Disentangling the Effects of Advisor Consensus and Advice Proximity

Stella K. Wanzel[1], Thomas Schultze[1,2] , and Stefan Schulz-Hardt[1,2]

[1]Georg-August-University Göttingen,

[2]Leibniz Science Campus Primate Cognition

Author Note

Stella K. Wanzel, Thomas Schultze, and Stefan Schulz-Hardt, Department of Economic and Social Psychology, Georg-August University Göttingen, Göttingen, Germany.

This research was partially financed by a grant from the German Science Foundation (Deutsche Forschungsgemeinschaft, DFG) to the last author, project no. SCHU 1279/18-1.

The data were presented as a research talk at the 50th Conference of the German Society for Psychology in Leipzig, Germany.

Correspondence concerning this article should be addressed to: Stella K. Wanzel, Department of Economic and Social Psychology, Georg-August University Göttingen, Göttingen, Germany.

E-mail: wanzel@psych.uni-goettingen.de

Abstract

When advice comes from interdependent sources (e.g., from advisors who use the same database), less information should be gained as compared to independent advice. On the other hand, since individuals strive for consistency, they should be more confident in consistent compared to conflicting advice, and interdependent advice should be more consistent than independent advice. In a study investigating the differential effects of interdependent vs. independent advice on a judge's accuracy and confidence (Yaniv, Choshen-Hillel & Milyavsky, 2009), advice interdependence was confounded with another variable, namely closeness of the advice to the judge's estimate: Interdependent advice was not only more consistent than independent advice but also closer to the judge's first estimate. The present study aimed at disentangling the effects of consensus and closeness of the advice by adding a third experimental condition in which interdependent (and, hence, consistent) advice was far from the judge's own estimate. We found that, as suggested by Yaniv et al., accuracy gains were indeed a consequence of advisor interdependence. However, in contrast to Yaniv et al.'s conclusions, confidence in the correctness of one's estimates was mostly a function of the advice's proximity to the participants' initial estimations, thereby indicating a social validation effect.

Keywords: advice, proximity, consensus, interdependent, independent

Disentangling the Effects of Advisor Consensus and Advice Proximity

We frequently receive advice, either because we actively seek it, or because others offer it unsolicitedly. Sometimes, we might feel that we need advice from several advisors before making a decision. For example, when making predictions regarding the development of a stock index – in order to know whether and how much to invest in a certain index fund – a person might consult more than one analyst. Another example is environmental legislation. Before deciding whether to regulate carbon dioxide emissions more strictly, a government might consult several experts on climate development. Typically, individuals seek more than one opinion in order to prevent following one piece of bad advice, or to receive sufficient confirmation that they are doing the right thing. Therefore, decision makers will hope for consensus among advisors, that is, the second (and third) advisor should support the first one. If advice is consistent, it is taken as accurate (Kahneman & Tversky, 1973).

Advisor consensus and accuracy

Advice from multiple advisors may not always be equally beneficial, however. An important variable that should affect the value that multiple pieces of advice can have for judgmental accuracy is whether the different advisors are *independent* or *interdependent*. If the financial analysts in the first example are freelancing, intuition holds that their advice may be independent, because the analysts do not influence each other, and they may use different sources of information. If, however, the financial analysts work for the same company (thus sharing a common background and applying similar methods), we should expect their advice to be interdependent. That is, the advice is – to some extent – redundant.

If a person aims to make accurate judgments, multiple pieces of advice are particularly valuable if they are independent. As Yaniv (2004) points out, a quantity estimate consists of the true value, an idiosyncratic bias (i.e., an individual's tendency to under- or overestimate target values), and a random error. The random errors of two individuals are (by definition) independent, but this is not necessarily true for their idiosyncratic biases. Interdependence of opinions means that the idiosyncratic biases are (positively) correlated. This has two consequences: On the one hand, interdependent advice must be – on average – more similar to each other than independent advice. The reason is that for independent opinions the only source of covariation (and thus, similarity) is the true value. In contrast, interdependent opinions co-vary because they share the true value and (part of) the idiosyncratic biases. Thus, interdependent opinions will appear more similar to one another on average – thereby suggesting greater consensus regarding the true value. Second, this greater consensus usually does not go along with greater accuracy – quite the contrary: In the case of independent opinions, aggregation will lead to cancellation of unsystematic (or random) error and idiosyncratic biases, since these biases are mutually independent. In interdependent opinions, only unsystematic errors are reduced in the same fashion, but this is not so for idiosyncratic biases. In essence, the higher the correlation of idiosyncratic biases (i.e., the greater the interdependence of the opinions), the less potential there is for cancellation. Therefore, assuming equal unsystematic errors, independent opinions are, on average, superior to interdependent opinions in terms of accuracy (see also Hogarth, 1978, for a formal analysis).

