal standard of comparison which resulted in relatively low amounts of cues being unlocked. Since it is unclear whether the exceeding difficulty in unlocking cues had any negative effects on participants' motivation or effort in the HMT task (or the measurement of these variables), we decided to replicate our findings with a more suitable baseline in Experiment 2.

Experiment 2

After having used a judgment task in Experiment 1, we wanted to test whether our results would replicate in a decision task, since judgment and decision tasks vary substantially in the way that advisors perceive advice taking. In a judgment task, advice taking can occur gradually by placing different weights on the advice and a judges' initial estimates. In a decision task, however, taking advice means to choose the advisors' proposed option. Depending on the type of task, advisors might have different expectations regarding sufficient advice taking. Hence, we think it is important to examine the assumed negative reactions both in a judgment and in a decision task context.

Experiment 2 differs from Experiment 1 in two aspects. Instead of estimating the daily mean temperature of a specific city, participants had to choose the city with the highest daily mean temperature in a set of four presented cities. Decisions had to be based on the same set of cues that was used in Experiment 1, with each cue being presented for all four cities in the set. As a second difference, HMT baseline measurements were conducted without the possibility of acquiring an additional incentive for high performance, in order to produce a more adequate baseline for cue acquisition. Otherwise, the procedure in Experiment 2 is exactly the same as in Experiment 1.

Method

Participants. Participants were recruited using ORSEE. 33 out of 159 participants had to be excluded for not completing the experiment due to technical errors concerning the HMT. Of the remaining 126 participants, 78 were female. Participants had an average age of 23.52 (SD = 4.26).

Results

Detailed results for all analyses of variance and covariance which are reported in this section can be found in Table 2.

Manipulation checks. All manipulation checks produced comparable results to the manipulation checks in Experiment 1. Participants in low advice taking conditions perceived a lower amount of advice taking than did participants in high advice taking conditions (M = 19.85, SD = 18.65 vs. M = 79.58, SD = 19.65), $t(124) = 17.49$, $p < .001$, $d = 3.12$.

Also, a significant effect of advice taking on dependency ratings was found, $F(1, 122) = 110.78$, $p < .001$, $\eta_p^2 = 0.48$. Participants in low advice taking conditions rated final decisions to depend less on their advice (M = 2.71, SD = 1.40) than participants in high advice taking conditions (M = 5.30, SD = 1.36).

Eighty-Seven percent of our participants correctly remembered whether or not they achieved a higher rank than their partner in the performance feedback following the training phase, which was clearly above chance levels, $\chi^2(1, N = 126) = 70.35$, $p < .001$.

Providing even stronger evidence for the success of our relative expertise manipulation than in Experiment 1, participants in conditions with more competent advisors perceived a performance difference in favor of their advisor (M = -0.41, SD = 1.83), while participants in conditions with more competent judges perceived a difference in favor of the judge (M = 0.87,

SD = 1.81), $t(124) = 3.94$, $p < .001$, $d = 0.70$. Thus, all our experimental manipulations were successful.

Effort in advice giving. Analyses of both HMT pressing durations and subjective effort ratings showed no significant effects of advice taking, relative expertise, or their interaction, all $Fs \leq 2.13$, $ps \geq .147$, $\eta_p^2 s \leq .02$.

Advice quality. The ANOVA for advice quality showed no significant effects of advice taking, relative expertise, or their interaction, all $Fs \leq 2.87$, $ps \geq .093$, $\eta_p^2 s \leq .02$.

Willingness to cooperate further. As in Experiment 1, a binary logistic regression showed a significant effect of advice taking, $B = 0.84$, ($SE = 0.28$), $OR = 2.32$, OR 95%-CI = [1.36, 4.04], $p = .005$. Low amounts of advice taking led to a higher rate of participants wishing to switch partners (50% in low advice taking conditions vs. 23% in high advice taking conditions). There were no effects of relative expertise, $B = -0.29$, ($SE = 0.28$), $OR = 0.75$, OR 95%-CI = [0.43, 1.30], $p = .300$, or the interaction of advice taking and relative expertise, $B = -0.02$, ($SE = 0.39$), $OR = 0.98$, OR 95%-CI = [0.45, 2.13], $p = .968$.

We also conducted an ANOVA that showed no effects of advice taking, relative expertise, or their interaction, on advisors' rated motivation to continue the cooperation with their current partner, all $Fs \leq 2.24$, $ps \geq .137$, $\eta_p^2 s \leq .02$. On a descriptive level, however, the differences in motivation ratings between high and low advice taking conditions matched our predictions (high advice taking $M = 3.75$, $SD = 1.89$ vs. low advice taking $M = 3.26$, $SD = 1.78$).

Perceived fairness. The analysis of fairness ratings showed a significant effect of advice taking, $F(1, 122) = 68.88$, $p < .001$, $\eta_p^2 = 0.36$. Advice taking behavior was rated as being generally more fair in high advice taking conditions ($M = 5.56$, $SD = 1.61$) than in low advice

taking conditions (M = 3.19, SD = 1.56). Neither the effect of relative expertise nor the interaction were significant, both $Fs < 1$, $ps \geq .582$, $\eta_p^2 s < .01$.

Table 2

Summary of computed ANOVAs in Experiment 2

Source	*df*	*F*	*p*	η_p^2
DV 1: HMT Pressing Durations				
Baseline Duration	1	106.88	< .001	.47
Advice Taking (AT)	1	< 1	.882	.00
Feedback (FB)	1	< 1	.530	.00
AT x FB	1	1.44	.233	.01
Residuals	121			
DV 2: Effort Ratings				
Advice Taking (AT)	1	2.13	.147	.02
Feedback (FB)	1	< 1	.567	.00
AT x FB	1	< 1	.688	.00
Residuals	122			
DV 3: Advice Quality				
Advice Taking (AT)	1	2.87	.093	.02
Feedback (FB)	1	1.70	.195	.01
AT x FB	1	< 1	.543	.00
Residuals	122			
DV 4: Motivation Ratings				
Advice Taking (AT)	1	2.24	.137	.02
Feedback (FB)	1	< 1	.514	.00
AT x FB	1	< 1	.972	.00
Residuals	122			
DV 5: Fairness Ratings				
Advice Taking (AT)	1	68.88	< .001	.36
Feedback (FB)	1	< 1	.582	.00
AT x FB	1	< 1	.967	.00
Residuals	122			
DV 6: Dependency Ratings				
Advice Taking (AT)	1	110.78	< .001	.48
Feedback (FB)	1	< 1	.418	.01
AT x FB	1	< 1	.364	.01
Residuals	122			

Discussion

The results of Experiment 2 largely resemble those of Experiment 1. Once more, participants whose advice was almost ignored felt treated more unfairly than participants whose advice was highly weighted by the judge. However, these feelings of fairness (or lack thereof) were independent of whether or not the alleged performance differences implied that the judge should neglect the advice – even if the participants' advice was rightfully neglected, they felt that this was unfair. We observed similar negative reactions for reduced willingness to cooperate further, and for lower ratings of perceived fairness when advice taking was low. Once again, these effects occurred independent of whether or not the amount of advice taking was justified by the alleged performance differences between judge and advisor. Measures of current effort, however, were again unaffected by advice neglect. Since we succeeded in obtaining a more suitable baseline for HMT performance in Experiment 2 (average number of acquired cues in Experiment 2: $M = 2.75$, $SD = 1.21$), the absence of effects for current effort cannot be explained by low motivation or a lack of sensitivity of this measure due to high difficulty of unlocking cues.

Regarding the general pattern of results, there was one minor difference between our first two experiments. There was a significant effect of advice taking on motivation ratings in Experiment 1, but not in Experiment 2. Due to this inconsistency, we abstain from further interpretation before having obtained more evidence on this effect in the subsequent experiments.

So far, our experiments leave us with two questions. The first is how to explain that low weights of advice do not reduce advisor's ongoing effort in a JAS. We can think of both theoretical and methodical reasons that might have prevented negative effects in the ongoing

cooperation. The second question concerns the lack of a moderating effect of relative expertise on the relationships between advice taking and our outcome measures. To further narrow down possible explanations for the observed pattern of results in our experiments so far, we decided to conduct two additional studies.

Experiment 3

Regarding the lack of advice taking effects on the ongoing cooperation in a JAS, there is the possibility that our previously used effort measures might not be sufficiently suited to detect existing but unconscious differences. Particularly, our HMT effort task is substantially affected by large variation in hand muscle strength within our samples. Differences in effort might go unnoticed due to this high interpersonal variance. Therefore, in Experiment 3 we switched from a physical measure to a dexterity task, which is less prone to inter-individual differences.

On another note, both our temperature judgment and our decision task were generally considered to be very difficult, and task difficulty might have interfered with negative reactions in the ongoing cooperation. High task difficulty could have limited participants' maximum effort in advice giving, for example, if participants perceive the benefit of additional cues to be negligible. In contrast, an easier task might allow for a higher effort due to an increased perceived benefit of additional cues. As a consequence, there would be more potential to lower the effort in the ongoing cooperation. Hence, a negative effect of advice taking on the ongoing cooperation might be observed under these conditions. To ensure that the absence of effects on the ongoing cooperation in Experiments 1 and 2 do not depend on task difficulty, we decided to use a less demanding type of judgment task in Experiment 3.

Method

Participants and design. Participants were recruited using ORSEE. Fourty-three out of 74 participants were female (58%) with an average age of $M = 25.11$ ($SD = 5.45$) years. No participants had to be excluded from the analyses.

Task and procedure. The procedure in Experiment 3 was the same as in Experiment 1, with the exception of using a new effort measure as well as a new type of judgment task. In each trial of the experiment, participants were shown an urn containing 100 red and blue balls. The proportions of red and blue balls inside the urn differed with every trial, and they were not made known to participants. Instead, based on a sample of balls drawn from and then returned to the urn, participants were asked to estimate the general probability of drawing a red ball from the presented urn. It was emphasized that seeing a larger sample of balls drawn from the urn would (on average) lead to more accurate probability judgments. In the training phase, the size of the sample was randomly varied to ensure that participants fully recognized the relevance of sample size for judgment accuracy.

In the main phase of the experiment, the size of a ball sample shown to participants was determined by participants' respective performance in our new measure of effort in advice giving. Our participants had to track a moving circle on a screen with a computer mouse for a fixed period of 30 seconds. Participants had to keep the cursor directly on the circle for as long as possible within the tracking period. Afterwards, we computed the proportion of time on target, in which the curser had been on the circle's surface, so that a higher proportion indicated better

tracking performance.[4] Tracking performance was used to determine the sample size of balls shown to participants in the subsequent urn judgment task. Sample sizes varied between 4 balls for tracking time proportions of 30 percent or below, and 20 balls for a tracking percentage of 100 percent. In Experiment 3, tracking performance was also used as our primary indicator of effort in advice giving.

Results

Detailed results for all analyses of variance and covariance which are reported in this section can be found in Table 3.

Manipulation checks. In accordance with our preceding studies, participants in low advice taking conditions perceived a significantly lower amount of advice taking (M = 27.00, SD = 20.17) than did participants in high advice taking conditions (M = 77.58, SD = 15.63), $t(72) = -12.01$, $p < .001$, $d = 2.79$.

The ANOVA for dependency ratings showed that participants perceived final judgments to depend less on their advice in low advice taking conditions than in high advice taking

---

[4] In an additional pilot study, we ensured that tracking performance was suited to measure motivational differences. We compared the tracking performance of 122 participants, of which 62 participants received a performance-linked compensation, and 60 participants received a fixed compensation. Our analysis showed a significantly higher proportion of time on target in the performance-linked fee group (M = 0.84, SD = 0.09), than in the fixed fee group (M = 0.74, SD = 0.15), $t(120) = 4.48$, $p < .001$, $d = 0.81$. Hence, higher task motivation (due to the performance-linked compensation) led to better tracking performance.

conditions (*M* = 3.39, *SD* = 1.41 vs. *M* = 5.39, *SD* = 0.99), *F*(1, 70) = 49.89, *p* < .001, η_p^2 = 0.42. There were no main or interaction effects of relative expertise, both $Fs \leq 2.49$, $ps \geq .119$, $\eta_p^2 s \leq .03$.

In our manipulation check for relative expertise, 74% of the participants correctly remembered whether they or their partner had achieved a higher rank in the bogus feedback after the training phase. Participants' recall performance was clearly above chance levels, $\chi^2(1, N = 74) = 17.45$, p < .001.

The difference between the rated performance of the judges and the self-rated performance of our advisors was also tested for differences between the relative expertise conditions. The analysis revealed that the participants felt to be more competent than their respective judge, both in conditions with more competent advisors and in conditions with more competent judges (*M* = -1.00, *SD* = 1.68 vs. *M* = -0.22, *SD* = 1.73). However, the difference was significantly larger in conditions with more competent advisors *t*(72) = 2.30, *p* = 0.024, *d* = 0.53, indicating that our participants did react to the manipulation of relative expertise. In sum, our experimental manipulations were once again successful.

Effort in advice giving. Participants unlocked an average sample size of *M* = 11.58 balls (*SD* = 2.83). As our new measure of effort in advice giving, tracking proportions were analyzed in a 2 x 2 ANOVA. In accordance with Experiments 1 and 2, there were no effects of the amount of advice taking, relative expertise, or their interaction on tracking performance, all $Fs < 1$, $ps \geq .394$, $\eta_p^2 s \leq .01$.

Also, no effects on subjective effort ratings were found, all $Fs < 1$, $ps \geq .599$, $\eta_p^2 s < .01$.

Advice quality. When analyzing participants' accuracy (MAE scores) in the urn estimation task, we found no significant effect of advice taking, relative expertise condition, or their interaction, all Fs < 1, ps $\geq$ .526, η_p^2s $\leq$ .01.

Willingness to cooperate further. A logistic regression with desire to switch partners revealed a significant effect of advice taking, *B* = 2.51 (*SE* = 1.13), *OR* = 12.27, *OR 95%-CI* = [1.90,243.92], *p* = .026. As in our previous experiments, there were more participants wishing to switch partners in low advice taking conditions (47%) than in high advice taking conditions (19%). Neither effect of relative expertise, *B* = -0.20 (*SE* = 0.66), *OR* = 0.82, *OR 95%-CI* = [0.22, 2.95], *p* = .758, nor the interaction were significant, *B* = -1.66 (*SE* = 1.31), *OR* = 0.19, *OR 95%-CI* = [0.01, 1.99], *p* = .207.

