Dunning-Kruger effect in climate change science communication

C. Braat

TUDelft

Dunning-Kruger effect in climate change science communication

by

C. Braat

as additional master thesis
at the Delft University of Technology.

Student number:	4313488	
Project duration:	September 19, 2019 – April 1, 2020	
Thesis committee:	dr. ir. R.W. Hut,	TU Delft, supervisor
	dr. ir. M.W. Ertsen,	TU Delft
	dr. S. Illingworth,	Manchester Metropolitan University

Contents

Abstract

The knowledge deficit model assumes that people make more informed decisions when they are presented with more information. This model is often used in communication strategies while it has received quite some critique from science communicators and is not well supported by social science theories. One of these critiques comes from the observed Dunning-Kruger effect, where individuals unskilled in a certain area do not possess the skills to realize their incompetence. This effect has not been researched extensively yet in relation to climate change science communication and is the topic of this paper. By means of an online questionnaire (316 respondents), respondent's knowledge and estimated knowledge on climate change is tested. The Dunning-Kruger effect has been detected for this group and suggests a critical re-evaluation of the knowledge deficit model, however additional research is necessary. An initial analysis into the influence of factors like age, gender and highest completed education level on actual and estimated scores and the discrepancy between these is also done to provide leads for further research.

1
Introduction

Many science communication researchers have pointed out that using the knowledge deficit model, assuming people would make informed decisions by merely providing more knowledge, is not well supported by social science theories [11, 18, 20, 25]. However, this model is still in use in many fields (which Simis et al. [25] links to less positive attitude towards the public, amongst other factors), while empirical research poses promising suggestions for more effective science communication [11, 12, 25, 28].

One of the arguments against using the knowledge deficit model is the existence of the Dunning-Kruger effect. This effect shows that people who are unskilled in a certain field do not posses the skills to realize they are in fact incompetent [16]. This false overconfidence leads to wrong decision making and sometimes even rejecting of expert opinions. This effect has been found in many areas, from people's beliefs concerning vaccinations and GMO food to people's confidence in emotional intelligence, logic and grammar [6, 16, 19]. Several popular science articles also suggest a link between climate change denial and the Dunning-Kruger effect [1, 4, 24, 26].

Other recent articles and studies suggest that climate change denial can not be merely attributed to the Dunning-Kruger effect as it is a highly polarized topic [3, 6, 8, 14, 15]. Fernbach et al. [6] showed a significant Dunning-Kruger effect related to GMO food and found a similar but not significant trend when testing for climate change, stating that the polarization of the topic could be the cause. Kahan [14] claims that beliefs individuals have reflect the influence of two different goals people want to achieve with their belief, being to gain access to scientific knowledge but also to be part of a community strengthening their identity. Kahan reveals that, by splitting these two influences, there is 'in fact little disagreement among culturally diverse citizens on what science knows about climate change', shifting the source of the controversy and inaction to cultural-identity motives [14]. Taddicken et al. [27] found varying knowledge and confidence levels over the five different climate change knowledge dimensions they measured, indicating that these might need different communication strategies of which the knowledge deficit model is suggested to be one. Another recent study shows that a German test group was able to quite accurately estimate their performance in climate literacy for statements that were true and underestimated their performance for statements that were false [7].

This paper aims to study the existence of the Dunning-Kruger effect and the influence of background factors in relation to climate change science with the goal of informing more effective communication. It does so by using an online questionnaire based on an established climate change literacy quiz developed with EU funding [2] after which the participants are asked to estimate their performance. The estimated performance is then compared with the actual performance for the whole group as well as for subgroups depending on political affiliation and other factors such as education level and gender.

2 Method

2.1. Hypothesis

The main hypothesis of this paper is inspired by the original first hypothesis used by Dunning and Kruger, which states 'Incompetent individuals, compared with their more competent peers, will dramatically overestimate their ability and performance relative to objective criteria' [16]. As we focus on climate change communication here, our hypothesis reads:

'Individuals incompetent in the field of climate change, compared with their more competent peers, will dramatically overestimate their ability and performance relative to objective criteria.'