The question is if decision-makers can recognize the value of independent advice, or whether a desire for advisor consensus pushes them towards interdependent advice. So far, only one published study has addressed this question: Yaniv, Choshen-Hillel, and Milyavsky (2009) investigated how advice from interdependent vs. independent sources affected opinion revision. Their idea was that individuals pay more attention to consensus cues (like the consistency among the advisors) than to interdependence cues (e.g., how the advice was sampled). This should lead to higher confidence due to consistent opinions, but simultaneously less accurate judgments when participants are dealing with interdependent information. According to Yaniv et al., people tend to overlook that consensus can be spurious, defined as “a set of consistent opinions produced by interdependent sources” (p. 561). In contrast, only consensus arising from independent opinions is a valid indicator of the correctness of this advice (cf. 559).

To investigate their hypotheses, Yaniv and colleagues (2009) let participants perform a series of quantitative estimation tasks in a one-factorial within-subjects design. In each trial, participants first estimated the caloric value of a specific food item as accurately as possible. Then, they received advice in the form of the estimates that three participants of a previous study had made for the same food item. Afterwards, participants gave a final – and possibly revised – estimate. Participants also indicated their confidence in the accuracy of their initial and their final estimate. The type of advice (independent or interdependent) was the independent variable. Independent advice was drawn randomly from a pool of 100 previously collected estimates. Interdependent advice, on the other hand, was generated by drawing those three estimates that were closest, 7^{th} closest and 15^{th} closest to participants’ initial estimates. This constraint ensured a high degree of advisor consensus. Depending on the condition participants were either informed that the estimates were selected from those that were closest to their own initial judgment, or that they were randomly drawn.

Confirming their hypotheses, Yaniv et al. (2009) found that participants’ judgments were less accurate when they received interdependent compared to randomly sampled (i.e., independent) advice. Furthermore, participants revised their judgments less often and displayed higher levels of confidence in the interdependent condition. According to the authors, these findings indicate a dissociation of accuracy and confidence, resulting from participants’ preference for advisor consensus and their failure to realize that the underlying dependence caused this spurious consensus.

Advisor consensus or social validation?

Although the authors' interpretation is plausible, their experimental manipulation leaves room for an alternative interpretation: by drawing advice that was close to the participants' own opinion in the interdependent condition, Yaniv et al. (2009) not only manipulated advisor consensus, but also the distance of the advice to the judge's initial estimate. Because of this confound, the findings could just as well be the result of the *proximity* of the interdependent advice to participants' initial estimates. This possibility suggests an alternative explanation for the higher levels of confidence along with lower levels of accuracy in the independent vs. the independent condition, namely social validation. Social validation means that a person's opinion is reinforced by others signaling that they hold similar beliefs or have come to similar conclusions (e.g., Mojzisch, Schulz-Hardt, Kerschreiter, Brodbeck, & Frey, 2008). Schultze, Rakotoarisoa, and Schulz-Hardt (2015) demonstrated that social validation also occurs in the judge advisor system. They found that advice that was more similar to the judge's own initial opinion led to higher levels of confidence and, at the same time, to less frequent adjustments of the initial estimates.

Thus, it seems reasonable to conclude that the high levels of confidence found by Yaniv et al. (2009) are not necessarily due to the high consistency among the advice, but may rather have (at least partially) resulted from the judge being socially validated by the three advisors. Interestingly, one particular result in the Yaniv et al. (2009) study itself already hints at the possibility of social validation effects: participants were more than twice as likely to retain their initial estimate after receiving interdependent advice (which was close to their own initial estimate), as compared to independent (and far) advice (65% vs. 30%). It seems that they were more convinced that their first estimates were accurate and needed no further adjustment when presented with advice similar to these estimates, which is in line with the social validation hypothesis.

The present study aimed to disentangle the effects of the two confounded[1] variables on judges' willingness to heed the advice, their accuracy gains due to taking advice, and increases in confidence. To this end, we replicated the original Yaniv et al. (2009) study and added a third experimental condition in which advisors' estimates were interdependent, but at the same time, notably different from the judge's initial opinion.[2] If the interpretation by Yaniv and colleagues holds true, judges' behavior in the two interdependent advice conditions should differ from that in the independent advisor condition in the same way as in the original study, but the two interdependent conditions should not be different from each other. If, in contrast, participants' behavior in the new condition (interdependent advice dissimilar to the participants' estimations) differs from the original interdependence condition, and more closely resembles behavior in the independence condition, the effects observed by Yaniv and colleagues (2009) were misattributed and, in fact, driven by advice distance rather than interdependence.