Similar to Experiment 2, we found no effect of advice taking, relative expertise, or their interaction on participants' rated motivation to continue working with their current partner, all Fs < 1, ps $\geq$ .623, η_p^2s < .01.

Perceived fairness. The ANOVA analyzing participants' ratings of their respective partner's fairness showed a significant effect of advice taking, $F(1, 70) = 69.35$, $p < .001$, $\eta_p^2 = .50$, with higher fairness ratings for participants in high advice taking conditions than for participants in low advice taking conditions (*M* = 5.97, *SD* = 0.99 vs. *M* = 3.61, *SD* = 1.39). Similar to the previous experiments, we found neither an effect of relative expertise nor the interaction, both Fs < 1, ps $\geq$ .521, η_p^2s $\leq$ .01.

Table 3

Summary of computed ANOVAs in Experiment 3

Source	*df*	*F*	*p*	η_p^2
DV 1: Tracking Proportions				
Advice Taking (AT)	1	< 1	.552	.01
Feedback (FB)	1	< 1	.394	.01
AT x FB	1	< 1	.609	.00
Residuals	70			
DV 2: Effort Ratings				
Advice Taking (AT)	1	< 1	.691	.00
Feedback (FB)	1	< 1	.599	.00
AT x FB	1	< 1	.796	.00
Residuals	70			
DV 3: Advice Quality				
Advice Taking (AT)	1	< 1	.882	.00
Feedback (FB)	1	< 1	.606	.00
AT x FB	1	< 1	.526	.01
Residuals	70			
DV 4: Motivation Ratings				
Advice Taking (AT)	1	< 1	.623	.00
Feedback (FB)	1	< 1	.704	.00
AT x FB	1	< 1	.718	.00
Residuals	70			
DV 5: Fairness Ratings				
Advice Taking (AT)	1	69.35	< .001	.50
Feedback (FB)	1	< 1	.521	.01
AT x FB	1	< 1	.753	.00
Residuals	70			
DV 6: Dependency Ratings				
Advice Taking (AT)	1	49.66	< .001	.42
Feedback (FB)	1	< 1	.489	.01
AT x FB	1	2.49	.119	.03
Residuals	70			

Discussion

Using a new type of measure of effort in advice giving and a different multiple cue judgment task, we replicated the general pattern of our first two experiments. Only a minor difference in the general pattern of results occurred with advice taking effects being absent for motivation ratings (significant effect in Experiment 1). As in Experiments 1 and 2, we observed negative consequences of disregarding advice for advisors' willingness to cooperate with the judge in the future, but not for the ongoing effort in advice giving. Although participants felt being treated unfairly if their advice was neglected, they did not lower their effort. Once more, relative expertise differences had no moderating effects on the observed consequences for fairness ratings and for willingness to cooperate in the future.

For our fourth experiment, we decided to take a closer look at this absence of moderating effects of relative expertise. There are two possible explanations for the pattern we observed in Experiments 1, 2, and 3. On the one hand, advisors might be fully aware of the implications of relative expertise differences, but choose to react negatively whenever their advice is disregarded. In other words, they know that weighting advice from a less competent source decreases judgment accuracy but they, nevertheless, expect a judge to follow their advice to ensure further cooperation. On the other hand, advisors might fail to realize the implications of relative expertise differences, although they are aware of their presence in a JAS. In this case, they would not understand that advice from someone who is vastly less competent than the judge should be almost neglected if one aims at making the best possible judgment, and that only advice from a competent source should be strongly weighted. Hence, instead of advisors deliberately choosing to ignore the implications of expertise differences, advisors might simply fail to realize these implications.

Experiment 4

In our fourth experiment, we wanted to test whether the findings of our previous experiments would still hold if we made it unequivocally clear to participants how strongly their advice should be weighted on rational grounds. Therefore, we did not manipulate perceived expertise differences, as we did in our previous experiments. Instead, we presented participants with direct feedback regarding relative accuracy of both their advice and the judge's initial opinion, as well as the optimal amount of advice taking after each trial. By directly showing participants whether - for the sake of judgment accuracy - their advice should have been weighted to a high amount or to a low amount, we made sure that the participants understood the rational basis (or lack thereof) of a high vs. low advice weighting by the judge.

If the pattern of results that we observed in Experiments 1, 2, and 3 replicates in Experiment 4, this would mean that advisors do not care about such a rational basis. If, however, we do observe moderating effects of this factor in Experiment 4 (i.e., participants react negatively towards advisors who neglect them if this neglect is not warranted on a rational basis), this would mean that advisors do, indeed, care about such rational considerations but have difficulties to infer these considerations from information about expertise differences.

Method

Participants and design. Again, participants were recruited using our ORSEE participant database. Of the 141 participants with an average age of $M = 23.5$ ($SD = 4.12$), 84 participants (60%) were female.

Task and procedure. The procedure in Experiment 4 was similar to Experiment 1 with the exception of using a new effort measure as well as a new type of performance feedback. Since we observed consistent effects of advice taking irrespective of the type of task, we decided

to use the temperature judgments from Experiment 1 again. Yet, instead of having to unlock the temperature cues by pressing the HMT, participants were presented with a set of 30 sliders which had to be moved to specific positions. Each slider position reflected a value ranging from 0 to 100. A target value was shown on the left side of the slider, while the value of the current slider position was displayed on the right side. In each trial of the slider task, participants had to set as many sliders as possible to their target values within 30 seconds. The more sliders had been set to their target value after 30 seconds, the more cues were shown in the temperature judgments. To unlock 1, 2, 3, 4, or 5 cues, participants had to set 4, 6, 8, 10, or 15 sliders to their specific target values, respectively. The slider task is a finely gradated measure of effort (Gill & Prowse, 2012) that bears low intrinsic motivation (Gerhards & Siemer, 2016), thereby making it an optimal choice for our purposes.

After each temperature judgment, participants received our new bogus accuracy feedback. We computed the alleged judge's initial and final judgments in relation to both the participant's advice and the true criterion value at the end of the trial. Subsequently, we provided participants with graphical feedback (see Figure 1 for an example) showing information on the relative positions of initial estimate, final estimate, advice, and true value on the judgmental continuum. This feedback visualized the weight a judge had placed on the advice as well as the accuracy of all judgments. In conditions with more accurate advisors, the advice was presented to be much closer to the correct temperature that the judge's initial opinion. In this case, a high amount of advice taking was warranted. In conditions with more accurate judges, the judge's initial opinion was computed to be much closer to the true temperature than the advice, warranting low advice taking. Hence, participants saw whether following an advice was beneficial or detrimental to the accuracy of the final judgment. To make the last aspect even

more salient, participants were also informed about the weight of advice that would have maximized the accuracy of the final estimate, since it would have resulted in the final estimate to be the true value. In addition, we also presented to them the actual amount of advice taking. Both optimal and actual advice taking were presented as a percentage value. If the optimal weight of advice took on negative values or values greater than 100%, it was displayed as "< 0" or "> 100". Since we manipulated relative accuracy by giving feedback after each trial, there was no initial rank feedback given to participants.

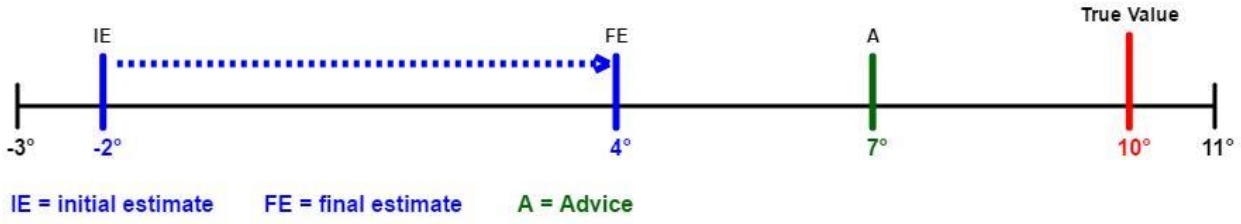

Figure 1. Example of the accuracy feedback shown to participants after each trial in Experiment 4.

Instead of measuring effort in advice giving through HMT pressing durations or tracking proportions, we calculated the mean number of sliders correctly set to the target value in each trial of the test phase. The quality of advice was again measured by advisors' MAPE values.

Results

Detailed results for all analyses of variance and covariance which are reported in this section can be found in Table 4.

Manipulation checks. As in our preceding studies, participants in low advice taking conditions perceived lower amounts of advice taking than did participants in high advice taking conditions (M = 15.28, SD = 17.24 vs. M = 71.34, SD = 17.33), $t(139) = 19.19$, $p < .001$, d = 3.24. Hence, our advice taking manipulation was successful.

As in Experiment 3, a 2 x 2 ANOVA for dependency ratings showed that participants perceived final judgments to depend less on their advice in low advice taking conditions than in high advice taking conditions (M = 2.71, SD = 1.04 vs. M = 4.93, SD = 0.91), $F(1, 137)$ = 184.23, $p < .001$, $\eta_p^2 = 0.57$. There were no effects for relative expertise, or the interaction of advice taking and relative expertise, both $Fs \leq 2.25$, $ps \geq .136$, $\eta_p^2 s \leq .02$.

Participants in conditions with more accurate judges perceived an accuracy difference in favor of the judge, while participants in conditions with more accurate advisors perceived an accuracy difference in favor of themselves (M = 2.40, SD = 1.59 vs. M = -1.08, SD = 1.50), $t(139) = 13.26$, $p < .001$, $d = 2.45$. Taken together, the experimental manipulation of both factors was successful.

Effort in advice giving. On average, participants in Experiment 4 unlocked M = 3.15 (SD = 0.73) out of five available cues by setting sliders to their target values. Analyzing slider task performance with an ANOVA of our experimental design, we did not find any significant effects of advice taking, accuracy feedback condition, or their interaction, all $Fs \leq 1.18$, $ps \geq .280$, $\eta_p^2 s \leq .01$. The same pattern of results was found for participants' subjective effort ratings, all $Fs \leq 2.02$, $ps \geq .158$, $\eta_p^2 s \leq .01$.

Advice quality. We analyzed participants' MAPE value as an indicator of advice quality. Conducting a 2x2 ANOVA, we found no effects of advice taking, feedback condition, or their interaction, all $Fs \leq 3.77$, $ps \geq .054$, $\eta_p^2 s \leq .03$.

Willingness to cooperate further. We analyzed desire to switch partners by conducting a binary logistic regression analysis with advice taking condition and relative accuracy as well as their interaction term as predictors. We found no main effect of advice taking, B = -0.71 (SE = 0.72), OR = 0.49, $OR\ 95\%\text{-}CI$ = [0.10, 1.87], p = .324. However, we did find a significant effect

of accuracy feedback, B = -2.57 (SE = 0.70), OR = 0.08, OR *95%-CI* = [0.16, 0.27], $p < .001$, which was qualified by a significant interaction of accuracy feedback and amount of advice taking, B = 2.63 (SE = 0.94), OR = 13.81, OR *95%-CI* = [2.36,99.1], $p = .005$. To disentangle the interaction, we conducted separate regression analyses for both accuracy feedback conditions. For participants in conditions with more accurate judges, the analysis showed no effect of advice taking, B = -0.71, SE = 0.72, OR = 0.49, OR *95%-CI* = [0.10, 1.87], $p = .324$. In contrast, there was a significant effect of advice taking for participants in conditions with more accurate advisors, B = 1.91, SE = 0.60, OR = 6.79, OR *95%-CI* = [2.20, 24.23], $p = .002$. As shown in Figure 2, when advisors were more accurate than the judge, they significantly more often wished to switch partners in the low advice taking condition than in the high advice taking condition (58% vs. 17%). However, when the judge was more accurate, there was no significant difference in the desire to switch partners between the low and the high advice taking condition (9% vs. 17%).

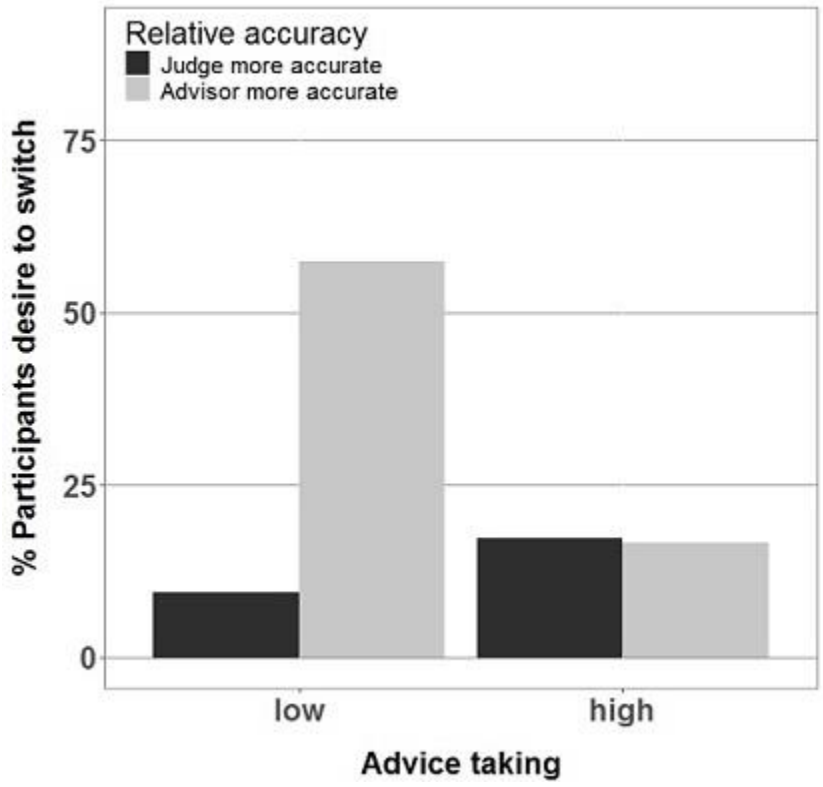

Figure 2. Number of participants desiring to switch partners depending on the amount of advice taking and relative accuracy in Experiment 4.

Participants' ratings regarding their motivation to continue working with their current partner revealed a significant effect of advice taking, $F(1, 137) = 4.69$, $p = .032$, $\eta_p^2 = .03$. Their motivation was significantly lower in low advice taking conditions than in high advice taking conditions ($M = 3.80$, $SD = 1.65$ vs. $M = 4.36$, $SD = 1.57$). There were no main or interactive effects of accuracy feedback condition, both $Fs \leq 2.11$, $ps \geq .148$, $\eta_p^2 s \leq .02$.

