Journal of Expertise
2022. Vol. 5(1)
ISSN 2573-2773

Vox Peritorum: Capitalizing on Confidence and Projection to Characterize Expertise

David J. Grüning[1] and Joachim I. Krueger[2]

[1]Psychology Department, Heidelberg University, Germany
[2] Cognitive, Linguistic, and Psychological Sciences, Brown University, USA

Correspondence: David J. Grüning., david.gruening@psychologie.uni-heidelberg.de

Abstract

Theory and research seek to isolate the properties of experts in judgment and decision-making tasks. Confidence in judgment and social projection have emerged as two important meta-judgmental markers, but the joint utility of these two indicators of expertise has not yet been considered. We show that the joint study of individual and contextual differences in confidence and projection offers new opportunities to understand expertise. Our theoretical premise is that experts can solve difficult tasks and do so with high confidence while knowing that few others accomplish this. In a re-analysis of data from Prelec and colleagues (2017) we show that expert judgments are accompanied by higher confidence and less social projection than judgments made by non-experts. Only among experts, are confidence and projection weakly correlated. Moreover, experts align their rate of projection with the difficulty of the judgment task. The present results support a new and integrative approach to the study of experts and expert judgment. We discuss the limitations of the present work and point to future research questions.

Keywords

Expertise, crowdsourcing, confidence, social projection

Introduction

Who's to say who's an expert? ~ Paul Newman

As societies evolve, so do systems of societal decision-making. In democratic societies, the aggregates of individual votes afford and sustain political decision making (Fiorina & Noll, 1979). Beyond political elections, there are myriad contexts for crowdsourcing judgments and decisions, which may serve as testing grounds for innovative methods that integrate available information in many different ways (e.g., Hertwig, 2012; Lang et al., 2016; Ranard et al., 2014; Wang et al., 2012). Of these methods, the majority rule (e.g., Hastie & Kameda, 2005), confidence weighing (e.g., Bahrami et al., 2010; Koriat, 2012ab), and meta-prediction models (e.g., Martinie et al., 2020; Palley & Satopää, 2020; Prelec et al., 2017) have received much research attention. How can these methods be used to gain a deeper understanding of expert judgment? Prelec and colleagues (2017, p. 532) note that some of the most prominent contemporary models "are biased for shallow, lowest common denominator information, at the expense of novel or specialized knowledge that is not widely shared." This concern raises the question of how experts may be characterized so that their unique and valuable knowledge can be used for the common good. We present a novel approach to identify the psychological properties of experts.

A New Way to Identify Expertise

There exist several theoretical ideas concerning the nature of expertise. Most of these models focus on a person's own decisions and one other psychological indicator (see Table 1 below for three prominent examples). Going beyond these models, we combine two critical indicators of expertise in order to capture expert judgment with greater psychological nuance. Specifically, we focus on a participant's degree of confidence in a particular judgment and their level of social projection, that is, their prediction of others' judgments given their own judgments. These two conceptually independent meta-judgments can open a unique window into the expert's mind.

Confidence

To identify decision makers with expertise, many theories rely on the decision maker's level of confidence (e.g., Bahrami et al., 2010; Koriat, 2012ab). Indeed, confidence is diagnostic for expertise in *kind* decision-making environments (Hertwig, 2012). Kind environments offer a close match between the context of information acquisition (learning) and the context of information application (prediction; Hogarth, 2001). In a kind environment, feedback is swift, accurate, and cheap. For crowdsourcing this means that the level of confidence in a prediction is positively correlated with its degree of accuracy. *High* confidence is a valid and valuable cue for a decision maker's expertise in such environments. By contrast, *wicked* environments (Hogarth, 2001) provide noisy feedback, which might also be delayed or hard to obtain. In this environment, confidence does not indicate accuracy, and hence it does not point to expertise.

We take task difficulty as a suitable proxy for the type of environment at hand—and assume that experts know this to be so. We expect confidence to be lower overall for difficult than for easy tasks and to be higher among experts than among nonexperts (e.g., Shanteau, 1992; Sundblad et al., 2009). Whereas the experts' high confidence on difficult tasks is justified, high confidence among the less informed is not. The latter, if confident, commit errors of overconfidence (Heck & Krueger, 2015; Moore, 2020). As shown in Table 1, Zhang (2021) offers an advanced perspective on confidence to increase its diagnosticity for expertise. But this method is costly and rarely applicable. We address the low diagnosticity of confidence in wicked environments (see e.g., Cocozza & Steadman, 1978) by assessing a second prominent indicator of expertise: social projection.

Social Projection

The term "projection" has enjoyed (or suffered) many a usage (Allport, 1924; Freud, 1894; Krueger & Grüning, 2021; Heck & Krueger, 2020; Krueger, 2007; Newman et al., 1997). A simple and robust way of looking at projection is to see it as an instance of inductive reasoning based on a sample of one. As a person has access to their own prediction, they can use this information to predict the predictions of others. A person who predicts "high" will (and should!) predict a greater proportion of "high" predictions than a person who predicts "low" to the extent that no other information about others' predictions is available (see Dawes, 1989; Krueger, 1998, for the Bayesian logic of this claim; Grüning & Krueger, 2021, for an applied scenario).