---

[1] Interestingly, Yaniv et al. (2009) were aware of their experimental conditions differing in two ways (proximity and advisor consensus), and they also acknowledged that individuals tend to give more weight to opinions that agreed with their own. However, eventually they decided to interpret the results as an effect of advisor consensus alone, while disregarding proximity to the judge's initial estimate.

[2] An even stricter test of our idea would be to implement a fully orthogonal 2x2-design in order to completely disentangle the effects of the two variables. However, the one particular cell that this experimental design would have in comparison to our design (namely a condition with independent advice close to the judge's initial estimate) is practically impossible to realize, as drawing advice that is close to the judge's estimate automatically means that it is interdependent (in terms of the definition of interdependence in the original study).

Method

Participants and design. Eighty[3] participants were recruited via the Online Recruitment System for Economic Experiments (ORSEE, Greiner, 2015). One person was excluded because she correctly guessed the purpose of the study, leaving 79 cases for further analyses. Fifty-two participants were female (66%), mean age was 24.89 years (*SD* = 5.79), and all participants were university students. As in the original study, we used a within-subjects design with the method of advice sampling (independent vs. dependent/far vs. dependent/close) being the independent variable.

Procedure. Our study was similar to that of Yaniv and colleagues (2009) in as many aspects as possible to ensure the comparability of the results. Participants worked on individual computers. They were informed that the experiment had two phases, in each of which they had to estimate the calorie content of various foods. We incentivized participants in the same ways as Yaniv et al.: On the one hand, participants learned that they would receive a bonus of 2 Cents for each estimate that fell in the range of +/- 12% of the true value. On the other hand, they were informed that they could gain an extra bonus of 40 Cents by betting on their answers on 15 out of 30 trials. They received the bonus if their judgment fell in an interval of +/- 12% of the correct answer. In each trial, participants were told how often they had bet so far, and how many

[3] A post-hoc power analysis (G*Power; Faul, Erdfelder, Lang, & Buchner, 2007) revealed a test power of .999 for medium effects (f = .25) when assuming a correlation of r = .50 among the repeated measures. For small to medium-sized effects (f = .18), power was .95. When testing the null-hypothesis with an alpha-level of α = .25 in order to avoid committing a type-II-error, we achieved a test-power of .80 for small effects (f = .10).

remaining bets they had. In the first phase, participants were asked to make initial estimates for the caloric value per 100g/ml of 30 different foods (such as zucchini or butter). As in the original study, we also asked them to provide confidence ratings for their estimates on a scale from 0% (not confident at all) to 100% (completely confident).

In phase two, participants were presented with the same 30 food items. They also received the estimates of three advisors and saw their own estimate from phase one. Phase two involved the three advice-sampling conditions (independent condition, dependent/far condition and dependent/close condition), and each participant passed through ten trials per condition. For each participant, the 30 food items were presented in the same sequence; each condition was presented in 10 trials; the order of conditions, however, was fully randomized for each participant in order to prevent confounding the experimental conditions with specific food items. In each trial, respondents saw a header, accurately informing them about how the three pieces of advice were sampled before they made their final judgments. The independent and the dependent/close condition were identical to those in Yaniv et al. study, whereas the new dependent/far condition was designed in order to disentangle consensus and proximity.

In the independent condition, the header stated that the three advisory estimates were selected randomly from a pool of 112 estimates, which had been made by participants in a previous study. In the dependent/close condition, it said that the advice came from persons who had given estimates close to the participant's opinion. Furthermore, they were informed that the advice was sorted according to its proximity to the judge's original estimate. Similar to the Yaniv et al. study, the computer picked the values that were closest, 7th closest, and 15th closest to the participants' initial estimation. In the dependent/far condition, we first computed an interval of one standard deviation (based on all estimates in the advisor pool for this specific trial). We then selected the piece of advice that was closest to either limit of this interval. This estimate then served as the anchor for drawing the three pieces of advice that were presented to the participant. Similar to the dependent/close condition, all estimates were sorted according to their proximity to this fourth participant's estimate with the selected pieces of advice having positions 1, 7, and 15. The header in the dependent/far condition informed participants that the advice were estimates close to the estimate of a fourth, allegedly randomly selected, participant (therefore, the three pieces of advice were also close to each other). After receiving the advice, participants provided a final, possibly revised, estimate and stated their level of confidence.