Perceived fairness. When analyzing fairness ratings, the results showed significant effects of advice taking, $F(1, 137) = 111.28$, $p < .001$, $\eta_p^2 = .45$, and for relative accuracy, $F(1, 137) = 6.61$, $p = 0.011$, $\eta_p^2 = .05$. Participants generally rated their partner's behavior as being more fair in high advice taking conditions than in low advice taking conditions ($M = 6.08$, $SD =$

0.91 vs. M = 3.98, SD = 1.41). Also, participants in conditions with more accurate judges perceived a higher average fairness than participants in conditions with more accurate advisors (M = 5.44, SD = 1.43 vs. M = 4.71, SD = 1.64). There was no interaction effect, $F(1, 137) = 1.11$, $p = .293$, $\eta_p^2 = .01$.

Table 4

Summary of computed ANOVAs in Experiment 4

Source	*df*	*F*	*p*	η_p^2
DV 1: Amount of Solved Sliders				
Advice Taking (AT)	1	< 1	.945	.00
Feedback (FB)	1	1.18	.280	.01
AT x FB	1	< 1	.781	.00
Residuals	137			
DV 2: Effort Ratings				
Advice Taking (AT)	1	2.02	.158	.01
Feedback (FB)	1	< 1	.592	.00
AT x FB	1	< 1	.679	.00
Residuals	137			
DV 3: Advice Quality				
Advice Taking (AT)	1	< 1	.910	.00
Feedback (FB)	1	3.77	.054	.03
AT x FB	1	< 1	.420	.00
Residuals	137			
DV 4: Motivation Ratings				
Advice Taking (AT)	1	4.69	.032	.03
Feedback (FB)	1	< 1	.985	.00
AT x FB	1	2.11	.148	.02
Residuals	137			
DV 5: Fairness Ratings				
Advice Taking (AT)	1	111.28	< .001	.45
Feedback (FB)	1	6.61	.011	.05
AT x FB	1	1.11	.293	.01
Residuals	137			
DV 6: Dependency Ratings				
Advice Taking (AT)	1	184.23	< .001	.57
Feedback (FB)	1	2.25	.136	.02
AT x FB	1	< 1	.882	.00
Residuals	137			

Discussion

In accordance with our previous experiments, our analyses revealed negative reactions to low amounts of advice taking in participants' rated motivation to continue working with their current partner and in their ratings of their partners' fairness. However, considering wishes to switch partners, only participants in our conditions with more accurate advisors showed increased intentions to switch partners when advice taking was low, while there was no such effect in conditions with more accurate judges. This moderating effect of relative accuracy indicates that, if participants have unambiguous information stating that low amounts of advice taking are justified in a JAS, they do not hold low weights of advice against the judge. However, the information we provided our participants with in this experiment will, under almost all circumstances that we can think of, not be available to advisors in real-world settings. Therefore, judges in a JAS usually should expect negative reactions to advice neglect even if it is justified in terms of relative expertise. In contrast to a judge's sensitivity for expertise differences (Harvey & Fischer, 1997), an advisor may not infer the justification of differential weighting from perceived expertise differences. By providing artificial accuracy feedback, we were able to clarify that this observed indifference is not based on a deliberate disregard of the implications of relative expertise differences.

General Discussion

In a series of four experiments, we tested whether and how advisors in a JAS react to judges (almost completely) neglecting their advice. We confronted our participants, taking on the role of advisors in a JAS, with judges showing either high or low amounts of advice taking. Additionally, we manipulated participants’ beliefs about the relative expertise of judge and advisor, thereby creating conditions under which low amounts of advice taking were either

rationally warranted or, to the contrary, were unwarranted. We were then able to observe the following pattern of results.

First, advisors perceived low weights of advice as being unfair. Second, neither the perceived unfairness nor the fact that their advice did not substantially impact the judge's final judgment or decision leads advisors to reduce their effort – or the quality of their advice – while working in the current JAS. Third, negative consequences of low weights of advice manifested in reduced willingness to engage in future cooperation. Finally, these negative reactions were unaffected by the relative expertise of judge and advisor. That is, unless advisors had clear evidence that considering their advice would have decreased the judge's accuracy (as in Experiment 4, where we created conditions that almost never occur in real life), they reacted negatively to low weights of advice even if the superior expertise of the judge justified such low weights.

We will discuss these findings in more detail in the following sections.

Effort in the Ongoing Cooperation

As Adams (1965) pointed out, partners who perceive unfairness in a social exchange might immediately adjust their own contributions to reestablish equity. Remarkably, however, using multiple effort measures we did not find any evidence of advisors lowering their effort in advice giving when judges exhibited low amounts of advice taking.

There are several possible explanations for the observed pattern of results which we partially addressed in our experiments. First of all, we accounted for the possibility of our measures not being fully suitable to show actual effort differences. We used objective effort indicators in addition to self-reported effort in case our participants were either not aware of or not willing to communicate lowered effort in the ongoing cooperation. By using three different

types of effort tasks, we also made sure that neither high inter-individual variance nor high intrinsic motivation interfered with our search for effort differences.

Like in many laboratory studies, we are not able to fully rule out the possibility of a demand effect resulting from the experimental situation. Participants might have feared that, by lowering their effort in advice giving, they would breach their obligations as participants in a psychological experiment, since they might believe that we want our participants to show high effort. However, we would expect such a demand effect to manifest itself in participants' motivation ratings as well (or even to a stronger extent). Since we did find effects of the amount of advice taking on the motivation ratings, we do not think that the absence of differences in effort ratings, rated importance to give good advice, and also our effort tasks can be explained with an experimental demand effect.

However, we do think that our participants might have been motivated to do their best in giving good advice by the rules that underlie social exchange. When establishing a reciprocal social exchange, mutual obligations based on social norms are generated for all participating parties (Cropanzano & Mitchell, 2005). According to Sniezek and colleagues (2004), judges and advisors enter a psychological contract based on their subjectively perceived mutual obligations. As we are just starting to investigate this psychological contract from the advisor's perspective, we can only speculate about the reasons for advisors not to adjust their effort after perceiving unfair judge behavior. Perhaps the psychological contract, being based on strong cooperation norms, is of such binding power that perceiving a norm violation is not a sufficient reason for advisors to breach the contract themselves. Alternatively, participants might generally be afraid to violate other social norms by lowering their effort, for example a norm to always give their best when helping others, or to always help when being asked to. Thus, deliberately giving less

than optimal advice might, from their perspective, not have been a viable option to counter the unfairness they evidently perceived. This explanation is also supported by Dana, Cain, and Dawes (2006), who suggest that experimental participants' actions are driven by a desire not to violate others' expectations. We are confident that additional research into this matter will allow us to gain important insights into the rules and motives that guide the interaction judge and advisor in a JAS.

Consequences for Future Cooperation

Whereas we did not find immediate negative consequences of judges exhibiting low amounts of advice taking, we did observe negative consequences regarding advisors' motivation and willingness for a continued cooperation with their alleged partner. In all of our experiments, advisors in low advice taking conditions wished to switch partners significantly more often than advisors in high advice taking conditions. Additionally, in three out of four experiments, advisors were less motivated to continue working with their current partner when this partner showed a low (as compared to high) amount of advice taking. Terminating a cooperation is one of the more severe consequences of rule violations in a social exchange (Cropanzano & Mitchell, 2005). In this respect, our findings support Harvey and Fischer's (1997) assumption that insufficient weighting might lead to advice not being offered in the future.

Regarding the question of why advisors choose an indirect approach to balance a perceived unfairness over a direct retaliation against their partner, we think that the rules of social exchange, as described above, offer a good explanation. By switching partners and avoiding future cooperation, our participants evade an unfair situation without having to violate norms of cooperation, or the expectations of their respective partners (c.f. Dana et al., 2006). Yet, the observed behavior gives rise to questions about effective punishment and immediate

compensation of suffered unfairness. As Fehr and Gächter (2002) argue, punishment is an important factor in producing human cooperation. Even noninvolved third parties engage in costly punishment of free riding and uncooperative behavior to enforce reciprocity in social exchanges (Fehr & Fischbacher, 2004a, 2004b). In contrast, avoiding future cooperation and switching partners are quite silent forms of punishment that will not necessarily come to a judge's attention. Also, by waiving a direct reaction, advisors miss the opportunity to immediately compensate the unfairness they perceive. Instead, advisors tolerate their partner's behavior until the ongoing cooperation ends. In sum, we think that the binding power of social norms and the psychological contract is a convincing explanation of our participants' behavior.

Relative Expertise Differences

We did not find any evidence of moderating effects of perceived expertise differences in our first three experiments. Only by providing advisors with artificial feedback about the warranted amount of advice taking, we were able to observe a mitigating effect on the negative consequences following advice neglect. Under more realistic circumstances, however, advisors will fail to infer the appropriateness of disregarding advice from expertise differences which they obviously perceive (as demonstrated by the success of all of our feedback manipulations). Furthermore, in many real-world situations, information on expertise differences might also not be available to an advisor, thus making it virtually impossible for him or her to infer whether advice neglect is justified or not.

As mentioned before, our findings suggest an important discrepancy between the judge perspective in a JAS, as assessed in previous research, and the advisor perspective that we are just beginning to investigate: While judges take relative expertise differences into account when they decide about whether and how to follow advice, advisors' expectations regarding sufficient

amounts of advice taking do not factor in these differences. We think that judges' and advisors' different handling of perceived expertise differences might originate from diverging perceptions of the psychological contract underlying the JAS. According to Sniezek and colleagues (2004), "judges and advisors will be likely to have different, often incompatible, perceptions of the psychological contract" (p. 188). Hence, while judges' perceptions of the psychological contract might focus on achieving the best outcome for the JAS, advisors' perceptions might be concentrated on an equal influence on the decision process for all JAS members. This could explain why advisors, without being vehemently pushed, do not draw the same conclusions from perceived expertise differences as do judges in a JAS. Unfortunately, due to the fact that the advisor perspective has not been addressed in previous research, we can only speculate about advisors' motives. This illustrates, once again, the importance of further investigating the advisor perspective.

Implications

Our results create an interesting puzzle when looking at a robust finding in previous research on advice taking, namely advice discounting. In the prototypical advice taking study, the mean weight of advice is around 30%, with judges frequently ignoring the advice altogether (Soll & Larrick, 2009). Based on our findings, we would expect this advice taking behavior to make advisors almost furious (even if they do not communicate their fury to the judge). The obvious explanation is that judges in previous studies – without exception – made their estimates privately or, at least, in the absence of the advisor. Assuming that the judge can take the perspective of the advisor and, thereby, anticipate potential negative reactions to low weights of advice, we would assume that advice discounting is far less pronounced when advisors know whether and to what extend the judge heeds the advice.

Regarding real world judge-advisor interactions, we propose that judges should seriously consider advisors' expectations regarding advice taking. The observations and assumptions made by Goldsmith and Fitch (1997) seem to prove true: Respect and gratitude, conveyed through the act of advice taking, are expected in exchange for advice. As our experiments show, advisors do react negatively if their expectations are not met, and even though these reactions seemingly do not affect an ongoing exchange, judges often have to ensure an advisor's willingness to cooperate in future interactions. When considering the amount of advice taking, judges should be aware that rational reasons to disregard advice might not be perceived by advisors if they are not extremely strong and convincing. However, in real world interactions judges have many more possibilities to value advice besides the amount of advice taking exhibited, for example through verbal appraisal or nonverbal communication. The role of advice taking as a means of conveying respect and gratitude might, therefore, be less important in a real-world JAS. In many cases, we would expect a judge, who received a low-quality advice, to separate advice taking and the compensation of an advisor from each other. For example, the advisor could verbally convey his gratitude and, at the same time, explain to an advisor why he will not follow the advice. In doing so, judges might be able to disregard advice without risking the negative consequences we observed in our experiments. To test this assumption, we would need to investigate advisors in more interactive settings that impose fewer restrictions on communication between judge and an advisor, as was also suggested by Sniezek and colleagues (2004).

Limitations and Directions for Future Research

In our view, the most important limitation of our approach to the advisor perspective in a JAS is the fact that no communication between judges and advisors was allowed, besides the mere exchange of judgments and estimates. As already stated, to fully understand the dynamics

of a JAS, its social context has to be accounted for. Nonetheless, we think that relying on an experimental setting that is similar to classical JAS studies made a lot of sense as a starting point for this line of research, because it would have been very difficult to control for all possible confounds stemming from free verbal communication between participants. For example, by exchanging their subjectively perceived expertise, our participants could have undermined our relative expertise manipulation. Another unwanted consequence of unrestrained verbal communication could have been that judges might have mitigated the negative effects of low amounts of advice taking by verbally appreciating received advice. However, as a next step in this line of research, we think it is useful to allow for more interaction between judges and advisors. Among other things, this would allow to investigate the consequences of low amounts of advice taking and judges' actions to mitigate such consequences in a realistic setting.

Regarding the generalizability of our findings, we think it is important to note that participants in our experiments believed not to know their partners personally. We think that exchange situations in which judges and advisors know each other differ considerably from situations in which they do not. Meeting the demands and rules of a social exchange might be even more important if the participating parties have a long-term relationship with each other. This might lead to judges feeling more pressure to conform with the rules of social exchange, on the one hand, and to advisors reacting more negatively and more intense to low amounts of advice taking, on the other hand. Therefore, we do not believe that our current findings easily generalize to JAS incorporating judges and advisors that are personally acquainted. Instead, future studies should also address differences in the social dynamics of JAS with different underlying conditions.

Conclusion

With our experiments, we took a first step towards the investigation of the advisor perspective in a judge-advisor system. Summarizing our results, we found substantial evidence for advisors reacting negatively to the rejection of advice. While these negative reactions were rather resistant to a rational justification for the judge's behavior, their consequences were also strictly limited to future cooperation. Hence, we think that Harvey and Fischer's (1997) assumption does hold: Judges in a judge-advisor system should not carelessly reject advice offered by an advisor, even if they think that the advice offers little in terms of new information. Advisors who have been previously rejected might withhold their advice at some point in the future when it is desperately needed. When it comes to advice, it may be wise not to look the proverbial gift horse in the mouth.

References

Adams, J. S. (1965). Inequity in social exchange. *Advances in Experimental Social Psychology*, *2*(267-299).