Related to this, we aim to find what the effect size for the Dunning-Kruger (DK) effect is when related to climate change science and what factors influence the actual knowledge, the estimated knowledge and the discrepancy between these two when a significant DK effect is found.

2.2. Questionnaire

To research this hypothesis, an online questionnaire will be used (see Appendix A). The first part of the survey measures the respondent's competence in climate science literacy, based on the established questionnaire on climate literacy funded by the EU [2]. The last two questions of this questionnaire were found to be framed in a way that invites a certain answer and were excluded. Other questions were slightly rephrased based on feedback gained in a pilot survey (10 respondents) to prevent unnecessary misinterpretation.

After the climate change knowledge questions, participants are asked to estimate their test performance relative to others. This is done in two ways. Participants are asked to estimate how they rank amongst '100 random people' from their country on a scale from 0 ('0: I performed in the lowest 10%') to 10 ('10: I performed in the top 10%'). They were also asked to estimate how many questions they had correct from 0 ('0: I answered 0 questions correctly') to 10 ('10: I answered all questions correctly'). Additionally, personal background information was collected, including 'Do you consider yourself an earth/climate scientist?' and 'Are you an English native speaker?'. Questions on political background asked the participants to rank themselves on a progressive-conservative and left-right scale (each with 7 options). Participants were also asked whether they voted during the last national election and on which party. To check whether participants are aware of the Dunning-Kruger effect, a list of concepts is presented as last question. The respondent is asked to select the concepts that are familiar to him/her.

The questionnaire is spread online mostly using Twitter and LinkedIn. The aim is to spread the survey to as many people as possible focusing on including different backgrounds, however the authors did expect certain biases from the way the data was collected in the area of highest level of education and political ideology.

2.3. Statistical analysis

The actual scores are compared with the performance estimates of the respondents using two tailed T-tests. The first analysis takes all respondents into account. In order to reduce the regression to the mean effect, the actual and estimated scores are compared per quartile. This is done for both percentile (how one scored relatively) and number (how one scored objectively) scores. The expected effect size is 0.2 and a significance level (α) of 0.01 is taken. To obtain a power of 0.9, our sample size should be 376. Other sample sizes for the same significance and power can be found per effect size in Appendix B, Figure 5.1.

The second analysis consists of breaking the available data down in different subgroups related to age, gender, highest completed education level and political ideology, amongst other factors (for a complete overview, see the questionnaire in Appendix A). The subgroups (sg) are compared with their corresponding not-subgroups (nsg) on estimated and actual scores and the discrepancies between these scores, again using four quartiles.

It should be noted here that as the same test is done to quite a high number of subsets of the same data, the probability of finding significant results while they are not true (type I error) increases. However undesirable, the multiple comparisons problem easily affects these types of research, even when this is not done intentionally [9]. Additionally, some subgroups in this study might be correlated with each other which separate T-test don't take into account. This analysis should be seen as a first exploration in the possible (relative) importance of factors influencing the Dunning-Kruger effect when related to climate change. Therefore, possibly insignificant results that show to be significant are not filtered out even though they could be due to chance. A more elaborate analysis can be deployed taking these considerations into account however this is outside the scope of this research.

To prevent fabricated results, the second analysis was only carried out after a significant result was found in the first analysis and the subgroups chosen were already decided upon when the questionnaire was set up. Only subgroups that turned out to have insufficient responses were left out or adjusted after the questionnaire results were in. To increase transparency and to give an indication of the influence of the multiple comparisons problem, significance for the straightforward but conservative Bonferroni corrected α level of $0.01/2/20 = 0.00025$ for a two-sided T-test is indicated separately in the results.

3
Results

3.1. Response

316 participants responded to the questionnaire. The effect size was found to be 0.85, so with alpha = 0.01 and a desired power of 0.9, the required sample size is 49 respondents (see Appendix B, Figure 5.2).

3.2. Respondents

When comparing the average real score (5.73/10) with the average estimated number score (6.96/10) of the whole group, the p-value is < 0.001. The average real percentile score (50 by definition) compared to the estimated average percentile score (73) gives again a p-value of < 0.001.