Results

In all analyses, one trial out of 30 had to be excluded for 7 persons in the independent condition, since they received advice from only two advisors due to technical problems.
In addition to frequentist tests, we also report Bayes Factors (BF) obtained using the free software JASP (2016). We used the default JZS priors for our analyses of variance (ANOVAs) and *t* tests (see Morey and Rouder, 2011; Rouder, Morey, Speckman, & Province, 2012; Rouder, Speckman, Sun, Morey, & Iverson, 2009). We report the BF_{10} for the Bayesian tests. The BF_{10} is an odds ratio stating how likely H1 is relative to H0 given the observed data. In line with Jeffrey (1961), we interpret a BF_{10} of 3 or greater as positive evidence for the alternative hypothesis, and a BF_{10} smaller then 1/3 as positive evidence in favor of the Null. In some analyses, we tested for differences in the change in certain variables (e.g., changes in accuracy or changes in confidence) using two-factorial ANOVAs with one factor representing time (pre-advice vs. post-advice). In those cases, we only report the BF_{10} for the interaction effect, as this is the effect of interest.[4]

[4] The interpretation of main or interaction effects in two-factorial ANOVAs in JASP does not correspond well to the frequentist significance test. The BF_{10} of an effect represents the relative likelihood of a model containing this effect

Manipulation checks

Advisor Consensus. First, we investigated the degree of advisor consensus in the three conditions by assessing the mean coefficient of variation (CoV). A repeated measures ANOVA showed a main effect for condition, $F(1.73$[5]$, 134.60) = 1309.00$, $p < .001$, $\eta_p^2 = .94$, $BF_{10} = 2.90 \times 10^{128}$. Paired *t*-tests revealed that, as expected, the CoV was higher in the independent condition (M=.61, SD = .09) than in the dependent/far (M = .17, SD = .05), $t(78) = 37.31$, $p < .001$, $d = 4.35$, $BF_{10} = 8.36 \times 10^{47}$ or the dependent/close condition (M =.13, SD = .04), $t(78) = 43.89$, $p < .001$, $d = 5.32$, $BF_{10} = 1.27 \times 10^{53}$. The difference between the two dependent conditions was also significant, $t(78) = 5.01$, $p < .001$, $d = 0.57$, $BF_{10} = 4.74 \times 10^{3}$; compared to the independent condition, however, the CoV was small in both dependent conditions.

Advice Proximity. As a measure of advice proximity, we calculated the absolute difference between judges' initial estimates and the mean of the advice for each trial, and we then averaged these absolute differences for each condition. A repeated measures ANOVA on advice proximity revealed an effect of condition, $F(1.58, 123.48) = 429.17$, $p < .001$, $\eta_p^2 = .85$, $BF_{10} = 4.64 \times 10^{74}$. As expected, proximity was higher in the dependent/close condition (M = 18.37, SD = 19.72) than in the independent condition (M = 113.19, SD = 46.88), $t(78) = 19.38$, $p < .001$, $d = 2.55$, $BF_{10} = 1.94 \times 10^{28}$, or the dependent/far condition (M = 161.23, SD = 27.76), $t(78) = 38.13$, *p*

---

and all subordinate effects to the null-model containing only an intercept. Since the effect of interest in the two-factorial ANOVAs we report is the interaction of time and experimental condition, we chose to report the BF_{10} for the a one-factorial ANOVA on the difference scores. A one-factorial test on the difference scores is equivalent to testing the interaction of time and condition, thus providing the desired BF_{10}.

[5] Whenever the assumption of sphericity was violated, we used the Greenhouse-Geisser correction and report the corrected fractional degrees of freedom.

$< .001$, $d = 4.36$, $BF_{10} = 4.06×10^{48}$. The difference between the independent and the dependent/far condition was also significant[5], $t(78) = -8.02$, $p < .001$, $d = - 0.093$, $BF_{10} = 1.01×10^9$. In summary, the manipulations of interdependence (consensus) and proximity were successful.

Main Analyses.