Back, M. D., Küfner, A. C. P., Dufner, M., Gerlach, T. M., Rauthmann, J. F., & Denissen, J. J. A. (2013). Narcissistic admiration and rivalry: Disentangling the bright and dark sides of narcissism. *Journal of Personality and Social Psychology*, *105*(6), 1013–1037. https://doi.org/10.1037/a0034431

Bednarik, P., & Schultze, T. (2015). The effectiveness of imperfect weighting in advice taking. *Judgment and Decision Making*, *10*(3), 265.

Bonaccio, S., & 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*, 127–151.

Cropanzano, R., & Mitchell, M. S. (2005). Social exchange theory: An interdisciplinary review. *Journal of Management*, *31*(6), 874–900. https://doi.org/10.1177/0149206305279602

Dana, J., Cain, D. M., & Dawes, R. M. (2006). What you don't know won't hurt me: Costly (but quiet) exit in dictator games. *Organizational Behavior and Human Decision Processes*, *100*(2), 193–201.

Emerson, R. M. (1976). Social exchange theory. *Annual Review of Sociology*, *2*, 335–362.

Fehr, E., & Fischbacher, U. (2004a). Social norms and human cooperation. *Trends in Cognitive Sciences*, *8*(4), 185–190. https://doi.org/10.1016/j.tics.2004.02.007

Fehr, E., & Fischbacher, U. (2004b). Third-party punishment and social norms. *Evolution and Human Behavior: Official Journal of the Human Behavior and Evolution Society*, *25*(2), 63–87. https://doi.org/10.1016/S1090-5138(04)00005-4

Fehr, E., & Gächter, S. (2002). Altruistic punishment in humans. *Nature*, *415*(6868), 137–140. https://doi.org/10.1038/415137a

Gerhards, L., & Siemer, N. (2016). THE IMPACT OF PRIVATE AND PUBLIC FEEDBACK ON WORKER PERFORMANCE—EVIDENCE FROM THE LAB. *Economic Inquiry*, *54*(2), 1188–1201. https://doi.org/10.1111/ecin.12310

Gill, D., & Prowse, V. (2012). A structural analysis of disappointment aversion in a real effort competition. *The American Economic Review*, *102*(1), 469–503.

Goldsmith, D. J., & Fitch, K. (1997). The normative context of advice as social support. *Human Communication Research*, *23*(4), 454–476. https://doi.org/10.1111/j.1468-2958.1997.tb00406.x

Greiner, B. (2015). Subject pool recruitment procedures: Organizing experiments with ORSEE. *Journal of the Economic Science Association*, *1*(1), 114–125. https://doi.org/10.1007/s40881-015-0004-4

Guay, F., Vallerand, R. J., & Blanchard, C. (2000). On the assessment of situational intrinsic and extrinsic motivation: The situational motivation scale (SIMS). *Motivation and Emotion*, *24*(3), 175–213. https://doi.org/10.1023/A:1005614228250

Harvey, N., & Fischer, I. (1997). Taking advice: Accepting help, improving judgment, and sharing responsibility. *Organizational Behavior and Human Decision Processes*, *70*(2), 117–133. https://doi.org/10.1006/obhd.1997.2697

Harvey, N., & Harries, C. (2004). Effects of judges' forecasting on their later combination of forecasts for the same outcomes. *International Journal of Forecasting*, *20*(3), 391–409. https://doi.org/10.1016/j.ijforecast.2003.09.012

Hertel, G., Kerr, N. L., & Messé, L. A. (2000). Motivation gains in performance groups: Paradigmatic and theoretical developments on the Koehler effect. *Journal of Personality and Social Psychology*, *79*(4), 580–601.

Kerr, N. L. (1983). Motivation losses in small groups: A social dilemma analysis. *Journal of Personality and Social Psychology*, *45*(4), 819. https://doi.org/10.1037/0022-3514.45.4.819

Molm, L. D., Takahashi, N., & Peterson, G. (2000). Risk and trust in social exchange: An experimental test of a classical proposition. *The American Journal of Sociology*, *105*(5), 1396–1427.

Muraven, M., Tice, D. M., & Baumeister, R. F. (1998). Self-control as a limited resource: Regulatory depletion patterns. In *Midwestern psychological convention, may, 1995, chicago, IL, US* (Vol. 74, p. 774). American Psychological Association. https://doi.org/10.1037/0022-3514.74.3.774

Murphy, R. O., Ackermann, K. A., & Handgraaf, M. (2011). Measuring social value orientation. *Judgement and Decision Making*, *6*(8), 771–781. https://doi.org/10.2139/ssrn.1804189

Rousseau, D. M. (1989). Psychological and implied contracts in organizations. *Employee Responsibilities and Rights Journal*, *2*(2), 121–139. https://doi.org/10.1007/BF01384942

Schmitt, M., Baumert, A., Gollwitzer, M., & Maes, J. (2010). The justice sensitivity inventory: Factorial validity, location in the personality facet space, demographic pattern, and normative data. *Social Justice Research*, *23*(2-3), 211–238. https://doi.org/10.1007/s11211-010-0115-2

Schultze, T., Rakotoarisoa, A.-F., & Schulz-Hardt, S. (2015). Effects of distance between initial estimates and advice on advice utilization. *Judgment and Decision Making*, *10*(2), 144.

Sniezek, J. A., & Buckley, T. (1995). Cueing and cognitive conflict in Judge-Advisor decision making. *Organizational Behavior and Human Decision Processes*, *62*(2), 159–174. https://doi.org/10.1006/obhd.1995.1040

Sniezek, J. A., Schrah, G. E., & Dalal, R. S. (2004). Improving judgement with prepaid expert advice. *Journal of Behavioral Decision Making*, *17*(3), 173–190. https://doi.org/10.1002/bdm.468

Soll, J. B., & Larrick, R. P. (2009). Strategies for revising judgment: How (and how well) people use others' opinions. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, *35*(3), 780–805. https://doi.org/10.1037/a0015145

Soll, J. B., & Mannes, A. E. (2011). Judgmental aggregation strategies depend on whether the self is involved. *International Journal of Forecasting*.

Watson, D., Clark, L. A., & Tellegen, A. (1988). Development and validation of brief measures of positive and negative affect: The PANAS scales. *Journal of Personality and Social Psychology*, *54*(6), 1063–1070.

Yaniv, I. (2004). The benefit of additional opinions. *Current Directions in Psychological Science*, *13*(2), 75–78. https://doi.org/10.1111/j.0963-7214.2004.00278.x

Yaniv, I., & Kleinberger, E. (2000). Advice taking in decision making: Egocentric discounting and reputation formation. *Organizational Behavior and Human Decision Processes* *83*(2), 260–281. https://doi.org/10.1006/obhd.2000.2909

Appendix

Additional measures. In Experiments 1, 2, and 3, we asked participants to anonymously assign an additional monetary reward to their alleged partner. They had to indicate a value between 0 and 3 € that would be added to the judges' alleged payout without the judge knowing that they had assigned this bonus. By including this bonus payment, we offered participants a way to silently react to their alleged partner's unfair behavior, without having to fear direct confrontation. To measure which amount of advice taking the participants considered to be sufficient, they were asked to state the average weight they would have placed on advice if they had been assigned the role of the judge, again on a scale ranging from 0 to 100 %. Additionally, participants were asked how important giving accurate advice in the experiment was to them (1 - *not important at all*, 7 - *very important*). At the end of the questionnaire we included a suspicion check and, after revealing that they had interacted with a simulated partner, asked participants to rate the credibility of both the interaction with their alleged partner and the bogus performance feedback.

In Experiment 1, we administered a computer version of the Positive and Negative Affect Schedule (PANAS; Watson, Clark, & Tellegen, 1988) directly before and after the test phase. The affective state before the main phase was then compared to the affective state after the main phase to test for any changes resulting from the cooperation with the simulated judge. We also included the german version of the Narcissistic Admiration and Rivalry Questionnaire (NARQ; Back et al., 2013) as a narcissism scale, and the german version of the Justice Sensitivity Inventory (JSI; Schmitt, Baumert, Gollwitzer, & Maes, 2010) at the end of our our experiment to explore possible effects of inter-individual differences. In Experiment 2, the NARQ as well as the JSI were replaced with an anagram task designed to measure ego depletion (for a detailed

description see Muraven, Tice, & Baumeister, 1998). In our third Experiment, we administered the Situational Motivation Scale (SIMS; Guay, Vallerand, & Blanchard, 2000) directly before and after the main phase of the experiment as an additional explorative measure. In Experiment 4, we included a measure of social value orientation (*SVO Slider Measure*; Murphy, Ackermann, & Handgraaf, 2011) as an additional questionnaire.

Appendix B

Manuscript 2

Disregarding Advice in Judge-Advisor Systems
With Multiple Advisors

Disregarding Advice in Judge-Advisor Systems With Multiple Advisors

Christian Treffenstädt, Thomas Schultze, and Stefan Schulz-Hardt

University of Goettingen

Abstract

Previous research suggests that disregarding advice in a judge-advisor system (JAS) can upset an advisor and thereby harm her or his willingness for future cooperation with a judge, while the ongoing interaction is spared. We follow up on this research by conducting two experiments to test whether this pattern of results would replicate under different circumstances in a multiple-advisor JAS. Participants, taking on the role of one of two advisors in a JAS, were confronted with judges either disregarding their advice in favor of another advisor or favoring them over another advisor. Consistent with previous findings, participants who experienced advice neglect evaluated the JAS interaction more negatively, were less motivated to cooperate further with the judge, and more often expressed a desire to switch partners in their JAS. Perceived expertise differences that would have justified lower amounts of advice taking did not ameliorate these negative consequences of disregarding advice. However, also consistent with previous research, participants did not lower their effort in the ongoing cooperation. The results of the present study yield further confirmation for a specific pattern of negative consequences of judges' disregarding advice.

Keywords: judgment, decision making, advice giving, judge-advisor system, multiple advisors

Disregarding Advice in Judge-Advisor Systems With Multiple Advisors

The Judge-Advisor-System (JAS; Sniezek & Buckley, 1995) is particularly well suited to investigate judgment and decision processes in which decision makers (also called judges) receive advice by one or more advisors. In the JAS paradigm, researchers commonly determine the impact of advice on the decision making process by comparing final judgments or decisions made after receiving advice to uninfluenced opinions expressed in advance of receiving advice. The difference between judges' initial and final opinions serves as an indicator of advice utilization or, in other terms, the weight of advice.

Although offering and receiving advice are, per se, acts of social interaction, past research has paid little attention to the social context underlying the JAS (Sniezek, Schrah, & Dalal, 2004). Instead, most of the previously conducted studies focused on judges and their perceptions and actions within a JAS. As a result, JAS literature holds ample findings regarding the judge's perspective within a JAS, but there is little knowledge on the advisor's perspective. As a first step in addressing this research gap, we conducted a series of experiments investigating advisors' reactions to their advices being disregarded (Treffenstädt, Schultze, & Schulz-Hardt, 2017). Even though advice discounting is one of the most robust findings in JAS research (Bonaccio & Dalal, 2006), the consequences of such advice discounting for an advisor in the JAS remain unknown. Therefore, our aim was to test whether advisors would react negatively to a low amount of advice taking, as it had previously been speculated by advice taking researchers (Goldsmith & Fitch, 1997; Harvey & Fischer, 1997).

In our above-mentioned series of experiments (Treffenstädt et al., 2017), we observed advisors in a computer-mediated judge-advisor-dyad who were confronted with judges utilizing their advice either to a high or to a low amount. We assumed that advisors would perceive low

amounts of advice taking as a violation of reciprocity rules. Subsequently, advisors might "retaliate" by lowering their effort in giving advice, and they might also show a diminished willingness to cooperate in the future. As we predicted, advisors in low advice taking conditions showed negative reactions towards the judge. Compared to advisors in high advice taking conditions, they were less motivated to continue the cooperation, and also more often wished to switch partners when further cooperating in a JAS. Additionally, we tested whether perceived expertise differences between judge and advisor had a moderating effect on the observed reactions. Our results show that perceived expertise differences moderated advisors reactions only when they were artificially emphasized by providing advisors with direct feedback on optimal amounts of advice taking. Under more realistic circumstances, however, advisors reacted negatively, independently of whether the judge had good reasons to discount their advice (due to being much more competent on the task).

Intriguingly, we did not observe any immediate negative reactions regarding advisors' effort during the ongoing interaction. Although advisors in low advice taking conditions did perceive judges' behavior as unfair, and even though they realized that their advice had little influence on JAS performance, they continued to show the same effort in advice giving as did advisors in high advice taking conditions. This pattern was not limited to a specific measure of effort but instead manifested on subjective effort ratings and on several behavioral measures of effort. These measures included pressing a hand muscle trainer for as long as possible, tracking a circle on a computer screen, or adjusting sliders on a computer screen as quickly as possible. Indicating considerable robustness, low weights of advice did not produce differences in effort on any effort measure in the four experiments. Apparently, judges' violation of reciprocity rules did not lead participants to violate the very same rule themselves. Arguably, a strong

psychological contract (Rousseau, 1989), or the desire not to disappoint others' expectations (Dana, Cain, & Dawes, 2006), might have prevented them from lowering their effort in reaction to low amounts of advice taking.

However, due to a methodological limitation inherent in the classical JAS with only one advisor, there is an additional explanation for the observed pattern of results. In such a JAS, discounting the recommendation of an advisor is always relative to the weight that the judge places on his or her own opinion. Hence, disregarding advice automatically means that the judge sticks to his or her initial opinion.

Now, in a JAS there is a clear hierarchical difference between the judge and the advisor. At the end of the day, the judge has to make the final decision, and take responsibility for whatever consequences might occur. This might grant judges something like a “right” to favor their own opinion, even if an advisor might be more competent. If judges have the responsibility, they might be "entitled"" to their own opinion.

However, what would happen if advice neglect does not take place in relation to the judge’s own opinion, but rather in relation to the advice given by another advisor? If there were no hierarchical differences between the two advisors, favoring the one over the other could not be subjectively explained by such differences. In other words, in such a situation it would be unequivocally clear to the focal advisor that the judge does not generally ignore advice, but instead inly neglects the focal advisor's recommendations. To create such a situation, we must move beyond the commonly used single-advisor JAS.