In general, and especially in highly uncertain environments, greater projection leads to higher accuracy. Empirically, however, it appears that people project too little (Hoch, 1987). Little is known about who or what type of person in what kind of context shows the highest degree of calibration in their level of social projection. This is where the issue of expertise provides traction. As shown in Table 1, some theories of expertise (Martinie et al., 2020; Palley and Satopää, 2020) assume that the most skilled individuals project most accurately (see Faulkner & Corkindale, 2009, for the success of innovations; Martin et al., 2004, for supreme court decisions). That is, experts might show adequate levels of projection where the less skilled tend to *over*project. However, projection is also comparatively low in people who have little or no relevant knowledge in the field of interest and who know so. If projection were the only indicator of expertise, these individuals would be confused with experts. Like confidence, projection is a fallible indicator of expertise when considered in isolation. If high confidence and low projection can be found among nonexperts, these two measures are not sufficient to indicate expertise on their own. But when put together they might be.

Table 1. Overview of models to identify expert individuals from a crowd of tested decision makers

	Variable used	Modeling of Expertise	Reference
Revealed confidence	Confidence	Robust beliefs in light of new information.	Zhang (2021)
Projection accuracy 1	Projection	The difference of own decision and own projection in relation to the sum of all differences of others' decisions and projections.	Martinie et al. (2020)
Projection accuracy 2	Projection	Projection accuracy as the difference between projection and the average estimation.	Palley & Satopää (2020)

A More Comprehensive Approach to Characterizing Experts

In light of the unique limitations of confidence and social projection to characterize experts, we propose a joint consideration of both (see for another combinatory approach, Shanteau et al., 2002). We also propose that expertise is an antidote against the effects of (over-)confidence. Experts, by accurately knowing their capability but also their limit, are not only less prone to overconfidence but, subsequently, also to overgeneralizing their personal experience to others (e.g., projecting their felt confidence). We further assume that only experts solve difficult tasks with confidence while knowing that few others can. Specifically, when prediction tasks are difficult, the combination of high confidence and low projection signals expertise, while none of the three alternative combinations do. When prediction tasks are easy, expertise is not defined, or, if it does exist, it cannot be detected.

Using data collected by Prelec et al. (2017), we evaluate the following hypotheses: Experts have high confidence in their choice (H1.a), and they project less than the average decision-maker (H1.b); confidence and projection are least correlated among experts (H2); Experts have a greater ability than others to tell easy from difficult questions.[1] Hence, experts more accurately adapt their predictions of other people answering correctly to the question's difficulty (H3).

Methods

We analyzed three samples of data collected by Prelec et al. (2017).[2] The samples differ in terms of item topics (i.e., "state capitals of the U.S.," "general knowledge," and "dermatological assessment of lesions") and the sampled respondents. We briefly summarize the methods here. For details, see Prelec et al.'s (2017) original work.

Participants and Procedure

U.S. State Capitals

Prelec et al. recruited 33 participants by mail. Participants were presented with 50 different items in the topical realm of "state capitals of the U.S." For each item, they read the statement "X is the capital of Y," where X was a city and Y a U.S. state.

General Knowledge

39 participants were recruited from Amazon Mechanical Turk (MTurk). Participants were

presented with 80 different items within their topics ranging from history to geography. For each item, respondents read a general knowledge statement (e.g., "A muon has negative electric charge.").

Dermatological Lesions

25 dermatologists were recruited by referral. From 80 pictures of lesions, each participant was presented with a sample of 60 pictures.

For every statement in all three topic domains, participants indicated (1) their assessment of a statement (i.e., true or false) or picture (i.e., benign or malignant), (2) their confidence in their assessment (from 50% to 100%), and (3) their estimate of the percentage of other respondents believing the statement to be true or the picture to show a benign lesion (from 0% to 100%). All three data sets were combined into one data set for analysis. The resulting data set comprised 97 participants making 6,270 decisions on 210 different items.

Results

Setting the Stage

Prelec and colleagues (2017) presented participants with two choice alternatives for each item: A written claim was to be judged as either true or false (i.e., questions about state capitals of the U.S. or general knowledge), or a lesion depicted visually was to be judged as either benign or malignant (i.e., dermatological lesions). For every item a crossing of the statistical status of indicating the majority or the minority response with the response's correctness or incorrectness yields four distinctive groups of responses. These four groups do not comprise individuals but individuals' choices on items. For example, on the statement "Berlin is the capital of Germany," (item 1) a respondent might correctly indicate "yes," while also falsely indicating "yes" in response to the statement "Marseille is the capital of France." (item 2). Respondent choosing "yes" for item 1 would find themselves in the correct majority if most others also chose "yes." Those saying "yes" to item 2, in contrast, would be in the incorrect minority if most others responded correctly by indicating "no." The incorrect majority or the correct minority were categorized in an analogous way.

After removing cases with missing values, individuals' answers to a question that were correct and were chosen by the minority of people were categorized as the correct minority (N = 862). Answers that were correct and were given by the majority of people were categorized as the correct majority (N = 2,783). All incorrect answers to a question chosen by individuals were, too, separated into two groups depending on the answer's popularity among all people ($N_{\text{incorrect minority}}$ = 1,017; $N_{\text{incorrect majority}}$ = 1,595).

As we assume that experts, by definition, respond correctly, we can expect expert choices to cluster in the correct majority and the correct minority. We can also expect the relative proportion of experts to be higher in the correct minority than in the correct majority. Hence, expert choices in the available sample seem to be best—though not exclusively—captured by the correct minority choices. For the study of what characterizes experts, a focus on the correct minority is thus necessary.