Accuracy. Table 1 depicts the mean absolute errors of the initial and final estimates, as well as the improvements (in percent) for each condition. A 2 (Phase 1 vs. 2) × 3 (Condition) analysis of variance with repeated measures on the mean absolute errors revealed a main effect of phase, $F(1, 78) = 25.59$, $p < .001$, $\eta_p^2 = .25$, no overall effect of condition, $F(2, 156) = 1.45$, $p = .24$, $\eta_p^2 = .02$, and a significant interaction $F(1.77, 148.97) = 5.66$, $p =.006$, $\eta^2 = .07$, $BF_{10} = 7.54$, indicating that accuracy gains differed between conditions. We analyzed the interaction more closely by investigating the accuracy gains in the three conditions. To this end, we calculated difference scores between the mean absolute errors of participants' initial and final estimates for each condition. Testing the difference scores against zero revealed significant gains in the independent condition ($M = 16.77$), $t(78) = 6.23$, $p < .001$, $d = 0.32$, $BF_{10} = 5.43×10^5$, and in the dependent close condition ($M = 4.64$), $t(78) = 3.05$, $p = .003$, $d = 0.09$, $BF_{10} = 8.80$. In the dependent/far condition, accuracy also increased significantly ($M = 8.52$), $t(78) = 2.15$, $p = .034$[6], $d = 0.16$, $BF_{10} = 1.09$, but since the BF_{10} did not exceed 3, we interpret this latter accuracy gain

---

[5] Since advice in the independent condition was randomly drawn, it can be close to the initial estimate by chance. However, this is impossible in the dependent/far condition.

[6] One person in the dependent/far condition had a particularly high accuracy gain of 114 points. When we exclude this person from the analysis, the gain in accuracy for this condition (M = 7.16) no longer reaches significance, $t(77) = 1.90$, $p = .061$, $d = 0.43$, $BF_{10} = .69$.

with caution. The results are in line with previous research showing that individuals benefit from advice (e.g., Soll & Larrick, 2009; Yaniv, 2004). However, the magnitude of the accuracy gains differed between the conditions: the independent condition yielded significantly larger gains than the dependent/close condition, $t(78) = 4.08$, $p < .001$, $d = -0.62$, $BF_{10} = 188.34$. This result replicates the finding by Yaniv and colleagues (2009). When comparing our new dependent/far condition with the other two, accuracy gains were also greater in the independent than in the dependent/far condition, $t(78) = 2.00$, $p = .049$, $d = -0.27$, $BF_{10} = 0.81$. Technically, while the *p*-value supports the H1, the BF_{10} tends toward the null and is in a range that Bayesian conventions consider inconclusive (Jeffreys, 1961). There was no significant difference between the two dependent conditions, $t(78) = 1.01$, $p = .316$, $d = -0.15$, $BF_{10} = 0.20$. Here, the BF_{10} shows positive evidence of the Null, suggesting that accuracy gains did not differ substantially in the two dependent conditions. In sum, the pattern is not completely unequivocal but, overall, the independent condition seems to be different from the two dependent conditions which, in turn, do not seem to differ reliably from each other. Hence, our findings suggest that, more or less, Yaniv et al. correctly attributed the accuracy findings to consensus. However, this was to be expected, because social validation, as outlined, should mainly affect confidence-related variables.

Frequency of revision. For each condition, we calculated the frequency of revision as the percentage of trials in which participants' initial and final estimates differed (see also Table 1). A repeated measures ANOVA revealed a significant effect of condition, $F(1.65, 128.60) = 92.79$, $p < .001$, $\eta_p^2 = .54$, $BF_{10} = 2.07 \times 10^{24}$. Participants changed their initial judgments less often in the dependent/close condition (44%) than in the dependent/far (82%), $t(78) = 11.29$, $p < .001$, $d = -1.28$, $BF_{10} = 1.24 \times 10^{15}$ or the independent condition (71%), $t(78) = 9.26$, $p < .001$, $d = -1.05$, $BF_{10} = 2.11 \times 10^{11}$. Furthermore, participants changed their initial judgments less often in the independent than in the dependent/ far condition, $t(78) = -4.86$, $p < .001$, $d = -0.55$, $BF_{10} =$

$2.73{\times}10^3$. Note that the values for the independent and the dependent/close condition are very similar to those found by Yaniv and colleagues (2009), which speaks for the comparability of our results to theirs. However, the fact that we observed the highest adjustment rates in the dependent/far condition suggests that reduced willingness to adjust in the dependent/close condition is due to advice proximity rather than interdependence.