To be able to compare the low scoring group with the rest of the respondents, the total group of respondents is split in four quartiles depending on their actual score. The p-values for the number scores of these four quartiles are < 0.001 for the first three quartiles and 0.016 ($> \alpha$) for the last quartile. For the percentile scores, the p-values are < 0.001, < 0.001, 0.152 and 0.658 respectively, making the difference between the estimated and real percentile scores of the first two quartiles significant for $\alpha = 0.01$.

Figure 3.1 compares the actual score with the estimated number score per quartile. The error bars represent one standard deviation. All quartiles on average overestimate themselves, however it is clear that respondents in the lowest quartile overestimate themselves most. Figure 3.1 also shows that the actual scores are not so spread out with the first quartile already having an average score of 4.02/10 and the last quartile 7.36/10.

Figure 3.2 presents the same comparison but then for percentile scores, showing that there is almost no difference in the estimated percentile score over the four percentiles. The last quartile on average even underestimates itself, an effect seen in earlier Dunning-Kruger studies [16, 21, 23].

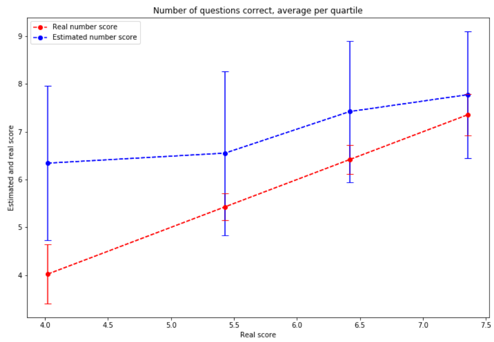

Figure 3.1: All respondents number score

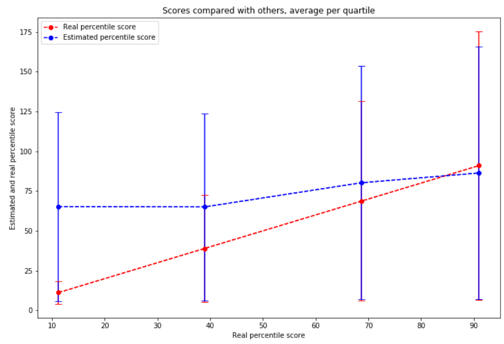

Figure 3.2: All respondents percentile score

3.3. Subgroups

In order to get a better picture of what influences the differences seen in real and estimated scores, different subgroups are compared. Only subgroups with a larger sample size than the required sample size of 49 respondents are discussed. As an extreme majority of 271 out of 310 respondents (85.8%) stated to have voted in the previous national elections, this factor was left out of consideration. Analysis based on the party that was last voted on in national elections was also left out as 41 respondents were not willing to share this information, forgot or did not specify down to one party and another 26 were not eligible to vote. The remaining 249 respondents voted on parties from different countries which lead to the biggest party being GroenLinks with merely 60 respondents. Only 7 respondents stated to be aware of the existence of the climate literacy quiz funded by the EU which again made an analysis on this factor not possible.

Appendix C gives a visual overview of the subgroup analysis, where the subgroups scores are compared to the average real and estimated scores. All subgroups show curves resembling the Dunning-Kruger effect with a significant difference between the real and estimated score for all first quartiles (both number and percentile score types). Only UK respondents do not have significantly different real and estimated scores in the second quartile (Q2), all other subgroups show a significant difference in Q2 as well (both score types). Q4 shows for most subgroups a sharp decrease in overestimation with underestimations for 3/19 number score subgroups and 14/19 percentile score subgroups. The number score behavior for the group that has a Bachelor degree as highest attained education level is the biggest exception to this observation, showing no relevant decrease in overestimation from Q2 to Q4. A complete overview of p-values for the whole group and all subgroups can be found in Appendix D.

Another analysis was done to check for significantly different estimate and real scores between the subgroups (sg) and their complement not-subgroups (nsg). This chapter describes the subgroups and gives an overview of the most notable sg-nsg trends. A complete overview of all p-values referred to can be found in Appendix E. The chapter ends with an elaboration on the influence of the Bonferroni corrected α on both subgroup analyses.