Advice Neglect in a Multiple-Advisor Setting

While advice discounting and maintaining a judge's initial opinion coincide in a single-advisor JAS, the same does not necessarily hold true in multiple-advisor JAS settings. When

forming a final judgment or decision based on an initial opinion and multiple pieces of advice, a judge might decide to favor one piece of advice over the other. This would be justified in the light of diagnostic information that depicts one advisor as more competent than the other. Alternatively, hierarchical differences between the advisors could serve as a reason to discount advice, but in the absence of such differences and without diagnostic information, judges are left without a proper justification for preferring one advisor over the other.

On the one hand, favoring one advisor over the other in such a situation might cause more negative reactions than disregarding advice in a single-advisor JAS. Advisors might not only be less willing to cooperate further, but also might lower their effort in advice giving. In this case, the fact that the favored advice does not represent a superior's opinion might explain the different reactions of the neglected advisor. On the other hand, if the prior pattern of results (lower willingness to cooperate, but no reduction in effort) replicates in the described situation, hierarchical differences cannot account for the specific pattern of advisors' reactions.

Because both alternatives appeared equally plausible to us, we decided to formulate the following competing hypotheses regarding the replication of our previous pattern of results in a multiple-advisor JAS:

Hypothesis 1a: Advisors who experience advice discounting in favor of another advisor will lower their willingness to cooperate further, as well as their ongoing effort to give good advice (as compared to advisors who experience that their own advice is favored over the other advisor's advice).

Hypothesis 1b: Advisors who experience advice discounting in favor of another advisor will lower their willingness to cooperate further, but not their ongoing effort to give good advice

(as compared to advisors who experience that their own advice is favored over the other advisor's advice).

On another note, disregarding an advisor within a multiple-advisor JAS, as described above, bears some similarities with experimental settings used in ostracism research. According to Williams (2007), selectively excluding one participant from even minimal group interaction, induces feelings of rejection. Furthermore, Williams pointed out that the typical operationalization of ostracism contains repeated interactions that are observed by the participants while enduring exclusion and being ignored. The JAS paradigm usually involves a series of tasks for judge and advisor to cooperate on. Also, in a prototypical JAS, which strictly limits communication between judge and advisor, an adequate weighting of advice is the only way for a judge to reciprocate and show appreciation (c.f., Treffenstädt et al., 2017). Hence, being disregarded in a JAS could be perceived as exclusion from an ongoing cooperation and from the cooperating group. This should lead neglected advisors to evaluate the interaction with their partners more negatively:

Hypothesis 2: Advisors in a multiple advisor JAS who experience advice discounting in favor of another advisor, will evaluate the interaction within the JAS more negatively than advisors who experience a high amount of advice taking.

As in our previous studies (Treffenstädt et al., 2017), we wanted to account for possible moderating effects of perceived expertise differences on the negative consequences that we predicted to result from low amounts of advice taking. Depending on the expertise difference between the two advisors, discounting the focal advisor's recommendations would either be justified from a performance perspective (focal advisor is the less competent advisor), or it would be counterproductive (focal advisor is the more competent advisor). Therefore, we

manipulated perceived advisor expertise in the multiple-advisor JAS. Specifically, through a bogus performance feedback at the beginning of our experiment, we induced perceived expertise differences between our participants and an alleged second advisor in a JAS. Since the multiple-advisor setting described above is different from our previous single-advisor setting, we think it is important to test whether our previous finding that negative reactions were insensitive to perceived expertise differences will replicate under the described circumstances. Furthermore, in case that we might find support for Hypothesis 1a, it would be important to clarify whether perceived expertise differences have a moderating effect on the negative consequences of advice discounting on advisors' effort to give good advice. Again, we decided to formulate a set of competing hypotheses:

Hypothesis 3a: Advisors who experience advice discounting in favor of another advisor will show negative reactions regardless of perceived expertise differences.

Hypothesis 3b: Advisors who experience advice discounting in favor of another advisor will show less negative reactions if the other advisor has more expertise and, hence, advice discounting is justified.

Experiment 1

The present experiment is based upon Experiment 2 reported in Treffenstädt et al. (2017). Participants were asked to work as advisors in a computer-mediated JAS to cooperate on a series of decision tasks. In contrast to our previous experiments, however, participants were told to be part of a JAS consisting of a judge and two advisors instead of a single-advisor JAS.

Our participants were led to believe they would interact with other participants placed in separate rooms, while they were, in fact, interacting with computer-simulated partners.[1] Our simulated judges showed either high amounts of advice utilization by frequently adopting participants' advice as their final decision, or low amounts of advice utilization by more often adopting the advice given by the second advisor in the JAS. Also, as mentioned earlier, we presented bogus performance feedback after a series of training trials to induce perceived expertise differences between the participant and the alleged second advisor in the JAS. Such expertise differences provide a sound rationale for selectively disregarding the advice of the less competent advisor.

Method

Participants and design. A total of 123 participants were acquired via email or phone using the Online Recruitment System for Economic Experiments (ORSEE; Greiner, 2015). The data of one participant had to be excluded because of a technical error occurring during the

[1] Although we avoid deception whenever it is possible, such deception was not evitable in our case. An orthogonal manipulation of our experimental factors requires some of the judges to constantly disregard advice of the clearly more competent of two advisors, and others to constantly place high weights on advice of the less competent of two advisors. It is highly unlikely that we would observe such advice taking behavior with real judges. To gain the highest possible degree of control over the interaction, all actions of both the judge and the second advisor were computer controlled. Of course, all of our participants were fully debriefed about this deception at the end of the experiment.

experimental session. The data of an additional ten participants were excluded, because these participants had, according to our suspicion check, correctly inferred the intentions behind our experimental manipulations. The remaining 113 participants were on average 23.17 (*SD* = 4.02) years old. Sixty-seven participants (59%) were female, and 45 participants (40%) were male. One participant did not state any gender. Participants were randomly assigned to one of four conditions in a 2 (amount of advice taking: high vs. low) x 2 (relative expertise: other advisor more competent vs. participant more competent) between-subjects experimental design.

Task and apparatus. In our decision task, participants were asked to choose the city with the highest daily mean temperature among four mostly unknown cities. To provide participants with an informational basis for their decision, they were given either full or partial access to a set of five cues: the four cities' altitudes above sea level, their positions in longitude, their positions in latitude, their monthly mean humidity rates, and their monthly precipitation rates.

In the experimental test phase, participants had to unlock cues by engaging in a real-effort task, namely by pressing a hand muscle trainer (HMT) for as long as possible. Pressing durations were automatically evaluated to determine the number of cues presented in each trial, with longer pressing durations unlocking more cues. The durations required to obtain a specific number of cues were tailored to participants' individual baseline capabilities. Baseline capabilities were defined as the mean of two initial HMT measurements that we conducted at the beginning of the experiment. Thresholds for unlocking 1 to 5 cues were 30%, 50%, 70%, 80%, and 100% of the baseline duration, respectively. After determining the number of unlocked cue sets, the specific cues were randomly selected from the list of five cues mentioned above. In a pilot study, which is reported in Treffenstädt et al. (2017), we established that participants

understood that the number of available cues was related to the accuracy of advice in the present task. Consequently, participants who are strongly motivated to give accurate advice should engage in unlocking as many cue sets as possible through HMT pressing. The validity of HMT pressing durations as indicators of effort was established in a second pilot study, also reported in Treffenstädt et al. (2017). Therefore, by combining the physical effort task of HMT pressing with the number of presented cue sets, we expected HMT pressing durations to be a valid indicator of participants' effort in giving advice.

Procedure. We invited participants in groups of six. At the beginning of each experimental session, participants were seated together for a few minutes to take note of the other participants being present. This way we ensured our participants believing they would be working with each other. Afterwards, we guided participants to separate rooms with prepared computer workplaces. After starting the experiment by entering their personal information, participants were presented with a bogus randomization procedure to determine their respective roles in the JAS. All participants were told that they had been assigned the role of an advisor, whereas one of their alleged partners had also been assigned the advisor role, and the other partner had been assigned the judge role in the JAS. Participants were then informed about the function of the HMT, performed the two baseline measurements, and received an explanation on the decision task and the five relevant cues. To become acquainted with the task, participants worked on ten trials in a trainings phase followed by the bogus feedback mentioned above. They were told that the performance feedback would be shown to all members of the JAS. The feedback itself was presented as a ranking of both the participant and the second advisor among 100 participants who had allegedly completed the same temperature decision tasks in an earlier study. Participants in feedback conditions in which they were allegedly more competent than the

other advisor learned that they had ranked 24^{th}, while the second advisor had only achieved ranked 85. These ranks were reversed in feedback conditions in which the other advisor was allegedly more competent. There was no information indicating the judge's performance in the training phase. After receiving the bogus feedback, participants were informed about cue acquisition with the HMT in the following test phase of the experiment.

Each trial started with a cue acquisition phase in which participants had to press the HMT. Afterwards participants received the cues they had unlocked and selected the city they thought had the highest daily mean temperature. Subsequently, their selection was allegedly being presented to the judge along with the recommendation of the second advisor. A waiting screen with a random time delay was presented while participants had to wait for their partners' alleged actions. Afterwards, they were shown a feedback screen containing information on both their own and the second advisor's advice, as well as information regarding the judge's initial and final decision. This way, our participants could discern whether a judge maintained his or her initial opinion, or whether the judge adopted any of the two offered advices. The judge's initial decision as well as the advisor's recommendation were randomly selected from the four available cities. If the two advices and the initial decision were identical (which happened in 1% of the trials), the judge would always stay with this decision. In all other cases, the final decisions were computed via different algorithms depending on the experimental condition. In high advice taking conditions, judges adopted participants' advices in 70 percent of the cases, while adopting the second advisors' advices only in 10 percent of the cases. Vice versa, judges in low advice taking conditions adopted participants' advices only in 10 percent of the cases and the second advisors' advices in 70 percent of the cases. In 20 percent of the cases in both advice taking conditions, judges stayed with their own initial decision.

After completing twenty decision task trials, participants were asked to answer several intermediary questions before supposedly continuing to work on five additional trials. Participants were asked to rate their motivation to continue working with their current judge on a seven-point Likert scale (1 - *very low*, 7 - *very high*). They were also asked whether they wished to switch partners to work together with another judge on the upcoming trials. After this question, participants were informed that the main phase of the experiment was already completed, and that they would not have to work on further decision trials. Instead, they were asked to answer some additional questions about the preceding cooperation in the JAS. As a check for the advice taking manipulation, we asked participants on two separate seven-point Likert scales (1 - *very rarely*, 7 - *very often*) to what extent their advice was adopted by the judge as well as to what extend the second advisor's advice was adopted by the judge. To check for the feedback manipulation, participants stated whether they or the other advisor had received the better performance rating after the training trials. Additionally, participants had to rate their own performance as well as the second advisor's performance in the test phase on separate seven-point Likert scales (1 - *very bad*, 7 - *very good*). We also asked participants to rate how integrated they felt into the group (1 - *not at all*, 7 - *very much*), and how much they enjoyed working on the decision tasks (1 - *not at all*, 7 - *very much*). Both measures have been used before to detect effects of exclusion in ostracism studies (Williams, 2007; Zadro, Williams, & Richardson, 2004).

On three additional seven-point Likert scales, we asked participants to rate how much the judge's decisions depended on their own advice (1 - *not at all*, 7 - *very much*), how important it was for them to give accurate advices (1 - *not at all*, 7 - *very important*), and how much effort they exerted in the main phase (1 - *very litte*, 7 - *very high*). After completing these questions,

participants stated their beliefs about the aims of the study (suspicion check). Subsequently, we fully debriefed our participants. After finding out about the computer-simulated partners, participants rated the credibility of both the bogus feedback and the interaction on a seven-point Likert scale (1 - *not believable at all*, 7- *very believable*). Afterwards, we thanked our participants and paid them a fixed fee of €8.

Dependent variables. As the main measure of advice giving effort, we analyzed participants' HMT pressing durations. Additionally, we analyzed participants' subjective effort ratings. Advice quality was measured as the number correct choices over the twenty trials in the experimental main phase. Participants' willingness to cooperate was measured twofold. First, the number of participants wishing to switch partners in each experimental condition was analyzed. Secondly, we evaluated participants' motivation ratings regarding further cooperation within their current group. The subjective ratings of integration into the group and enjoyment were analyzed to account for the evaluation of the JAS interaction.

To account for multiple testing of the same hypothesis, we used Bonferroni correction to adjust the type-one-error for our two measures of effort in the ongoing cooperation, our two measures of willingness to cooperate further, and both measures of evaluation of the JAS interaction (adjusted $\alpha = .025$ for all six variables).

Results

In the following section, we focus on the key variables to our hypotheses. However, the full dataset as well as all test materials can be requested at any time from the first author. Also, detailed results for all analyses of variance and covariance which are reported in this section can be found in Table 1.

Manipulation checks. To test for the success of our advice taking manipulation, we analyzed participants' ratings on how often their own advice had been adopted, as well as how often the second advisor's advice had been adopted. Regarding their own advice, participants in the high advice taking conditions reported higher levels of advice utilization than did participants in the low advice taking conditions (M = 4.94, SD = 0.99 vs. M = 2.37, SD = 1.31), $t(111)$ = 11.84, $p < .001$, d = 2.24. In contrast, participants in high advice taking conditions reported lower levels of advice utilization for the second advisor than did participants in low advice taking conditions (M = 3.87, SD = 1.08 vs. M = 5.53, SD = 1.08), $t(111) = 8.12$, $p < .001$, d = 1.53.

We conducted a 2 (amount of advice taking: high vs. low) x 2 (relative expertise: other advisor more competent vs. participant more competent) ANOVA of participants' ratings regarding how much the final decisions in the JAS depended on their advice. We found a significant effect of advice taking, $F(1, 109) = 22.43$, $p < .001$, $\eta_p^2 = .17$, indicating that participants perceived final decisions to depend less on their advice in the low advice taking conditions than in the high advice taking conditions (M = 2.84, SD = 1.21 vs. M = 3.90, SD = 1.18). There were no significant effects of relative expertise or the interaction of advice taking and relative expertise, both $Fs \leq 1.90$, $ps \geq .170$, $\eta_p^2 s \leq .02$. In summary, all three of the above-mentioned analyses indicate the success of our advice taking manipulation.

Regarding the recall of the performance feedback presented at the beginning of the experiment, 69 percent of our participants correctly remembered whether they or the second advisor had achieved a higher ranking. Participants recalled the correct bogus feedback above chance level, $\chi^2(1, N = 113) = 16.33$, $p < .001$. However, participants in high relative expertise conditions did not rate their own performance in the test phase more favorably than participants

in low relative expertise conditions (M = 4.36, SD = 1.17 vs. M = 3.96, SD = 1.22), $t(111)$ = 1.78, p = .078, d = 0.33. This result could indicate that our manipulation of perceived relative expertise weakened during the test phase.