Weighting of advice. We analyzed an additional variable not included in the analyses by Yaniv et al. (2009), namely the weight of advice, calculated as the difference between the final and initial estimates, divided by the difference between the mean of the three advisors and the initial estimate. This measure corrects for the systematic difference in advice distance between the three experimental conditions. It is structurally similar to the advice taking measure Harvey and Fischer (1997) developed for the case of single advisors. Weights of advice usually range from 0 to 1. In line with previous studies, we truncated weighting scores greater than 1 to a value of 1, and negative scores to a value of 0 (Schultze et al., 2015; Soll & Larrick, 2009). A repeated measures ANOVA revealed a significant effect of condition, F (1.72, 134.20) = 31.77, $p < .001$, $\eta_p^2 = .29$, $BF_{10} = 1.55{\times}10^9$. Paired t-tests showed greater weights in the dependent/far condition (M = 0.43) than in the independent condition (M = 0.32), $t(78) = -5.31$, $p < .001$, $d = 0.63$, BF_{10} = 9785.36, or the dependent/close condition (M = 0.24), $t(78) = 6.71$, $p < .001$, $d = 0.79$, $BF_{10} = 2.57{\times}10^6$. Furthermore, weighting was greater in the independent than in the dependent/close condition, $t(78) = 3.62$, $p = .001$, $d = 0.41$, $BF_{10} = 40.84$. This pattern supports the findings for revision rate, indicating that differential weighting of advice is a function of advice proximity rather than advisor consensus.

Confidence and betting. We analyzed confidence ratings in a 2 (Phase 1 vs. 2) × 3 (Condition) within-subjects ANOVA. This analysis revealed main effects of phase, $F(1, 78)$ = 19.48, $p < .001$, $\eta^2 = .20$, and condition, $F(1.74, 135.49) = 16.66$, $p < .001$, $\eta_p^2 = .18$, both of which were qualified by an interaction effect, $F(1.76, 136.99) = 28.36$, $p < .001$, $\eta_p^2 = .18$, $BF_{10} = 2.65\times10^8$. Paired *t*-tests revealed that confidence was higher after receiving advice (as compared to before) in the independent condition ($M_{\text{final}} = 51.18$ vs. $M_{\text{initial}} = 46.71$), $t(78) = -3.80$, $p < .001$, $d = -0.24$, $BF_{10} = 75.16$, and in the dependent/close condition ($M_{\text{final}} = 57.10$ vs. $M_{\text{initial}} = 47.80$), $t(78) = -6.61$, $p < .001$, $d = -0.49$, $BF_{10} = 2.64\times10^6$. However, in the dependent/far condition, confidence did not change significantly ($M_{\text{final}} = 49.03$ vs. $M_{\text{initial}} = 47.60$), $t(78) = -1.10$, $p = .274$, $d = -0.08$, $BF_{10} = 0.22$. Note that the BF_{10} for this comparison shows positive evidence in favor of the Null. Additional *t*-tests with confidence gains (i.e., the difference between final and initial confidence) as the dependent variable revealed greater gains in the dependent/close condition ($M = 9.30$) than in the independent condition ($M = 4.47$), $t(78) = -5.62$, $p < .001$, $d = 0.65$, $BF_{10} = 4.89\times10^4$. This finding parallels the corresponding result in the Yaniv et al. (2009) study. In the new dependent/far condition, we found even smaller confidence gains ($M = 1.43$) than in the independent ($M = 4.47$), $t(78) = -2.85$, $p = .006$, $d = -0.32$, $BF_{10} = 5.15$, or the dependent/close condition ($M = 9.30$), $t(78) = -6.53$, $p < .001$, $d = -0.74$, $BF_{10} = 1.85\times10^6$. The strong difference between the dependent/close and the dependent/far condition shows that, contrary to what Yaniv et al. assumed, it is not the interdependence (the consensus) among advisors that drives the confidence findings – rather, it is the proximity of the advice to the participants' initial estimates.

Similar to Yaniv and colleagues (2009), we also analyzed participants' betting behavior. A χ^2-test for equal distribution of frequencies showed a significant difference between conditions, $\chi^2(2) = 95.16$, $p < .001$, $BF_{10} = 4.04\times10^{18}$. Supporting the results for the subjective confidence ratings, and in line with the original study, participants bet more often in the dependent/close (69%) condition than in the independent condition (50%), $\chi^2(1) = 52.02$, $p < .001$, $BF_{10} = 1.56\times10^{18}$. When adding the dependent/far condition, we found that participants bet less in this condition (44%) than in the dependent/close condition (69%), $\chi^2(1) = 88.29$, $p < .001$, $BF_{10} = 4.04 \times 10^{18}$. They also bet less in the dependent/far condition than in the independent condition, but this difference was relatively small at 6 percentage points, $\chi^2(1) = 4.83$, $p = .028$, $BF_{10} = 0.74$. Also, contrary to the p-value, the BF_{10} tends toward the Null, but suggest the results are inconclusive. Consistent with our reasoning above, we chose the frequentist interpretation but – due to the weak evidence – refrain from drawing strong conclusions from the observed difference. Nevertheless, because the dependent/close condition clearly differs from the other two ones, the results suggest that – once more – proximity is the main driving force here.