Age 3 of the 316 responses on age were invalid. The remaining 313 respondents are split in three age groups: ≤ 25 years, 25-45 years and > 45 years. The first group (88 respondents) shows over all quartiles a lower actual score than the average (only significantly different from nsg on average for percentile scores). The estimated scores per quartile are not significantly different from the nsg for both number as percentile scores. The second group (152 respondents) shows significantly different estimate scores compared to the nsg for the forth quartile and overall for percentile scores. The third group (73 respondents) shows higher scores for the first quartile and close to nsg overestimation for the number scores. For the percentile scores however, the estimated score are much lower and vary much less than average over the four quartiles, from 55.1 in the second quartile to 72.2 in the last (compared to 65.0 in the second quartile to 86.4 in the last on average). The percentile score for the fourth quartile shows even a significant underestimation.

Gender 313 respondents filled in the 'To which gender identity do you most identify?' question, of which 174 (55.6%) identified as male, 134 identified as female and 5 as non-binary. This last group is not taken into account for this comparison as the sample size is too small. Male respondents on average have a significantly higher actual score compared to the nsg. Male respondents also overestimate themselves more than the average, most notably in the first quartile (however only significant on average). Female respondents on the other hand overestimate themselves significantly less than the average, most notably in the last quartile. For both score types they even underestimate themselves in Q4, showing a significant difference to the nsg. These findings are in line with existing literature on the effect of gender on confidence in science knowledge [5, 13].

Country of residence All 316 respondents answered the question 'In which country do you currently live?'. 180 (57%) of the respondents currently lives in the Netherlands, 65 in the UK. Other countries had too few responses to include in the analysis. Residents of The Netherlands on average have a smaller spread of actual scores and a significantly higher average for actual scores. They overestimate themselves more than the average, only significant on average. Interestingly, UK residents overestimate themselves much less than the average and even underestimate themselves in the last quartile for both the number and percentile scores (only significant for percentile scores). The second quartile in both number and percentile scores shows similar estimate and real scores, making the first quartile stand out even more when it comes to overestimation. The sg actual scores are very close to the nsg actual scores, however their estimate scores are significantly different for both types on average and for number scores in each separate quartile as well except Q3.

Native English speaker 313 respondents answered the question 'Are you an English native speaker?'. A similar trend to UK residents can be found for native speakers, as from the 102 native English speakers (32.6%), 61 are currently residing in the UK. For both score types, the average estimate scores as well as the real scores were significantly lower for native English speakers, for not native English speakers it was significantly higher. The native English speaking group shows some interesting differences between the third quartile of the number and percentile scores, however subtle. They on average overestimate themselves a bit more than average when it comes to guessing the amount of questions they answered correctly, while they on average overestimate themselves less than average when they compare themselves with others from their country.

Highest attained degree 7 respondents preferred not to answer the question on highest attained degree of education, leaving 309 respondents. The group having primary or secondary education as their highest attained degree is too small to include in this analysis (1 and 30 respondents respectively). This already shows that the education levels found here do not represent actual society well, as 87.9% of the respondents finished at least their bachelors degree. The 105 respondents that had a Bachelor as their highest attained degree scored and estimated close to average. The forth quartile does not show a decrease in overestimation from the second and third quartile for the number score, showing no increase in self knowledge as real scores increase. Also for the percentile score, the forth quartile overestimates itself compared to the average, however the effect is less profound here as the real and estimated scores are not significantly different. The 116 respondents with a Master degree show a bit an opposite trend, mostly overestimating themselves more than average except the forth quartile. For both score types and for both Bachelor and Master degree holders, the estimated and real scores are not significantly different from the average. The 57 Doctorate or Professional school degree holders on average score better, however only significantly for number scores. Except for the second quartile, they overestimate themselves more than the nsg, again most notably for both first quartiles and Q3 and Q4 for the number scores (all significant).

Field of study 271 respondents answered the question 'In what field did you study?'. 17 people answered with a more specific study field than asked for and were not taken into account. 25 respondents answered Arts and Humanities and another 36 answered Social sciences. The only two groups big enough to analyse were Engineering (101 + 1 Engineering and Social sciences, 37.6%) and Natural sciences (92 + 1 Natural sciences and Arts and Humanities), showing another bias in the total group of respondents towards these two fields of study. The Engineering group shows close to average behavior, however with small differences in estimated score over all quartiles for number scores and significantly for Q2 and Q3 for percentile scores. They score significantly higher than their nsg. The Natural science group shows remarkably similar behavior with the Doctorate and Professional school degree holders. The 93 respondents in this group constitute of 36 Doctorate or Professional school degree holders, 36 Master degree holders and 21 Bachelor degree holders. On average, they score significantly better than their nsg but they overestimate themselves even more, significantly for all quartiles in the number scores and Q1 and Q4 in the percentile scores.