Differences In Confidence and Projection

We tested our first hypothesis by comparing the average confidence ratings (H1.a) and projection rates (H1.b) among the four groups. As the average perceived item difficulty is expected to be different in majority and minority groups, a comparison of the average confidence between the correct minority and the two majority groups is uninformative. We therefore compared the mean judgments of the two minority groups with each other and separately compared the two majority groups with each other.[3] The means of the four groups are displayed in Figure 1. As expected, the confidence level displayed in the correct minority (M = .71) was higher than confidence in the incorrect minority (M = .68), with ΔM = .03, $t(1877) = -4.12$, $p < .001$, showing a small effect, d = -.19. There was also a significant difference in confidence between the incorrect (M = .73) and the correct majorities (M = .79), with a ΔM = .06, $t(4376) = -10.67$, $p < .001$, and a small effect, d = -.34. This is consistent with our hypothesis. Experts—who reside in the correct groups—have higher confidence than other decision makers. However, we focus our analysis on the correct minority because that group carries a higher proportion of experts.

As shown in Figure 2, projection was weaker in the correct minority (M = .52) than in the incorrect minority (M = .55), ΔM = .03, $t(1877) = 3.51$, p = .003, with a small effect, d = .16. There was no significant difference, $t < 1$, p = .480, d = .02, between the projection rates in the correct majority (M = .60) and the incorrect majority (M = .60). Again, the comparison of the correct minority with the majority groups provides little information, as the projection rate is influenced by the perceived difficulty of an item, which differs between minority and majority. The critical pattern was observed.

In summary, judgments by the correct minority were characterized by comparatively high confidence and low social projection (H1). This pattern, we submit, is the footprint of expertise. As shown in *Table A1* in the Appendix, nonparametric Mann-Whitney U tests were consistent with the parametric analyses for confidence level and projection rate.

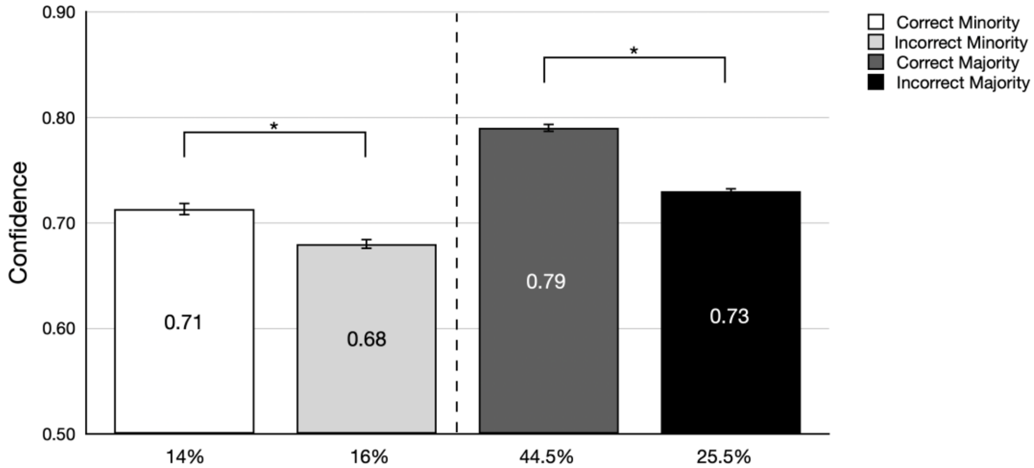

Figure 1. Judgmental confidence: Means and standard errors for the four groups' confidence levels

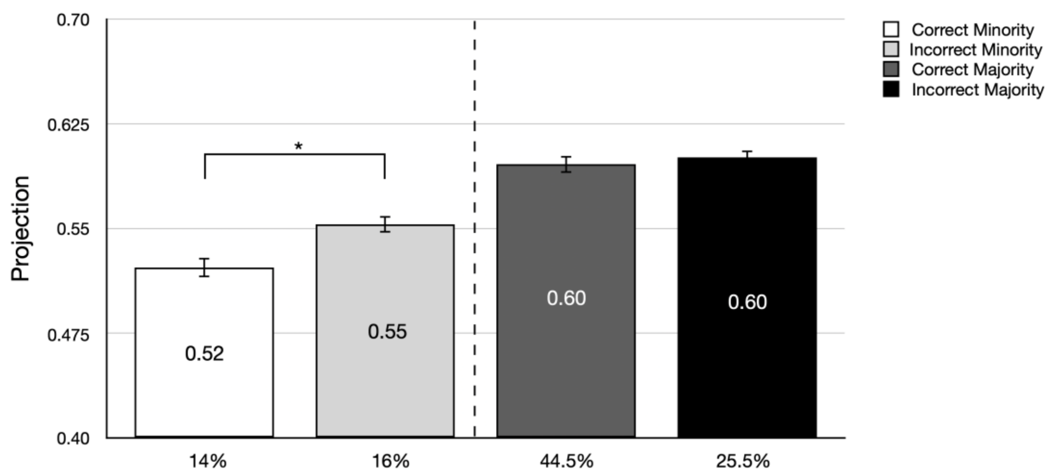

Figure 2. Social projection: Means and standard errors for the four groups' projection rates