The two previous analyses did not demonstrate the success of our perceived expertise manipulation as clearly as we expected based on the results we obtained in previous experiments. We would have expected even more of our participants to correctly remember their received expertise feedback, and we would have expected significant differences in performance ratings to follow from this bogus feedback. However, we decided to conduct our manipulation checks for the feedback manipulation at the end of the experiment to avoid any unintended influence on our participants' behavior in the test phase. Therefore, the long interval between the feedback manipulation in the instruction phase and the manipulation checks could explain participants' lower awareness of the bogus feedback and its implications for perceived expertise.

To account for this possibility, we decided to conduct an accompanying study involving a less temporally delayed test of whether our bogus feedback manipulation was successful. We presented 84 participants with the same multiple-advisor JAS scenario we used in Experiment 1. Paralleling the original procedure, participants received bogus feedback regarding their own performance and the performance of a second advisor after completing a ten-trial training phase. In contrast to the original experiment, we subsequently conducted an immediate manipulation check and asked participants to rate their own expertise as well as the other advisor's expertise on a seven point Likert scale (1- *very bad*, 7 - *very good*). Furthermore, we asked participants how much they would expect a judge to utilize both their own advice and the advice of the second advisor. Again, both ratings were given on a seven point Likert scale (1 - *not at all*, 7 - *very*

much). Afterwards, participants were asked to recall the performance feedback they just received.

74 out of 84 participants (88%) correctly recalled the performance feedback they had just been given, which is above chance level, $\chi^2(1, N = 84) = 49.86$, $p < .001$. In addition, we found significant differences between participants who received favorable expertise feedback and participants who received unfavorable expertise feedback on all four rating variables, all $t \geq 3.38$, $p \leq .001$, $d \geq 0.74$. Compared to advisors in low relative expertise conditions, advisors in high relative expertise conditions rated their own expertise to be higher ($M = 3.12$, $SD = 1.35$ vs. $M = 4.19$, $SD = 1.33$) and the other advisor's expertise to be lower ($M = 4.80$, $SD = 1.25$ vs. $M = 3.09$, $SD = 1.46$), they expected their own advices to receive a higher weight ($M = 3.29$, $SD = 1.54$ vs. $M = 4.40$, $SD = 1.45$), and they expected the other advisor's advices to receive a lower weight ($M = 4.93$, $SD = 1.31$ vs. $M = 3.53$, $SD = 1.55$). These results clearly speak to the success of our feedback manipulation prior to the test phase.

Effort in advice giving. We analyzed HMT pressing durations with a 2 (amount of advice taking: high vs. low) x 2 (relative expertise: other advisor more competent vs. participant more competent) ANCOVA with participants' baseline HMT pressing performance as a

covariate. The analysis yielded no significant effects of advice taking, relative expertise or their interaction, all $Fs \leq 1.02$, $ps \geq .315$, $\eta_p^2 s \leq .01$.[2]

A 2x2 ANOVA was computed on participants' subjective effort ratings revealed neither an effect of relative expertise nor the interaction of advice taking and relative expertise, both $Fs < 1$, $ps \geq .471$, $\eta_p^2 s \leq .01$, but there was a significant main effect of advice taking, $F(1, 109) = 7.71$, $p = .006$, $\eta_p^2 = .07$. Participants in low advice taking conditions rated their own effort to be lower than participants in high advice taking conditions ($M = 5.33$, $SD = 1.31$ vs. $M = 5.95$, $SD = 1.03$).

Advice quality. No significant effects were found in an analysis of variance examining the influence of amount of advice taking, relative expertise, and their interaction on advice quality, all $Fs < 1$, $ps \geq .528$, $\eta_p^2 s < .01$. On Average, participants' advice was correct in 36% of all cases, which equals about seven correct recommendations over the course of 20 trials.

Willingness to cooperate further. In a binary logistic regression, advice taking predicted participants' desire switch partners after the 20 trials, $B = 1.14$, ($SE = 0.31$), $OR = 3.12$,

---

[2] In contrast to other experiments in which we used the HMT, we decided to present participants their current pressing time while acquiring cue sets to improve motivation when working with the HMT. Since we used fixed hurdles for cue acquisition, we unintendedly provided participants with a way to maximize their cue acquisition through accurately timing their pressing durations. Indeed, 56 out of 113 participants achieved an average number of four or more acquired cues per trial, indicating a slight ceiling effect in cue acquisition which could have somewhat weakened effects of our experimental manipulations on HMT pressing durations.

OR 95%-CI = [1.72, 5.94], $p < .001$. In low advice taking conditions, 51% of participants expressed a wish to switch partners, compared to only 18% in high advice taking conditions. Neither relative expertise, B = -0.29, (SE = 0.31), OR = 0.74, OR 95%-CI = [0.39, 1.37], p = .348, nor the interaction of advice taking and relative expertise, B = -0.18, (SE = 0.44), OR = 0.83, OR 95%-CI = [0.34, 1.97], p = .681, were significant predictors.

An additional 2x2 ANOVA of the experimental design on participants' rated motivation to continue working with their current partners yielded no significant effects, all $Fs \leq 1.23$, $ps \geq .270$, $\eta_p^2 s \leq .01$.

Evaluation of JAS interaction. We subjected both ratings of integration and enjoyment to an ANOVA of the experimental design. The only significant effect in these analyses was a main effect of advice taking on participants' ratings of how integrated they felt into their group, $F(1, 109) = 18.86$, $p < .001$, $\eta_p^2 = .15$. Participants in high advice taking conditions felt more integrated than those in low advice taking conditions (M = 4.13, SD = 1.53 vs. M = 2.88, SD = 1.48). All other effects were insignificant, all $Fs \leq 1.90$, $ps \geq .171$, $\eta_p^2 s \leq .02$.

Table 1

Summary of computed ANOVAs in Experiment 1

Source	*df*	*F*	*p*	η_p^2
DV 1: HMT Pressing Durations				
Baseline Duration	1	262.16	< .001	.71
Advice Taking (AT)	1	< 1	.783	.00
Feedback (FB)	1	1.02	.315	.01
AT x FB	1	< 1	.990	.00
Residuals	108			
DV 2: Effort Ratings				
Advice Taking (AT)	1	7.71	.006	.07
Feedback (FB)	1	< 1	.471	.00
AT x FB	1	< 1	.546	.00
Residuals	109			
DV 3: Advice Quality				
Advice Taking (AT)	1	< 1	.678	.00
Feedback (FB)	1	< 1	.528	.00
AT x FB	1	< 1	.700	.00
Residuals	109			
DV 4: Motivation Ratings				
Advice Taking (AT)	1	1.23	.270	.01
Feedback (FB)	1	< 1	.716	.00
AT x FB	1	1.01	.316	.01
Residuals	109			
DV 5: Integration Ratings				
Advice Taking (AT)	1	18.86	< .001	.15
Feedback (FB)	1	< 1	.459	.01
AT x FB	1	1.36	.246	.01
Residuals	109			
DV 6: Enjoyment Ratings				
Advice Taking (AT)	1	< 1	.899	.00
Feedback (FB)	1	1.90	.171	.02
AT x FB	1	< 1	.973	.00
Residuals	109			
DV 7: Dependency Ratings				
Advice Taking (AT)	1	22.43	< .001	.17
Feedback (FB)	1	< 1	.757	.00
AT x FB	1	1.90	.170	.02
Residuals	109			

Discussion

In Experiment 1, we were able to replicate three crucial findings from our previous series of experiments (2017). First, in accordance with our previous observations in the single-advisor JAS, advisors in our new multiple-advisor JAS setting showed particular negative reactions to low amounts of advice taking. In particular, we found a reduced willingness to cooperate in the future, as indicated by the increased desire to switch partners in low advice taking conditions. Although we did not find a significant advice taking effect on our second measure of willingness to cooperate further, namely participants' motivation ratings, the results indicate that negative consequences of disregarding advice do not only arise in single-advisor but also in multiple-advisor JAS settings. We will try to obtain further evidence regarding our hypotheses in Experiment 2.

Second, and most notably, we found no effect of advice taking condition on our participants' HMT pressing durations. This means that, when confronted with low amounts of advice taking, advisors did not lower their objective effort in advice giving, although they obviously realized their advice had less influence on the final decisions in the JAS than the other advisor's advice had (as indicated by the dependency ratings). As argued above, it was unclear whether this finding would replicate in a multiple-advisor JAS, because of the principle differences between single-advisor and multiple-advisor settings. Advisors in a single-advisor JAS might feel less offended by advice discounting because it usually means placing a higher weight on the opinion of a superior JAS member. In our experimental setting, however, advice discounting was not attributable to the hierarchical superiority of the highly weighted opinion. In this multiple-advisor setting, advisors might have reacted more negatively. However, the result for our objective effort measure suggests that the presence or absence of negative reactions to

advice discounting is not guided by considerations of hierarchical differences, supporting Hypothesis 1b. In contrast, and in accordance with Hypothesis 1a, we observed an effect of advice taking on our participants' subjective effort ratings. Because this is the very first time we found an effect of advice taking on any effort measure in both our previous and our current experiments, and because no such effect was found in our objective effort measure, we think it is important to test whether it will replicate in Experiment 2.

The third important replication is the absence of advice taking effects on actual advice quality. Although advisors subjectively reported less effort in advice giving, the quality of advice remained constant over all experimental conditions and is, therefore, in accordance with the observations made regarding the objective effort measure (HMT pressing durations). In sum, we think our evidence so far does yield more support for Hypothesis 1b (no negative consequences for the ongoing cooperation) than for Hypothesis 1a (negative consequences also in the ongoing cooperation). Nonetheless, it is crucial to collect additional data to decide between our first set of rivalling hypotheses.

Regarding Hypothesis 2, we found a significant effect of advice taking on participants' integration ratings. Advisors in low advice taking conditions felt less integrated into their JAS group than advisors in high advice taking conditions. In contrast, we did not find a similar effect for participants' enjoyment ratings, which we collected as a second measure of experiencing ostracism. In Experiment 2, we will further investigate advisors' subjective evaluation of the JAS to obtain a clearer pattern of results.

Our second set of competing hypotheses concerns possible moderating effects of perceived expertise differences. Since we observed neither main effects of our feedback manipulation nor any interaction effects of feedback and advice taking, our data so far support

Hypothesis 3b (no moderating effect of relative expertise). However, there is a possible limitation: In the present experiment, we found a lower proportion of participants to correctly remember the feedback they had received than in our previous series of experiments. Also, we did not find a significant effect of the received feedback on participants' expertise ratings regarding the test phase, which would have indicated a successful and lasting manipulation of perceived expertise differences. Although we demonstrated the immediate effects of our feedback manipulation in an accompanying experiment, it is not entirely clear whether we did not observe moderating effects due to a short-lived manipulation, or because of expertise differences which were actually perceived but ignored by our participants in the experimental test phase. This issue will also be addressed in our second experiment.

Experiment 2

In Experiment 2, we want to further investigate our assumption that reactions to advice neglect only have a negative impact on future cooperation between judges and advisors but not on their ongoing cooperation. Also, we want to test for the presence of moderating effects of a demonstrably successful manipulation of relative expertise differences between the two advisors in our multiple-advisor JAS. Therefore, we decided to replicate Experiment 1 with only a small amount of minor changes.

Method

Participants and design. The initial sample of Experiment 2 consisted of 159 participants. One person who did not speak German sufficiently well to understand the instructions had to be excluded. Another 15 participants were excluded after inspection of the suspicion check, because they correctly inferred the design and hypotheses. This left us with a

total of 143 participants, of which 82 (57%) were female. Participants were on average 24.15 (*SD* = 3.88) years old. As in Experiment 1, we used a 2 (amount of advice taking: high vs. low) x 2 (relative expertise: other advisor more competent vs. participant more competent) between-subjects experimental design.

Task and procedure. The procedure of Experiment 2 differed in the following aspects. First, while maintaining the decision task we already used in Experiment 1, we replaced the HMT pressing task with a slider task that we had previously used in Treffenstädt and colleagues (2017; Experiment 4). With the slider task, it is less likely to run into ceiling effects than with the HMT task.[3] Hence, instead of pressing a HMT, our participants received a set of 30 sliders which had to be set to specific positions. Each slider covered a range of values from 0 to 100. Within 30 seconds, participants had to set as many sliders as possible to specific target values which varied for each slider. The number of correctly set sliders was then used in the main trials of Experiment 2 to determine the number of cues that would be presented in the decision task. To unlock 1, 2, 3, 4, or 5 cues, participants had to correctly set 4, 6, 8, 10, or 15 sliders to their specific target values, respectively. The slider task is a finely gradated measure of effort which produces low amounts of intrinsic motivation (Gill & Prowse, 2012), which both are valuable features for investigating effort.

Dependent variables. Except for HMT pressing durations and enjoyment ratings, all dependent variables from Experiment 1 were also collected in Experiment 2. Instead of HMT

---

[3] As mentioned before, our findings in Experiment 1 suggest a slight ceiling effect in cue acquisition, probably because feedback on pressing time allowed participants to effectively pace themselves.

pressing durations, we used the number of sliders that were set to the correct target value as an indicator of effort. We decided to replace enjoyment ratings with ratings regarding judges' fairness. In addition to the previously collected dependent variables, we asked participants to rate how often they would have adopted their own advices had they been assigned the role of the judge in the JAS. Answers were given on a Likert scale ranging from 1 (*very seldom*) to 7 (*very often*). As an additional measure of behavioral consequences resulting from low amounts of advice taking, participants were asked to freely and anonymously assign a bonus payment ranging from zero to three euros to the judge in their JAS. This bonus assignment task offered participants a way to react anonymously to fair or unfair behavior perceived in the preceding JAS interaction. As with Experiment 1, we focused on the key variables to our hypotheses. Therefore, no analyses will be reported for both additional measures.

We used Bonferroni correction to adjust alpha for our two measures of effort in the ongoing cooperation, for both measures of willingness to cooperate further, and for our two measures of advisors' evaluation of the JAS interaction, to account for multiple testing of the same hypothesis (adjusted $\alpha = .025$ for all six variables).

Results

Detailed results for all analyses of variance and covariance which are reported in this section can be found in Table 2.