Earth/climate scientist Only respondents that answered Engineering, Natural sciences or something outside the four defined categories in the previous question were led on to this question. From these 210, 208 respondents answered the question 'Do you consider yourself an earth/climate scientist?'. Unfortunately, only 38 said yes, making the group too small to analyse. The other 170 said no (81.7%), showing significantly different actual scores for both score types. The number scores were only significantly higher on average, while the percentile score is significantly higher over all quartiles except Q4.

Political ideology Different analyses were done to get a picture of the influence of political ideology on overestimation behavior. Leaving moderate (M) voters out, everyone that answered being slightly progressive, progressive and extremely progressive are seen as progressive (P). From these, leaving the slightly progressive group gives the extra progressive (EP) group. The conservative group (C), extra conservative group (EC), right group (R) and extra right group (ER) were too small to be considered. Combining the progressive-conservative (P-C) and left-right (L-R) scale to make PL, PC, CL and CR groups was also not possible with this data set as only the PL group was big enough (195 against 20, 0 and 16 respectively). Another big bias in the data thus exists here. The overwhelming progressive majority of 250 out of 303 (82.5%) respondents show unsurprisingly very average behavior with real and estimated scores higher only on average. Leaving out the 67 slightly progressive respondents changes little. 200 out of 300 respondents (66.7%) identified as politically left shows again similar close to average behavior, as the left and progressive groups largely overlap (195). On average, these groups tend to score a bit better than average (however not significalty) and also overestimate themselves a bit more (only significant on average).

Dunning-Kruger effect To rule out any effects on overestimation levels, the respondents were asked to tick any concepts they were familiar with from a list including the Dunning-Kruger effect. 94 out of 316 (29.7%) respondents stated to be familiar with the DK effect and this had no significant impact on the overestimation for both score types. However for the percentile scores, where respondents are asked to compare themselves with their peers, respondents aware of the DK effect overestimate themselves less over all quartiles. A more flat distribution of overestimation over the quartiles is the result, showing no significant difference already in Q3 (and also in Q4) between the real and estimated score. The real scores of the respondents aware of the DK effect is significantly higher for both score types. This suggests that for this group, knowledge about the topic and about one's own abilities compared to others is linked, as described by Dunning and Kruger in their original paper Kruger and Dunning [16].

Bonferroni correction Most subgroups also show significant differences between the real and estimated score in Q1 - Q3 when compared to the Bonferroni corrected $\alpha = 0.00025$, for both number and percentile scores (see Appendix D). Some significant differences between subgroup and not-subgroup estimate and real scores are not significant for the corrected α, however most are still significant (61.5%). When looking only at sg-nsg differences averaged over all four quartiles, 70.6% of all significant findings remain significant when using the corrected α level. Most notably, significant findings related to age subgroups are all not significant anymore with Bonferroni correction, suggesting a relatively small influence of this factor on differences in real and estimated scores (see Appendix E).

4 Discussion

4.1. Response

While feedback was gathered on the questions used by a pilot survey, still eleven respondents commented on unclear questioning and/or discussing the accuracy of the multiple choice answers given for the climate change knowledge questions. Others also commented on the questions regarding people their backgrounds being too specific or not applicable to their situation. Producing a clear survey that is understandable to everyone was thus not possible and for this report, nothing was done with the contributions of these respondents. However it remains important to strive for non-ambiguous questions and a more thorough pilot with more respondents could have been done.

The questionnaire was intended for an English speaking audience as the questions were only phrased in English. This could have led to additional misinterpretation for not native English speakers. Differences were indeed found between the native English speaking group and the rest, however it is not clear from this survey whether this is because of their skills in English. The finding that the actual scores of native speakers was lower than for non native speakers for both score types suggests that there was either not a significant language barrier for the non native speakers or that there are other factors at play.