Correlational Analyses

We expected a lower correlation between confidence and projection within the correct minority (i.e., among the experts) than within the other three groups (H2). As shown in *Figure A2* in the Appendix, the correlation in the correct minority was small, $r(860) = .10$, but significant, $p = .003$, across item topic groups. After Fisher's Z transformations, we found that the association between projection and confidence in the correct minority was significantly lower than in the incorrect minority, $r(1015) = .345$, $Z = -5.59$, $p < .001$, the incorrect majority, $r(1593) = .326$, $Z = -5.62$, $p < .001$, and the correct majority, $r(2781) = .302$, $Z = -5.52$, $p < .001$, respectively. For one of the three domains of questions, namely, the questions about state capitals, the correlation of confidence and projection in the correct minority was near zero and nonsignificant, $r = -.035$, $p = .579$, while the relationship remained highly significant in the other three groups.

Sensitivity to Item Difficulty

We predicted that experts' projections are more sensitive to item difficulty than others' (H3). We therefore compared the differences in projection rates among the four groups, contrasting easy with difficult items. The percentage of respondents choosing the correct option for an item represented the item's difficulty. Difficulty is greatest when no one chooses the correct option (0%), and lowest when everyone chooses the correct option (100%). By definition, the correct minority and the incorrect majority only contain items that were answered correctly by the minority (< 50%). The incorrect minority and the correct majority only contain items that were answered correctly by the majority (> 50%).[4] To test for projection sensitivity to item difficulty in the former two groups, we compared choices for items that less than 11% of the participants answered correctly (highest difficulty in the two groups) and for items which were answered correctly by 39 to 49 percent (lowest difficulty). For the latter two groups, we compared choices for items which were answered correctly by 51 and 61 percent of participants (high difficulty) and for items which more than 89 percent of the participants answered correctly (low difficulty). For the very difficult items, the correct minority only contained a few cases ($N = 14$).

While the results of Frequentist and Bayesian analyses have a high matching probability in general (Krueger & Heck, 2018), the Bayesian tradition has one important advantage. It allows to judge if, in such cases of low group size, our data were informative enough to evaluate our hypothesis (Lee & Wagenmakers, 2014). Due to missing prior knowledge we, according to suggestions (Jeffreys, 1948; Rouder, 2009), use the default Cauchy distribution centered around zero with the ability to test bidirectionally as our prior for t-testing of the independent samples. We found strong evidence, $BF_{10} = 14.60$, indicating that the average rate of projection in the correct minority adapted to difficult ($M = .35$, $N = 14$) and easy items ($M = .51$, $N = 446$; see Figure 3, a). That is, expert respondents (i.e., the correct minority) were sensitive to the item's difficulty when predicting other respondents' judgments. For the incorrect majority there was extremely strong evidence, $BF_{10} = 179.83$, in the opposite direction (Figure 3, b). Projection was stronger for high ($M = .67$, $N = 141$) than for low ($M = .59$, $N = 540$) item difficulty. There was moderate evidence, $BF_{10} = .20$, for the interpretation that there is no difference in projection between high ($M = .56$, $N = 329$) and low ($M = .58$, $N = 41$) item difficulty in the incorrect minority (Figure 3, c). That is, projection in the incorrect minority – that is among nonexperts – was *not* sensitive to item difficulty. In contrast, there was extreme evidence, $BF_{10} = 2357^{e+39}$, for the claim that projection is weaker for difficult ($M = .50$, $N = 421$) than for easy items ($M = .72$, $N = 845$) in the correct majority (Figure 3, d). In short, projection in the correct minority, as in the correct majority, was sensitive to the difficulty of an item.

In summary, there was clear evidence for projection sensitivity to item difficulty in both groups that included experts, the correct minority and the correct majority. There was no evidence for such sensitivity in the two groups that lack experts, the incorrect minority and the incorrect majority.

Additionally, Figure A3 in the Appendix shows the confidence ratings of high and low item difficulty in the four different groups. Only in the expert group (i.e., correct minority) did the item difficulty not change the participants' confidence level significantly.

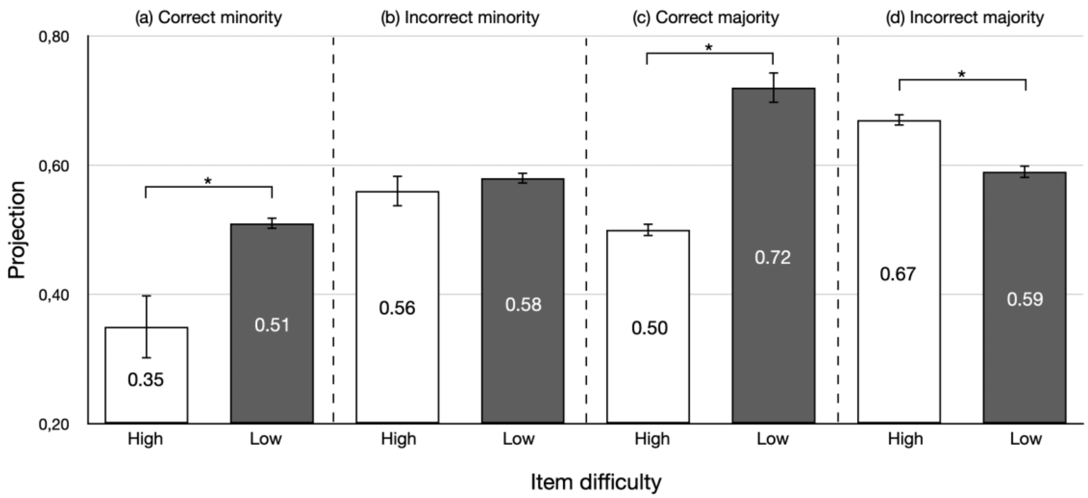

Figure 3. Projection on difficult and easy items difficulties across the four groups