Manipulation checks. As expected, participants in low advice taking conditions perceived that their advice had been adopted less often than did participants in high advice taking conditions ($M = 2.92$, $SD = 1.25$ vs. $M = 6.13$, $SD = 0.95$), $t(141) = 17.10$, $p < .001$, $d = 2.87$. Also, participants in low advice taking conditions perceived that the judge had adopted the second advisor's advice more often than did participants in high advice taking conditions ($M =$

5.11, *SD* = 1.14 vs. *M* = 2.94, *SD* = 1.15), $t(141) = 11.28$, $p < .001$, $d = 1.89$. Furthermore, a 2x2 ANOVA on the influence of advice taking and relative expertise on dependence ratings yielded a significant main effect of advice taking, $F(1, 139) = 54.01$, $p < .001$, $\eta_p^2 = .28$. As in Experiment 1, participants perceived judges to depend less on their advice in low advice taking conditions than in high advice taking conditions (*M* = 3.68, *SD* = 1.29 vs. *M* = 5.25, *SD* = 1.24). There were no main or interactive effects of relative expertise, both Fs < 1, ps $\geq .352$, η_p^2s $\leq .01$. Hence, our advice taking manipulation was successful.

Regarding the manipulation of relative expertise, 88% of participants correctly remembered whether they had received favorable or unfavorable performance feedback after the training phase, which was significantly above chance level, $\chi^2(1, N = 143) = 83.24$, $p < .001$. Furthermore, at the end of the experiment, participants who received favorable performance feedback rated their own performance in the test phase to be significantly better than participants who received negative feedback (*M* = 4.55, *SD* = 1.00 vs. *M* = 3.71, *SD* = 1.44), $t(141) = 4.06$, $p < .001$, $d = 0.68$, indicating a lasting effect of our feedback manipulation throughout the test phase.

Effort in advice giving. Two ANOVAs on the experimental design with slider task performance and participants' subjective effort ratings as the dependent variables showed no significant effects (Bonferroni-adjusted alpha levels of .025), all Fs ≤ 3.87, ps $\geq .051$, η_p^2s $\leq .03$.

Advice quality. As in Experiment 1, a 2x2 ANOVA revealed no significant differences in advice quality between our experimental conditions, all Fs < 1, ps $\geq .593$, η_p^2s < .01.

Willingness to cooperate further. As our first indicator for participants' willingness to cooperate, we analyzed desire to switch partners after 20 trials in the main phase. A binary

logistic regression with amount of advice taking and relative expertise as well as their interaction as predictors showed that the amount of advice taking had a significant influence on desire to switch partners, B = -1.86, (SE = 0.80), OR = 0.16, OR 95%-CI = [0.02, 0.64], p = .021. Whereas 33% of participants in low advice taking conditions expressed a desire to switch partners, only 6% of participants in high advice taking conditions desired to switch. We found neither a main effect of expertise differences, B = 0.28, (SE = 0.49), OR = 1.32, OR 95%-CI = [0.50, 3.48], p = .572, nor an interaction effect, B = -0.37, (SE = 1.14), OR = 0.69, OR 95%-CI = [0.07, 7.23], p = .744.

Participants' motivation to continue working with their current partners was analyzed in a 2x2 ANOVA, with advice taking and relative expertise as factors. Similar to the intentions to switch partners, we found a significant main effect of advice taking, $F(1, 139) = 8.47$, p = .004, η_p^2 = .06. Participants in high advice taking conditions were more motivated to continue working with their current partners than participants in low advice taking conditions (M = 4.84, SD = 1.41 vs. M = 4.14, SD = 1.46). Once again, there were no main or interactive effects of expertise differences (Bonferroni-adjusted alpha levels of .025), both $Fs \leq 4.41$, $ps \geq .038$, $\eta_p^2 s \leq .03$.

Evaluation of JAS interaction. Both for ratings of integration into the JAS group and for ratings of perceived fairness, a 2x2 ANOVA revealed a significant main effect for advice taking, both _F_s(1,139) $\geq$ 16.84, $p < .001$, $\eta_p^2 \geq .11$. Participants felt less integrated in low advice taking conditions than in high advice taking conditions (M = 3.61, SD = 1.58 vs. M = 5.12, SD = 1.47). Also, participants rated their judges to be less fair in low advice taking conditions than in high advice taking conditions (M = 4.25, SD = 1.43 vs. M = 5.21, SD = 1.32). There were no other significant effects, all $Fs < 1$, $ps \geq .694$, $\eta_p^2 s < .01$.

Table 2

Summary of computed ANOVAs in Experiment 2

Source	*df*	*F*	*p*	η_p^2
DV 1: Slider Task Performance				
Advice Taking (AT)	1	1.37	.244	.01
Feedback (FB)	1	< 1	.550	.00
AT x FB	1	3.87	.051	.03
Residuals	139			
DV 2: Effort Ratings				
Advice Taking (AT)	1	< 1	.706	.00
Feedback (FB)	1	< 1	.354	.01
AT x FB	1	< 1	.444	.00
Residuals	139			
DV 3: Advice Quality				
Advice Taking (AT)	1	< 1	.593	.00
Feedback (FB)	1	< 1	.774	.00
AT x FB	1	< 1	.641	.00
Residuals	139			
DV 4: Motivation Ratings				
Advice Taking (AT)	1	8.47	.004	.06
Feedback (FB)	1	< 1	.875	.00
AT x FB	1	4.41	.038	.03
Residuals	139			
DV 5: Integration Ratings				
Advice Taking (AT)	1	34.05	< .001	.20
Feedback (FB)	1	< 1	.729	.00
AT x FB	1	< 1	.761	.00
Residuals	139			
DV 6: Fairness Ratings				
Advice Taking (AT)	1	16.84	< .001	.11
Feedback (FB)	1	< 1	.954	.00
AT x FB	1	< 1	.694	.00
Residuals	139			
DV 7: Dependency Ratings				
Advice Taking (AT)	1	54.01	< .001	.28
Feedback (FB)	1	< 1	.352	.01
AT x FB	1	< 1	.370	.01
Residuals	139			

Discussion

The results of Experiment 2 revealed a very consistent pattern. We observed negative reactions to low amounts of advice taking, both for the desire to switch partners and for participants' subjective motivation to further work together with the judge. Disregarding advice clearly affected advisors' willingness to cooperate further with a judge who repeatedly neglected their advice. At the same time, participants' effort in the ongoing advice-giving interaction seemed to be unaffected by low advice taking, thereby replicating the exact pattern of results that we observed in our single-advisor JAS studies (2017). Compared to participants in high advice taking conditions, participants in low advice taking conditions did not lower their actual effort in the slider task, neither did they report lower subjective ratings of effort. Also, as in Experiment 1, we did not find any differences in advice quality between our experimental conditions. In summary, the results of Experiment 2 strongly favor Hypothesis 1b (negative consequences of low advice taking only for future cooperation) over Hypothesis 1a (negative consequences for both ongoing and future cooperation).

Regarding Hypothesis 2, we replicated the effect of advice taking on participants' integration ratings that we already observed in Experiment 1, thereby supporting the hypothesis. Additionally, we found a comparable effect of advice taking for participants' fairness ratings, which adds to the similarity of the result patterns we observed in single-advisor and multiple-advisor JAS settings.

Again, perceived expertise differences did not moderate any of the observed effects. Furthermore, we did not observe any main effects of our relative expertise manipulation on our dependent variables. Since we are confident that our expertise manipulation was effective on the grounds of our accompanying experiment to Experiment 1, these results strongly support

Hypothesis 3a (no moderating effect of expertise differences) over Hypothesis 3b (moderating effect of expertise differences). In other words, participants' willingness to cooperate further with a judge who largely neglected their advice was reduced independently of whether neglecting the advice might have been justified based on their alleged performance. Hence judges in a multiple-adivsor JAS should not expect advisors to take mitigating circumstances into account when reacting to advice discounting.

General Discussion

Extending our previous investigation on negative consequences following low amounts of advice taking in single-advisor JAS, we conducted two experiments to test whether our earlier findings would replicate in a multiple-advisor JAS setting. We observed participants' reactions to low amounts of advice taking when serving as advisors in a JAS together with a bogus judge and a second bogus advisor. As with our previous experiments, low amounts of advice taking were either warranted or unwarranted based on perceived differences in expertise within the JAS.

With our present studies, we were able to show that a very distinctive pattern of negative reactions to low advice taking generalized to the multiple-advisor JAS. Specifically, our participants reacted negatively to their advices being disregarded, but the negative consequences were limited to future cooperation only. Again, and most remarkably, following advice neglect, advisors did not lower their effort in the ongoing cooperation despite recognizing the decreased influence of their advice. This was true even when discounting their advice was rationally unwarranted. At the same time, neither participants' reduced willingness to cooperate with the judge in the future nor their feelings of being insufficiently integrated into the interaction and being treated unfairly depended on whether their performance (relative to the other advisor) gave

the judge good reason for preferring the other advisor's recommendations. In the following, we will discuss our results and their implications in more detail.

The consequences of disregarding advice in a multiple-advisor JAS

Harvey and Fischer (1997) suggested that low amounts of advice taking might evoke negative reactions from advisors. This seems to apply to multiple-advisor JAS, as well. Not only should judges be aware that largely neglecting advice in favor of their own opinion might bear negative consequences for future interaction with this advisor. Rather, they should also be aware that favoring one advisor at the expense of the other might lead to similar consequences.

When we transferred both our advice taking manipulation and our bogus feedback manipulation to a multiple-advisor JAS, opposite predictions were plausible regarding whether our previous findings would replicate here. In the new advice discounting situation, our bogus judges favored advice received by another advisor, which cannot be attributed to hierarchical differences in the JAS. In our previous study, these differences presented an alternative explanation for the pattern of results we observed, namely advisors limiting their negative reactions to advice discounting to future cooperation. It is both theoretically and practically remarkable that in the absence of this possible justification for advice neglect, advisors continue to avoid hurting the ongoing cooperation, even though they perceive the judge's behavior to be unfair and evaluate the interaction in the JAS more negatively.

From a theoretical perspective, we took an additional step to shift the focus from the judge towards the advisor in JAS research. To the best of our knowledge, the present experiments constitute the first attempt to investigate a multiple-advisor JAS from the advisor's perspective. Second and most importantly, our results along with previous research (2017) speak to the robustness and generalizability of a particular behavioral pattern in response to low advice

taking, namely reduced willingness to cooperate with the judge in the future, while the effort in the current interaction remains unscathed.

On the practical side, our results indicate that judges both in single-advisor as well as multiple-advisor JAS settings must face a pervasive problem: Since the negative consequences we observed do not arise in the ongoing cooperation, there are no warning signs which could trigger preventive measures such as a change in advice taking levels. By showing no negative reactions in the ongoing cooperation, advisors avoid confrontation and instead try to silently leave the cooperation by switching partners. This quiet exit corresponds to the idea that participants in a social exchange try not to disappoint their partners' expectations (c.f., Dana et al., 2006). Therefore, judges might be surprised to learn that, because of their previously insufficient advice taking, advisors do not wish to cooperate with them in future encounters. Hence, to ensure advisors' willingness to cooperate in the future, judges might have to sufficiently weight received advice across the board (or to acknowledge this advice by other means – which was outside of the focus of the present study), and not only if they see fit. Adhering to the rules of reciprocity might prevent unpleasant surprises in future JAS interactions.

Furthermore, we think that our findings speak to the existence of strong social norms and maybe even an irrevocable psychological contract underlying the judge-advisor interaction. We observed advisors keeping up their advice-giving effort despite their advice being continuously disregarded, and this was true although our data clearly showed that participants knew that their advice had only little influence on the final decisions. Based on motivational models, for example expectancy theory (Vroom, 1964), we would predict lower effort when advisors recognize the diminished expectancy of their advice for the JAS. Additionally, research on

motivation losses in groups has demonstrated that group members' effort depends on their perceived impact on the group outcome (e.g., Kerr & Bruun, 1983). In contrast, the advisors in our Experiments did not react on their perceived dispensability. One explanation for this could be the existence of social norms or a psychological contract prohibiting the reduction of effort.

The moderating role of relative expertise

With respect to relative expertise differences, we feel confident to say that judges in JAS should not expect advisors to consider these as mitigating or aggravating circumstances when reacting to low amounts of advice taking. On the one hand, the negative consequences we observed for the willingness to cooperate further were not ameliorated when perceived expertise differences would have warranted low amounts of advice taking. On the other hand, low amounts of advice taking did not cause negative reactions in the ongoing cooperation even when they were completely unwarranted on the grounds of relative expertise. Instead, matching findings from single-advisor JAS settings (2017), we found no moderating effects of perceived expertise differences on any of our outcome variables.

We think that the absence of moderating effects of perceived expertise places judges in a multiple-advisor JAS in a difficult position. Prior research shows that judges use trimming strategies to exclude outlier advice in a JAS with multiple advisors to increase decision quality (Yaniv & Milyavsky, 2007). Hence, they decide to ignore some pieces of advice in favor of others, just as in our present Experiments. Since our experiments show that disregarding advice is potentially contra productive for future cooperation, judges might face the dilemma of improving accuracy vs. meeting advisors' expectations when deciding about whether and how to weight advice. They might, in fact, feel pressured to place more weight on what they perceive to be low-quality advice to ensure the future cooperation with an advisor. By weighting advice

more equally despite existing accuracy differences, judges might reduce disharmony and feelings of social exclusion which has also been suggested by research on the *equality bias* in collective decision making (Mahmoodi et al., 2015).

Limitations and directions for future research

In our present experiments, we selectively limited our investigation to a very specific situation, namely judges clearly favoring one advisor over the other. Although we outlined why we think that this particular situation is of critical interest for the present research, we recognize that our present experiments can only be a small first step into investigating advisors' negative reactions to low amounts of advice taking in multiple-advisor settings.

One question that could be of particular interest for future research is whether advisors evaluate the weight judges place on their advice in absolute terms or in relation to the weight they place on other advice in the JAS. For example, advisors could encounter a situation in which a judge neglects both their own advice and another advisor's advice. If advisors evaluate advice taking in absolute terms, the difference to the situation we investigated (neglecting advice in favor of another advice) should be of no consequence to advisors' reactions. In contrast, if advisors' reactions in our experiments depended on the perception that another advisor was favored over the focal advisor's own recommendations, a situation in which two advisors are equally neglected might cause less negative reactions.