Discussion

In their study on interdependent versus dependent advice, Yaniv et al. (2009) showed a dissociation between accuracy and confidence. Compared to independent advice, receiving interdependent advice led to less accurate judgments but, paradoxically, also to increased confidence in the accuracy of those judgments. In their experiment, Yaniv et al. focused on advisor consensus, resulting from interdependence among advisors, to explain this dissociation. However, since interdependent advice in their study was not only more consistent, but also closer to the judge's first estimate, one cannot firmly draw the conclusion that advisor consensus is responsible for the effects. It is possible that these effects are caused by the proximity of the advice to the participants' initial estimates. In the present study, we conducted an extended replication of the Yaniv et al. study. By introducing a third condition where advice had similarly high consensus as in the interdependent condition of the original study, but was far from the participants' initial estimates, we disentangled consensus and distance.

First of all, it is important to note that we were able to replicate the main findings of the Yaniv et al. (2009) study. Compared to the independent condition, participants were less accurate after receiving interdependent and close advice, while simultaneously displaying higher levels of confidence and betting more often on their final estimates. Furthermore, they revised their initial estimates less frequently. However, we reach a different conclusion than Yaniv et al. when we take the results of the new dependent/far condition into account. The fact that we found the lowest (and non-significant) confidence gains in the dependent/far condition suggests that proximity, not consensus, is the driving factor here, since otherwise we would have expected roughly the same confidence gains in the two interdependent conditions. Apparently, when the dependent advice is far from one's own estimate (and the distance was greatest in the dependent/far condition), thus not validating the judge's opinion, this results in lower (or even no) confidence gains. This finding is in line with the result of Schultze et al. (2015) that proximity serves as a source of social validation. These conclusions are further supported by participants' reduced willingness to bet on their final estimates in the independent or the dependent/far condition, as compared to the dependent/close condition.

For revision rates, we observed an interesting pattern: Participants revised their estimates most frequently in the dependent/far condition, and least frequently in the dependent/close condition. This is in line with the idea that individuals use consensus as a cue for validity (see Kahneman & Tversky, 1979). When advice is far from one's own opinion, but the advisors are consistent, this results in opinion change. When advice is consistent and close, however, this encourages individuals to keep their opinion. These findings suggest that consensus is always used, but the rule of conduct it signals differs.

In sum, the results so far suggest that greater confidence as well as the lower revision rates after receiving dependent advice in the original study were most likely effects of advice proximity, since our two conditions with dependent advice differed on both confidence measures in the way a social validation account would suggest. Regarding our newly added measure of advice weighting, the results also speak to proximity as the driving force. Supporting the findings for revision rates, participants weighted the advice most in the dependent/far condition and least in the dependent/close condition. Thus, the probability that advice is weighted is higher when it does not resemble individuals' own opinion. The pattern is completely in line with an effect of proximity. However, we can not rule out that proximity in combination with consensus affects our measure in a particular way: when advice is not only far from one's own estimate but also more consistent, this may appear as an especially good reason to adjust to the advice.

With respect to accuracy gains, our results draw a somewhat different picture. While accuracy gains were lower in the two interdependent conditions than in the independent condition, the two interdependent conditions did not differ. In accordance with its lower informational value, interdependent advice leads to less accurate final estimates as compared to independent advice. Descriptively, participants' accuracy improved somewhat more in the dependent/far than in the dependent/close condition, and this pattern makes sense, given that participants retained their initial estimates less often in the latter condition and, accordingly, had less opportunity to improve on their initial estimates. Thus, our data suggest that Yaniv et al. (2009) correctly attributed the lower accuracy gains to the interdependence of advice.

In sum, our results suggest that, while consensus may play a role in explaining some of the findings by Yaniv and colleagues (2009), it is certainly not the only driving force. Quite the contrary: For the majority of variables measured in the Yaniv et al. study (namely confidence, revision rates, and betting), and also for our newly added weighting variable, we showed that proximity rather than consensus drives the effects. Hence, the dissociation between confidence and accuracy Yaniv et al. reported is due to a combination of both proximity and consensus effects, with the former contributing somewhat more than the latter.