Note. Item difficulty in correct minority and incorrect majority: high (under 11% of people answered correctly) vs. low (39% to 49% answered correctly). Item difficulties in incorrect minority and correct majority: high (51% to 61% answered correctly) vs. low (over 89% answered correctly).

Discussion

The aim of this study was to test three hypotheses offered by an integrative psychological approach to expertise. Analysis of data collected by Prelec and colleagues (2017) corroborated these hypotheses. Expert choices are associated with high confidence and a comparatively low level of social projection (H1). Confidence and projection are conceptually and statistically discriminable. As predicted, these two indicators are less positively correlated among experts than among nonexperts (H2). Experts are most sensitive to an item difficulty, and they project their own responses to difficult items less strongly to others (H3). Together, these findings suggest a more nuanced view of expertise as they point to the relevance of two variables, confidence and projection, that are commonly used in studies of expertise but whose joint effects have thus far not been explored.

We wish to remind the reader that we analyzed expert choices (i.e., correct minority) not expert individuals. Respondents answered choice problems, and they indicated their confidence in their own response as well as their (projective) predictions of the responses of others. Thus, the results do not reflect on the characteristics of individuals. Although the present data sets by Prelec and colleagues (2017) could be rearranged to the individual level, the resulting samples sizes (*N*s < 40) would lack the statistical power necessary for confident conclusions. Future research is needed to extend the present analyses by treating individual respondents as units of analysis. Such an endeavor would mark an additional, necessary test for the expert theory we propose in the present paper. It should be noted that the present analyses address the description of expertise and do not include its prediction. A next crucial step is to build and test a model that formalizes confidence, social projection, and their relation mathematically to predict expertise in decision-makers.

Recognition of the roles of kind and wicked environments is critical when asking questions of expertise. We addressed this issue with the analysis of items of different difficulties. However, there are many ways to proceed to ascertain what this theory can and cannot tell us in kind versus wicked environments. Comparing these types of environment is critical because testing a theory mostly in kind or in wicked environments leads to a myopic view on this theory (Grüning & Krueger, 2021). Specifically, a focus on wicked environments risks an undervaluation of human rationality. A better test of any assumption or intervention (Grüning et al., in press) addresses both environments. We call for such testing for the present theory.

Future research might also explore how best to use identified experts for decision-making. The literature on the wisdom of the crowd (Krueger & Chen, 2014; Surowiecki, 2004) shows that aggregated choices tend to be most accurate, and particularly so if aggregation is limited to known experts (see e.g., Mannes et al., 2014; for an applied example, Kattan et al., 2015). Averaging multiple expert opinions instead of listening to a single expert's voice presents the advantage that individual cognitive biases—which afflict experts, too—are partly cancelled out by the crowd. Following the tradeoff between expertise and diversity described by Olsson and Loveday (2015), the wisdom of the *expert crowd* might advance decision-making by integrating important insights from expertise research into the common *wisdom-of-the-crowd* approach.

As noted by a reviewer, future studies should test the present model in different decision-making scenarios that go beyond the simple dichotomy of static statements like true vs. false and benevolent vs. malignant. Asking participants to make decisions for action vs. inaction may be one possible endeavor to advance the model's applicability. Beyond manipulating the type of a decision, it may prove interesting, too, to vary the format of a decision. For instance, the model could be tested on choices between numeric values like percentages or scores.

As societies evolve, so do systems of societal decision-making. Experts have always had a position of influence in this process, and rightly so. However, with the accelerating speed in which the informational ecology grows more complex, it gets increasingly complicated to identify experts for certain domains, especially for domains that have just been born. We have offered here a first step towards a more complex theory of multiple indicators of expertise. This approach blends in the recent understanding that identifying expertise is a complex endeavor that calls for a well-fledged and detailed psychological theory—and welcomes efforts to refine indicators with more psychological detail (e.g., Zhang, 2021; Sziklai, 2018). We suggest that accuracy can be improved by relating multiple indicators of expertise to each other in a formalized manner. Our combination of indicators considers confidence and social projection as expertise markers. Experts are not just confident in their decision but also more sensible in projecting what others might decide. Their confidence has less influence on their projective behavior and this behavior is sensitive to the environment's wickedness (i.e., difficulty of items).

Endnotes

1. We make no claim about moderately difficult items but only about the two extreme groups, low vs. high difficulty for a group of decision-makers.
2. We thank Drazen Prelec for sharing the data with us.
3. Tests for normality revealed significant skews. As the differences between the means and the medians were small, we focus on the results of parametric tests, noting the nonparametric tests did not change any of our conclusions. We revisit the results of the appropriate non-parametric tests later.
4. Note that no question resulted in 50% of people answering correctly and incorrectly.