The effect size of 0.85 found for the DK effect related to climate change science seems relatively high. However, existing literature on the DK effect rarely report the found effect size (as far as the authors are aware of) making comparison impossible.

4.2. Respondents

Many biases exist in the respondents group, some more clear than others. Virtually no politically right or conservative leaning respondents were present and a distinct majority of the respondents is highly educated (Bachelor degree or more). The average and median age are also way lower than seen in society at large, mostly under-representing middle age and old people. This makes it hard to generalize the findings to a bigger population. On the other hand, a group that could have been targeted more specifically for this survey are Earth and Climate scientist. This group was unfortunately too small to make any sensible conclusions, which would have added to the understanding of the relation between actual and estimated knowledge. The spread of the number scores show that, next to attracting more respondents that score relatively high, even more effort is needed to attract respondents that score relatively low. This way, the four quartiles would give a better overview of the possible scores one could attain, assuming more extreme scores to be present in society.

4.3. Subgroups

The subgroup standing out most from the rest are the UK residents, highly overlapping with the native English speaking group. Other factors that enhanced estimation skills were checked for the UK group (65 respondents) as post-hoc analysis, such as being aged ≤ 25 (17) or > 45 (18), female (33) or an earth scientist (0). These factors could not explain the differences in an univocal way. One (indirect) reason could be the way the respondents were recruited, being highly dependent on a handful of people with different backgrounds targeting only residents of the Netherlands or the UK.

Difference number and percentile scores Another finding relates to the two types of measuring overestimation. Many subgroups show similar behavior for both number and percentile scores, however being aged > 45 seems to show how people can interpret these questions differently. Respondents in this group seem to overestimate themselves more or less around the same percentile score. It seems that for this group, while overestimation in Q1 and Q2 is still significant, they more accurately estimate their knowledge level when this is related to their peers.

Q4 underestimation Some underestimation for Q4 is expected, as it is hard to overestimate an almost perfect score (the regression to the mean effect). Literature on the DK effect suggests that the opposite of what happens in Q1 can hold for Q4: the more knowledgeable one is on a certain topic, the better one understands the complexity of the subject and also the better ones 'metacognitive skill to recognize competence and incompetence in others' is developed [16]. This could explain why this Q4 underestimation holds true especially when respondents are asked to compare themselves with their peers. Additionally, the false-consensus effect (the belief that ones peers have the same level of knowledge) could be at play and was already related to Q4 underestimation by the original Dunning-Kruger study [16].

Relation overestimation and highest attained degree The Q4 effect could also partly explain why respondents with a Bachelor degree as highest attained educational level don't see a drop in overestimation for Q4 towards a more realistic or underestimated guess for both score types. These students have started learning about a certain topic however mostly not yet fully in depth when compared with Master or Doctorate and Professional degree holders. Nevertheless, Doctorate and Professional degree holders in this survey highly overestimate themselves compared to others, most notably for the number scores.

Another subgroup that was thought to increase estimation accuracy, being in the field of Natural Sciences, again had an opposite effect with significantly higher overestimations over all the quartiles for both score types except for one quartile. A reason for these above average overestimations could again be a version of the DK effect, where people that are highly educated (in anything) or are educated in Natural Sciences (of which Earth and Climate Science are a part) falsely believe they are more knowledgeable on a certain topic they in fact have insufficient knowledge about. Both Doctorate and Professional degree holders and Natural Scientists score above average, but this is more than offset by the extra overestimation. The effect of other factors outside the scope of this survey would need to be studied more elaborately to substantiate these hypotheses and explain this behavior.

Relation overestimation and political affiliation Personal values and political affiliations are brought forward by Kahan [14] and Fernbach et al. [6] to why a Dunning-Kruger effect might not show in relation to climate change knowledge. Including political affiliation in the factor effect study was unfortunately not possible with this data set. We recommend taking this into account as it could help better understand the suggestions by Kahan [14] and Fernbach et al. [6] and the conclusions taken recently by Fischer et al. [7] who did not find a DK effect related to climate change science. The sample used in Fischer's study was merely balanced in terms of gender, age and geographical distribution within Germany and did not look at political background. Depolarizing the science behind climate change and shifting the focus on the actually relevant discussions does indeed seem like a fruitful direction.