Limitations and directions for future research. As an extended replication, our study shares most of the limitations of the original study. First, we only investigated opinions on matters of fact; thus, the results may differ for matters of taste. Second, the operationalization of interdependence was somewhat artificial, in that it was based solely on statistical calculations. Such a procedure is unlikely to map onto everyday life when individuals usually solicit advice. Outside the laboratory, interdependence can arise from advisors belonging to the same group or using shared information. For example, doctors who work in the same hospital or medical center may influence each other when they discuss certain cases or they may underlie a common practice and, therefore, may draw similar conclusions even when they come from different fields. If interdependence arises from realistic circumstances, it might be easier to detect, since individuals are more familiar with these situations. They might intuitively conclude that they do not necessarily receive new information when the advisors all belong to the same institution and pursue the same agenda, or when they have the same data sources. Therefore, an interesting avenue for further research would be to investigate whether the findings still hold true when an ecologically valid operationalization of advisor interdependence is used.

***Conclusion*.** In conclusion, we can state that proximity and consensus are both at play in creating the effects found by Yaniv and colleagues (2009). While social validation can explain why individuals are guided by the closeness of the advice to their own prior opinion, it is rather unclear why it is so difficult for judges to understand the implications of interdependence. Our data suggest that in the absence of proximity (that is, when the advice is far from one's opinion), individuals pay unwarranted attention to (spurious) consensus. Compared to the independent condition, individuals changed their final judgments more often and weighted the advice – which was in fact also far from one's opinion, but more consistent (albeit less beneficial) – more strongly. Individuals apparently feel that there is safety in consensus, even though this consensus is the result of poor sampling methods. Since individuals have a significant disadvantage when they only pay attention to consensus cues, it would be wise to investigate possibilities to train individuals how to use meta-information (like the consensus cues and sampling information) correctly.

Furthermore, our findings tell us that it is also important to take into account the advice's proximity to the advice seeker's own opinion when studying the effects of interdependent vs. independent advice. In real-life situations, advice might be close to the judge's own opinion when the latter also comes from the same institution as the advisors (e.g., a co-worker) or uses the same database, etc. As we have seen, proximity influences individuals' confidence in their own judgements after receiving interdependent vs. independent advice, as well as the extent to which they use the advice. Thus, in such cases it does not suffice to consider only consensus when explaining differential effects in interdependence vs. independence conditions.

References

Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, *39*, 175-191.

Greiner, B. (2015). Subject pool recruitment procedures: organizing experiments with ORSEE. *Journal of the Economic Science Association*, *1*, 114-125.

Hogarth, R. M. (1978). A note on aggregating opinions., *Organizational Behavior and Human Performance 21*, 40-46.

JASP Team (2016). JASP (Version 0.8.0.0)[Computer software].

Jeffreys, H. (1961). *Theory of probability*. Oxford, England: Oxford University Press.

Kahneman, D., & Tversky, A. (1973). On the psychology of prediction. *Psychological Review*, *80*, 237-251.

Mojzisch, A., Schulz-Hardt, S., Kerschreiter, R., Brodbeck, F. C., & Frey, D. (2008). Social validation in group decision-making: Differential effects on the decisional impact of preference-consistent and preference-inconsistent information. *Journal of Experimental Social Psychology, 44*, 1477-1490.

Morey, R. D., & Rouder, J. N. (2011). Bayes factor approaches for testing interval null hypotheses. *Psychological Methods, 16*, 406-419.

Schultze, T., Rakotoarisoa, A., Schulz-Hardt, S. (2015). Effects of distance between initial estimates and advice on advice utilization. *Judgment and Decision Making, 10*, 144-171.

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*, 780-805.

Rouder, J. N., Morey, R. D., Speckman, P. L., & Province, J. M. (2012). Default Bayes factors for ANOVA designs. *Journal of Mathematical Psychology, 56*, 356-374.

Rouder, J. N., Speckman, P. L., Sun, D., Morey, R. D., & Iverson, G. (2009). Bayesian t-Tests for accepting or rejecting the null hypothesis. *Psychonomic Bulletin & Review, 16*, 225-237.

Yaniv, I. (2004). Receiving other people’s advice: Influence and benefit. *Organizational Behavior and Human Decision Processes, 93*, 1-13.

Yaniv, I., Choshen-Hillel, S., & Milyavsky, M. (2009). Spurious consensus and opinion revision: Why might people be more confident in their less accurate judgments? *Journal of Experimental Psychology: Learning, Memory, and Cognition, 35*, 558-563.