Acknowledgements

We thank Drazen Prelec for sharing the data with us and two anonymous reviewers for their constructive feedback.

Authors' Declarations

The authors declare that there are no personal or financial conflicts of interest regarding the research in this article.

The authors declare that they conducted the research reported in this article in accordance with the Ethical Principles of the Journal of Expertise.

The authors declare that they are not able to make the dataset publicly available but are able to provide it upon request.

ORCID iDs

David J. Grüning
https://orcid.org/0000-0002-9274-5477

Joachim I. Krueger
https://orcid.org/0000-0001-9607-1695

References

Allport, F. H. (1924). *Social psychology*. Riverside.

Bahrami, B., Olsen, K., Latham, P. E., Roepstorff, A., Rees, G., & Frith, C. D. (2010). Optimally interacting minds. *Science, 329*(5995), 1081-1085. https://doi.org/10.1126/science.1185718

Cocozza, J. J., & Steadman, H. J. (1978). Prediction in psychiatry: An example of misplaced confidence in experts. *Social Problems*, *25*(3), 265-276.

Dawes, R. M. (1989). Statistical criteria for establishing a truly false consensus effect. *Journal of Experimental Social Psychology, 25*(1), 1-17. https://doi.org/10.1016/0022-1031(89)90036-X

Faulkner, M., & Corkindale, D. R. (2009). Are experts better than potential users in predicting the uptake of an innovation? Extending the use of the Juster Scale. *Technological Forecasting and Social Change*, *76*(7), 910-916. https://doi.org/10.1016/j.techfore.2008.11.001

Fiorina, M. P., & Noll, R. G. (1979). Majority rule models and legislative elections. *The Journal of Politics*, *41*(4), 1081-1104. https://doi.org/10.2307/2129734

Freud, S. (1894). *The neuro-psychoses of defence*. Hogarth Press.

Grüning, D. J., & Krueger, J. I. (2021). Strategic thinking: A random walk into the rabbit hole. *Collabra: Psychology*. https://doi.org/10.1525/collabra.24921

Grüning, D. J., & Krueger, J. I. (2021). The experience heuristic. *Journal of Expertise*, *4*(3).

Grüning, D. J., Panizza, F., & Lorenz-Spreen, P. (2022, January 25). The importance of informative interventions in a wicked environment. https://doi.org/10.31234/osf.io/azsbn

Hastie, R., & Kameda, T. (2005). The robust beauty of majority rules in group decisions. *Psychological Review*, *112*(2), 494-508. https://doi.org/10.1037/0033-295X.112.2.494

Heck, P. R., & Krueger, J. I. (2015). Self-enhancement diminished. *Journal of Experimental Psychology: General, 144*, 1003-1020. http://dx.doi.org/10.1037/xge0000105

Heck, P. R., & Krueger, J. I. (2020). Self-enhancement error motivates social projection. *Social Cognition*, *38*(5), 489-522. https://doi.org/10.1521/soco.2020.38.5.489

Hertwig, R. (2012). Tapping into the wisdom of the crowd—with confidence. *Science*, *336*(6079), 303-304. https://doi.org/10.1126/science.1221403

Hoch, S. J. (1987). Perceived consensus and predictive accuracy: The pros and cons of projection. *Journal of Personality and Social Psychology*, *53*(2), 221-234. https://doi.org/10.1037/0022-3514.53.2.221

Hogarth, R. M. (2001). *Educating intuition*. University of Chicago Press.

Jeffreys, H. *(1948), Theory of probability (2nd ed.)*. Oxford University Press.

Kattan, M. W., O'Rourke, C., Yu, C., & Chagin, K. (2016). The wisdom of crowds of doctors: Their average predictions outperform their individual ones. *Medical Decision Making, 36*(4), 536-540. https://doi.org/10.1177/0272989X15581615

Koriat, A. (2012a). The self-consistency model of subjective confidence. *Psychological Review*, *119*(1), 80-113. https://doi.org/10.1037/a0025648

Koriat, A. (2012b). When are two heads better than one and why?. *Science*, *336*(6079), 360-362. https://doi.org/10.1126/science.1216549

Krueger, J. (1998). On the perception of social consensus. *Advances in experimental social psychology*, *30*, 163-240. https://doi.org/10.1016/S0065-2601(08)60384-6

Krueger, J. I. (2007). From social projection to social behaviour. *European Review of Social Psychology, 18*, 1-35. https://doi:10.1080/10463280701284645

Krueger, J. I., & Chen, L. J. (2014). The first cut is the deepest: Effects of social projection and dialectical bootstrapping on judgmental accuracy. *Social Cognition*, *32*(4), 315-336. https://doi.org/10.1521/soco.2014.32.4.315

Krueger, J. I., & Grüning, D. J. (2021). Psychological perversities and populism. In J. P. Forgas & K. Fiedler (Eds.). *The social psychology of populism [The 22nd Sydney Symposium on Social Psychology]*. Taylor & Francis.

Krueger, J. I., & Heck, P. R. (2018). Testing significance testing. *Collabra: Psychology*, *4*(1), 11. https://doi.org/10.1525/collabra.108

Lang, M., Bharadwaj, N., & Di Benedetto, C. A. (2016). How crowdsourcing improves prediction of market-oriented outcomes. *Journal of Business Research*, *69*(10), 4168-4176. https://doi.org/10.1016/j.jbusres.2016.03.020

Lee, M. D., & Wagenmakers, E. J. (2014). *Bayesian cognitive modeling: A practical course*. Cambridge University Press.