A second question pertains the restricted social interaction within the JAS setting. We chose a classical JAS setting for our research because such a setting is particularly suited to investigate advice giving and taking without the interference of actual social interaction. However, as Sniezek et al. (2004) argue, ignoring the social context of a JAS limits the insights we can gain regarding real-world advice giving and taking. In a real-world setting, judges would

be able to communicate reasons for disregarding advice. Also, judges could convey respect and gratitude despite discounting an advice, and thereby mitigate the consequences of advice neglect. It stands to reason that we might observe very different reactions to advice neglect in a real-world decision process.

Conclusion

Expanding on a series of studies on single-advisor JAS settings, the current two experiments further demonstrate advisors' negative reactions to low amounts of advice taking in a multiple-advisor JAS. As in the earlier experiments, advisors' reactions were limited to the prospect of future cooperation and did not affect the ongoing interaction. When their advice was neglected (relative to the recommendations given by another advisor), our participants more often desired to switch partners, showed a lower motivation regarding the cooperation in further trials, and they evaluated the interaction within the JAS more negatively. However, advisors did not lower either their subjective or their objective advice giving effort in the ongoing JAS interaction.

Combining the findings reported by Treffenstädt et al. (2017) with the present experiments, a specific pattern of results was established not only with different types of judgment and decision tasks as well as different measures of real-effort in the ongoing cooperation, but even over different types of JAS settings, suggesting substantial robustness of the observed pattern of effects. These results speak to the importance of understanding the advisor's perspective in any type of JAS, since advisors' evaluations of the interaction can have a massive impact on their further cooperation with a judge.

References

Bonaccio, S., & 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*, 127–151.

Dana, J., Cain, D. M., & Dawes, R. M. (2006). What you don't know won't hurt me: Costly (but quiet) exit in dictator games. *Organizational Behavior and Human Decision Processes*, *100*(2), 193–201.

Goldsmith, D. J., & Fitch, K. (1997). The normative context of advice as social support. *Human Communication Research*, *23*(4), 454–476. https://doi.org/10.1111/j.1468-2958.1997.tb00406.x

Greiner, B. (2015). Subject pool recruitment procedures: Organizing experiments with ORSEE. *Journal of the Economic Science Association*, *1*(1), 114–125. https://doi.org/10.1007/s40881-015-0004-4

Harvey, N., & Fischer, I. (1997). Taking advice: Accepting help, improving judgment, and sharing responsibility. *Organizational Behavior and Human Decision Processes*, *70*(2), 117–133. https://doi.org/10.1006/obhd.1997.2697

Kerr, N. L., & Bruun, S. E. (1983). Dispensability of member effort and group motivation losses: Free-rider effects. *Journal of Personality and Social Psychology*, *44*(1), 78–94.

Mahmoodi, A., Bang, D., Olsen, K., Zhao, Y. A., Shi, Z., Broberg, K., … Bahrami, B. (2015). Equality bias impairs collective decision-making across cultures. *Proceedings of the National Academy of Sciences of the United States of America*, *112*(12), 3835–3840. https://doi.org/10.1073/pnas.1421692112

Rousseau, D. M. (1989). Psychological and implied contracts in organizations. *Employee Responsibilities and Rights Journal*, *2*(2), 121–139. https://doi.org/10.1007/BF01384942

Sniezek, J. A., & Buckley, T. (1995). Cueing and cognitive conflict in Judge-Advisor decision making. *Organizational Behavior and Human Decision Processes*, *62*(2), 159–174. https://doi.org/10.1006/obhd.1995.1040

Sniezek, J. A., Schrah, G. E., & Dalal, R. S. (2004). Improving judgement with prepaid expert advice. *Journal of Behavioral Decision Making*, *17*(3), 173–190. https://doi.org/10.1002/bdm.468

Treffenstädt, C., Schultze, T., & Schulz-Hardt, S. (2017). *You better listen to me - consequences of disregarding advice in Judge-Advisor systems*. Manuscript in preparation.

Vroom, V. (1964). *Motivation and work*. New York: Wiley.

Williams, K. D. (2007). Ostracism. *Annual Review of Psychology*, *58*, 425–452. https://doi.org/10.1146/annurev.psych.58.110405.085641

Yaniv, I., & Milyavsky, M. (2007). Using advice from multiple sources to revise and improve judgments. *Organizational Behavior and Human Decision Processes*, *103*(1), 104–120. https://doi.org/10.1016/j.obhdp.2006.05.006

Zadro, L., Williams, K. D., & Richardson, R. (2004). How low can you go? Ostracism by a computer is sufficient to lower self-reported levels of belonging, control, self-esteem, and meaningful existence. *Journal of Experimental Social Psychology*, *40*(4), 560–567. https://doi.org/10.1016/j.jesp.2003.11.006

Appendix C

Curriculum Vitae

Christian Treffenstädt

Curriculum Vitae

Persönliche Daten

Name	Christian Treffenstädt
Geburtstag	14. September 1985
Geburtsort	Aachen

Schule, Studium & Beruflicher Werdegang

seit 2011 **Wissenschaftlicher Mitarbeiter,**
Abteilung für Wirtschafts- und Sozialpsychologie,
Georg-Elias-Müller-Institut für Psychologie, Georg-August-Universität, Göttingen.

- 2014-2016 Projektstelleninhaber im DFG-Projekt SCHU 1279/18-1 "Urteilsprozesse in Judge-Advisor-Systemen mit mehreren Ratgebern"
- Seit 2011 Promotionsstudium im Rahmen der Georg-August University School of Science (GAUSS)

2009-2011 **Studium M.Sc. Psychologie**, *Georg-August-Unversität*, Göttingen.

- Abschlussnote: 1,1 (mit Auszeichnung)
- Titel der Abschlussarbeit: Konsequenzen der Ratschlagsgewichtung für die Arbeit eines Ratgebers im Judge-Advisor-System
- 2010-2011 Studentische Hilfskraft, Abteilung für Wirtschafts- und Sozialpsychologie

2006-2009 **Studium B.Sc. Wirtschaftspsychologie**, *Leuphana Universität*, Lüneburg.

- Abschlussnote: 1,3
- Titel der Abschlussarbeit: Der Zusammenhang von Sicherheitskultur und menschlichen Fehlerkonzepten
- 2009 Absolvierung der Coaching-Ausbildung "Systemisches Coaching"
- 2008-2010 Studentische Hilfskraft,
  Institut für experimentelle Wirtschaftspsychologie (Lünelab)
- 2007-2008 Studentische Hilfskraft,
  UNESCO-Projekt ELAN III - "Einführung in die Nachhaltigkeit"

2005 **Abitur**, *Walddörfer Gymnasium*, Hamburg.

- Abschlussnote: 1,5
- Schwerpunktfächer: Englisch & Mathematik

treffenstaedt@psych.uni-goettingen.de • *ctreffe*

Stipendien & Preise

2014 **Georg-Elias-Preis für die beste Lehre im Jahr 2013 (im Bachelor), 3. Platz**. Seminar Testtheorie

2013 **Georg-Elias-Preis für die beste Lehre im Jahr 2012 (im Bachelor), 3. Platz**. Seminar Wirtschaftspsychologie I: Arbeitssicherheit

2010-2011 **Stipendiat der Studienstiftung des deutschen Volkes**.

Publikationen

Gades, R., Marquardt, N., Treffenstädt, C., and Höger, R. (2011). Sicherheitskultur - ein wirtschaftlicher Erfolgsfaktor? In Schöning, S., Richter, J., and Pape, A., editors, *Kleine und mittlere Unternehmen: Finanz-, Wirtschafts- und andere Krisen. Forschungsbeiträge*, pages 317–326. Lang, Frankfurt am Main.

Marquardt, N., Gerstmeyer, K., Treffenstädt, C., and Gades, R. (2012a). Cognitive Performance Limitations in Operating Rooms. In de Waard, D., Merat, N., Jamson, A. H., Barnard, Y., and Carsten, O. M. J., editors, *Human Factors of Systems and Technology*, pages 317–326. Shaker Publishing, Maastricht, the Netherlands.

Marquardt, N., Gerstmeyer, K., Treffenstädt, C., and Gades-Büttrich, R. (2012b). Analyse menschlicher Fehlerursachen im OP-Saal. In Athanassiou, G., Schreiber-Costa, S., and Sträter, O., editors, *Psychologie der Arbeitssicherheit und Gesundheit: Sichere und gute Arbeit erfolgreich gestalten – Forschung und Umsetzung in die Praxis, 17. Workshop*, pages 105–108, Kröning. Asanger.

Marquardt, N., Gerstmeyer, K., Treffenstädt, C., and Gades-Büttrich, R. (2014). Mentale Beanspruchung bei Ärzten und kritische OP-Ereignisse. In Eigenstetter, M., Kunz, T., Portuné, R., and Trimpop, R., editors, *Psychologie der Arbeitssicherheit und Gesundheit – Psychologie der gesunden Arbeit – 18. Workshop*, pages 409–412, Kröning. Asanger.

Marquardt, N., Treffenstädt, C., Gades-Buettrich, R., and Gerstmeyer, K. (in press). Human error causation in medicine: A medical care perspective. In Scott, G., editor, *Human Error: Preventive Measures, Analysis and Improvement Strategies*, pages X–X. Nova Science Publishers, New York.

Marquardt, N., Treffenstädt, C., Gerstmeyer, K., and Gades-Buettrich, R. (2016). Mental workload and cognitive performance in operating rooms. *International Journal of Psychology Research*, 10(2):209.

Treffenstädt, C., Marquardt, N., Gerstmeyer, K., and Gades-Büttrich, R. (2012). Ursachen menschlicher Fehler im OP-Saal aus Sicht von Pflegekräften. In Athanassiou, G., Schreiber-Costa, S., and Sträter, O., editors, *Psychologie der Arbeitssicherheit und Gesundheit: Sichere und gute Arbeit erfolgreich gestalten – Forschung und Umsetzung in die Praxis, 17. Workshop*, pages 109–112, Kröning. Asanger.

Lehrtätigkeit

SoSe 17	Seminar Wirtschaftspsychologie I: Arbeitspsychologische Basisskills (Bachelor, 2 SWS)
Wise 16/17	Seminar Angewandte Diagnostik: Eignungsdiagnostik (Master, 2 SWS)
Wise 16/17	Modul Arbeitspsychologie: Arbeitssicherheit & Gesundheit (Master, 4 SWS)
SoSe 16	Seminar Wirtschaftspsychologie I: Arbeitspsychologische Basisskills (Bachelor, 2 SWS)
Wise 15/16	Modul Arbeitspsychologie: Arbeitssicherheit & Gesundheit (Master, 4 SWS)
WiSe 14/15	Modul Arbeitspsychologie: Arbeitssicherheit (Master, 4 SWS)
SoSe 14	Modul Gruppenlernen (Master, 4 SWS)
WiSe 13/14	Seminar Testtheorie (Bachelor, 2 SWS)
WiSe 13/14	Modul Arbeitspsychologie: Arbeitssicherheit (Master, 4 SWS)
SoSe 13	Modul Gruppenurteile, Gruppenentscheidungen und Gruppenleistung (Master, 4 SWS)
WiSe 12/13	Seminar Testtheorie (Bachelor, 2 SWS)
SoSe 12	Seminar Wirtschaftspsychologie I: Arbeitspsychologische Basisskills (Bachelor, 2 SWS)
WiSe 11/12	Modul Arbeitspsychologie: Arbeitssicherheit (Master, 4 SWS)
Sonstige	Regelmäßige Durchführung von Mitarbeiterschulungen und Workshops zur Datenanalyse mit R sowie zur Erstellung von Computerexperimenten mit Python

Betreute Abschlussarbeiten

Master

Rumkamp, J. (2016). Der Einfluss von Zielen auf den Rolleneffekt von Käufern in distributiven Verhandlungen

Rothaler, S. (2016). Motivationsgewinne und Motivationsverluste in distributiven Verhandlungen

Jünger, J. (2015). Multiple Ratgeber in Judge-Advisor-Systemen: Wie nutzen wir aggregierte Ratschläge?

Sievers, L. (2015). Die Gewichtung sequentiell präsentierter Informationen

Sänger, M. (2014). Wie wirken sich angemessene und nicht angemessene Gewichtungen auf die Arbeit des Ratgebers im Judge-Advisor System aus?

Bachelor

Mehner, C. (2016). Soziales Lernen im Judge-Advisor-Paradigma

Held, F. (2016). Untersuchung der Leistung bei sequentieller Ratschlagsnutzung

Espe, L. (2016). Überprüfung von Modellen zur sequentiellen Ratschlagsnutzung

Reiber, F. (2014). Die Gewichtung sequenziell präsentierter Ratschläge

Rodewald, H. (2014). Auswirkungen der Bekanntgabe von Initialschätzungen auf Ratschläge und Finalurteile in Judge-Advisor-Systemen

Köhne, A. (2013). Der Einfluss kultureller Normen auf Gesetzestreue

Stahlecker, C. (2013). Der Einfluss der Darbietungsform multipler Ratschläge auf die Urteilsbildung

Sievers, L. (2013). Messung von Leistungsmotivation bei Online-Experimenten

Kuhne, L. (2012). Der Einfluss von Ratschlägen auf Finanzprognosen - Ratschlagsgewichtung und Urteilsqualitä

Dreyer, R. (2012). Die Gewichtung von Ratschlägen im Rahmen einer Multiple-Cue-Judgment Aufgabe

Diplom

Goetzke, B. (2012). Messmethoden mentaler Beanspruchung im Kontext der Arbeitssicherheit

Fachvorträge

2015 The consequences of insufficient advice taking in a judge-advisor system (IAREP – SABE Joint 2015 Conference, Sibiu, RO)

2014 Alfred – A library for rapid experiment development (56. Tagung experimentell arbeitender Psychologen, Gießen)

2014 Auswirkungen niedriger Ratschlagsnutzung auf die Zusammenarbeit von Ratgebern und Entscheidungsträgern (49. Kongress der Deutschen Gesellschaft für Psychologie, Ruhr-Universität Bochum)

2012 Auswirkungen niedriger Ratschlagsnutzung auf die Zusammenarbeit von Ratgebern und Entscheidungsträgern (48. Kongress der Deutschen Gesellschaft für Psychologie, Universität Bielefeld)

2012 Ursachen menschlicher Fehler im OP-Saal aus Sicht von Pflegekräften (17. Workshop Psychologie der Arbeitssicherheit und Gesundheit, Maikammer)