Mannes, A. E., Soll, J. B., & Larrick, R. P. (2014). The wisdom of select crowds. *Journal of Personality and Social Psychology*, *107*(2), 276-299. https://doi.org/10.1037/a0036677

Martin, A. D., Quinn, K. M., Ruger, T. W., & Kim, P. T. (2004). Competing approaches to predicting supreme court decision making. *Perspectives on Politics*, *2*(4), 761-767. doi:10.1017/S1537592704040502

Martinie, M., Wilkening, T., & Howe, P. D. (2020). Using meta-predictions to identify experts in the crowd when past performance is unknown. *PloS one*, *15*(4), e0232058. https://doi.org/10.1371/journal.pone.0232058

Moore, D. A. (2020). Perfectly confident: How to calibrate your decisions wisely. Harper Collins.

Nemet, G. F., Anadon, L. D., & Verdolini, E. (2017). Quantifying the effects of expert selection and elicitation design on experts' confidence in their judgments about future energy technologies. *Risk Analysis*, *37*(2), 315-330. https://doi.org/10.1111/risa.12604

Newman, L. S., Duff, K. J., & Baumeister, R. F. (1997). A new look at defensive projection: Thought suppression, accessibility, and biased person perception. *Journal of Personality and Social Psychology*, *72*(5), 980-1001. https://doi.org/10.1037/0022-3514.72.5.980

Olsson, H., & Loveday, J. (2015). A comparison of small crowd selection methods. In D. C. Noelle, R. Dale, A. S. Warlaumont, J. Yoshimi, T. Matlock, C. D. Jennings, & P. P. Maglio (Eds.). *Proceedings of the 37th Annual Meeting of the Cognitive Science Society*. Cognitive Science Society.

Palley, A., & Satopää, V. (2020). Boosting the wisdom of crowds within a single judgment problem: Selective averaging based on peer predictions. https://dx.doi.org/10.2139/ssrn.3504286

Prelec, D., Seung, H. & McCoy, J. A solution to the single-question crowd wisdom problem. *Nature 541***,** 532-535 (2017). https://doi.org/10.1038/nature21054

Ranard, B. L., Ha, Y. P., Meisel, Z. F., Asch, D. A., Hill, S. S., Becker, L. B., Seymour, A. K., & Merchant, R. M. (2014). Crowdsourcing—harnessing the masses to advance health and medicine, a systematic review. *Journal of General Internal Medicine*, *29*(1), 187–203. https://doi.org/10.1007/s11606-013-2536-8

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

Shanteau, J. (1992). The psychology of experts: An alternative view. In G. Wright & F. Bolger (Eds.), *Expertise and decision* *support* (pp. 11–23). Plenum Press. https://doi.org/10.1007/978-0-585-34290-0_2

Shanteau, J., Weiss, D. J., Thomas, R. P., & Pounds, J. C. (2002). Performance-based assessment of expertise: How to decide if someone is an expert or not. *European Journal of Operational Research*, *136*(2), 253-263. https://doi.org/10.1016/S0377-2217(01)00113-8

Sundblad, E. L., Biel, A., & Gärling, T. (2009). Knowledge and confidence in knowledge about climate change among experts, journalists, politicians, and laypersons. *Environment and Behavior*, *41*(2), 281-302.

Surowiecki, J. (2004). *The wisdom of crowds*. Random House.

Sziklai, B. (2018). How to identify experts in a community?. *International Journal of Game Theory*, *47*, 155-173. https://doi.org/10.1007/s00182-017-0582-x

Wang, J., Zhao, P., & Hoi, S. C. (2012). Exact soft confidence-weighted learning. *Proceedings of the 29th International Conference on Machine Learning*, *8*(1), 1-32. https://doi.org/10.1145/2932193

Zhang, Y. (2021). Identify experts through revealed confidence: An application to wisdom of crowds. https://dx.doi.org/10.2139/ssrn.3739192

Received: 9 October 2021
Revision received: 12 March 2022
Accepted: 16 March 2022

Appendix

Table A1. Nonparametric analyses via Mann-Whitney U test for the differences in confidence level and projection rate between the four groups

Variable	Comparison	*U*-value	*p*-value	$r_{r\text{-}bi}$
Confidence level	Incorrect vs. correct minority	394993.50	< .001	-.10
	Incorrect vs. correct majority	1790000000.00	< .001	-.19
Projection rate	Incorrect vs. correct minority	472088.00	.003	.08
	Incorrect vs. correct minority	2234000.00	.710	< .01

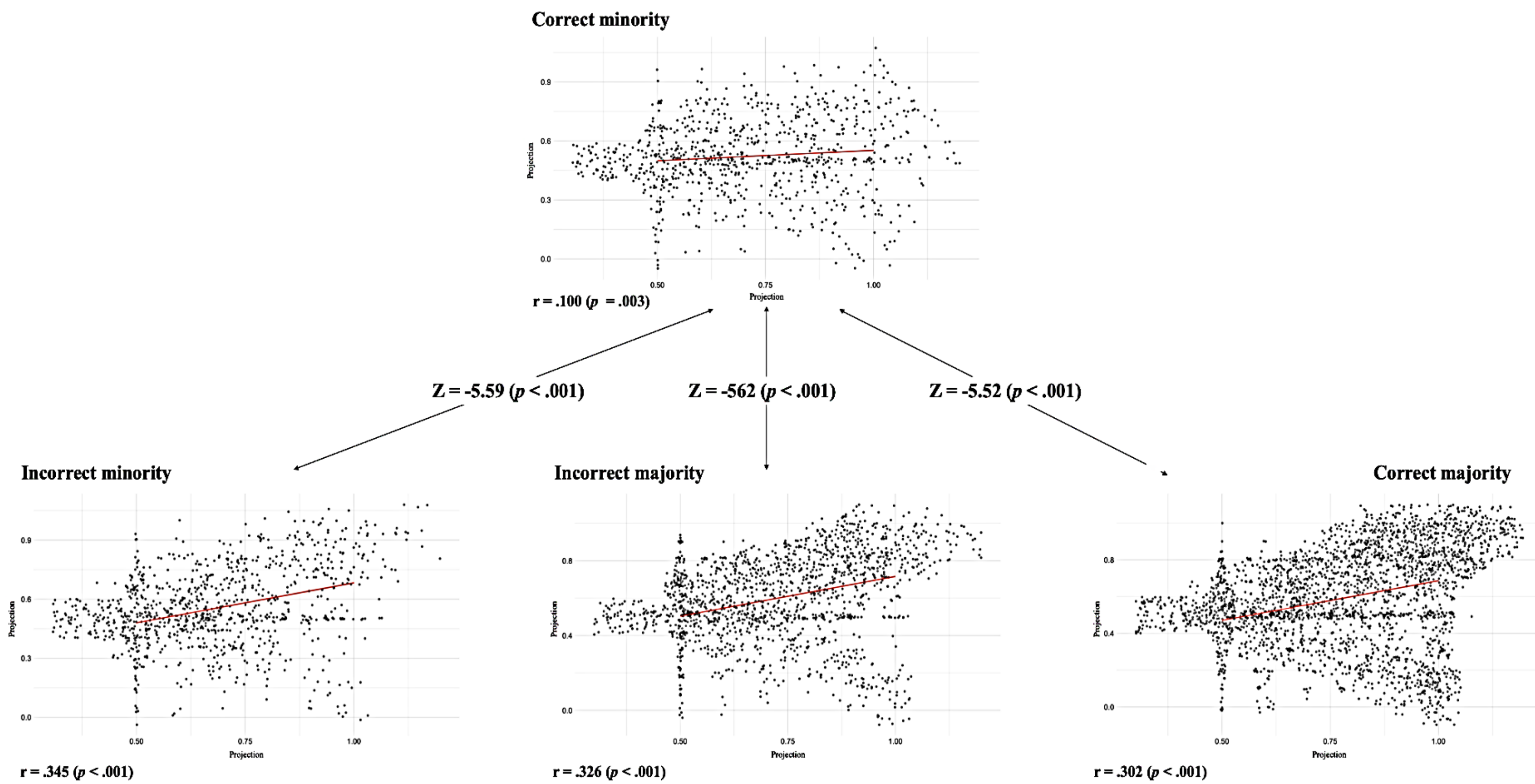

Figure A2. Visualization of correlations between projection (y-axis) and confidence (x-axis) across all four groups (correct minority, incorrect minority, incorrect majority, and correct majority); and tests for significant differences between the correlation in the correct minority and the other three groups' correlations

Note. $N_{\text{correct minority}} = 862$; $N_{\text{incorrect minority}} = 1017$; $N_{\text{incorrect majority}} = 1595$; $N_{\text{correct majority}} = 2783$.

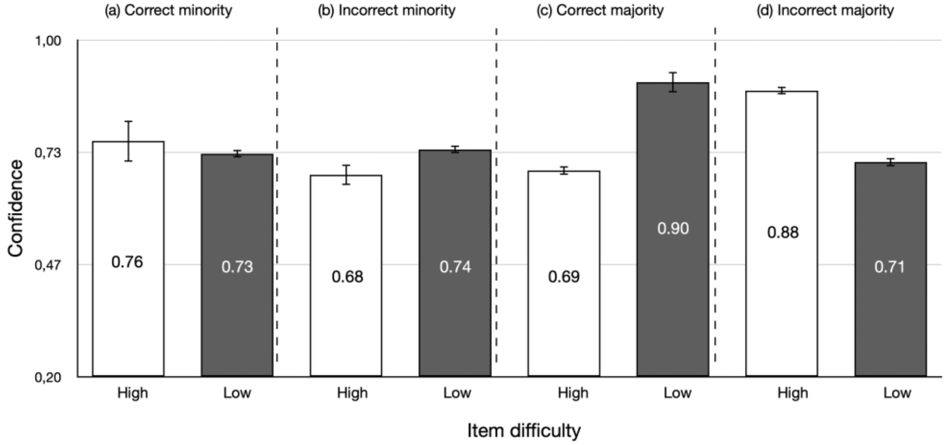

Figure A3. Visualization of confidence differences between item difficulties across the four groups

Note. Item difficulties in correct minority and incorrect majority: high (under 11% of people answered correctly) vs. low (39% to 49% answered correctly). Item difficulties in incorrect minority and correct majority: high (51% to 61% answered correctly) vs. low (over 89% answered correctly).