Subgroup statistical analysis The Bonferroni correction is added but not strictly taken into account in the results analysis. Some seemingly significant results are not significant when comparing to $\alpha = 0.00025$. For a first analysis, it feels more useful to identify too many possible influences than too little. As Rothman [22] points out, 'scientists should not be so reluctant to explore leads that may turn out to be wrong that they penalize themselves by missing possibly important findings' (p. 43). In other words, typically Type I errors are less of concern as the null hypothesis is rarely thought to be true [10]. For transparency, the Bonferroni correction is used to label observations that are still significant under the strict and conservative Bonferroni assumptions.

Additionally, only limited intersectional analysis is done, which would combine different subgroup traits to form new subgroups on the 'intersection'. Next to including more background information of the respondents, this could be a worthwhile lead towards a better understanding of the differences seen between the current subgroups. However, as seen in the political ideology analysis, combining two or more subgroups requires enough respondents from all subgroups under study. The authors recognize that a more robust analysis should be performed in the future, taking into account the multiple comparisons problem and possible relations between the different factors. Examples are given by Gelman et al. [10] and Lehmann and Romano [17].

5 Conclusion

Current ways of communicating the science and urgency of climate change relies heavily on knowledge transfer and its efficiency is shown to be far from optimal [11, 12, 18, 20, 25]. The author believes the current emphasis on knowledge transfer in climate change science communication should be critically examined and changed according to the presented and related findings.

Helpful recommendations for a more effective communication strategy are done by Treisse, Debbie and Weigold [28], Hart and Nisbet [11], Simis et al. [25] and Hut et al. [12]. How the knowledge deficit model can possibly be useful for certain types of climate change knowledge dimensions is suggested by Taddicken et al. [27]. This paper aims to contribute and complement these insights by researching the Dunning-Kruger effect in relation to climate change science. For all score types and subgroups studied, the first quartile shows the largest significant overestimation compared to other quartiles. This overestimation is significant, even when the conservative Bonferroni correction is applied. This statement holds even for all second quartiles, except for UK residents and native English speakers (two highly overlapping subgroups). For the whole group and a majority of the subgroups, the estimated scores approach or even underestimate the actual score when moving to Q3 and Q4. These observations compare well with existing literature on the DK effect ([16, 21, 23] amongst others), showing the presence of the DK effect related to climate change science. The initial exploration into the factors influencing the real and estimated scores and the discrepancy between these also show some significant results and are interesting leads for further research.

However, major biases in the respondent group and an insufficient multiple comparisons problem analysis make it hard to generalize the current findings to society at large. This can best be overcome by redesigning the way respondents are recruited and incorporating a more robust system that can take into account overlapping and possibly related subgroups and multiple comparisons. Also, not all effects found can as of now be explained, leaving room for further research into additional factors.

It should be stated that the emphasis on background factors is not to promote differentiated communication tending to different (political) groups as this could easily trigger negative consequences (making climate change science even more politicized, driving people even more into different bubbles e.g.). Further researching the influence of political affiliation as suggested by Kahan [14] and Fernbach et al. [6] as well as other factors relating to the DK effect gives us a better understanding of effective climate change science communication that does not have to fully rely on the imperfect knowledge deficit model.

Appendices

Appendix A: Questionnaire

Note: texts in italics indicate correct answers.

1. Which statement describes weather (and the science of meteorology) best?
    - It's local, based on historical averages and it focuses on long-term trends
    - *It gives precise information and predicts only in the short-term*
2. What is the role of the Greenhouse Effect (GHE) on climate change?
    - *The GHE regulates the temperature of the Earth*
    - The GHE allows fruits and vegetables to grow easily
    - The GHE traps gases produced by human activities
3. How does the GHE work?
    - Particles in the Earth's ground retain the heat at the surface
    - Water in the oceans absorbs the heat and distribute it over the planet
    - *Gases in the atmosphere absorb the heat and send it back to the Earth's surface*
4. What would be the Earth's average temperature without the GHE?
    - 25 degrees Celcius, or 77 degrees Fahrenheit
    - -3 degrees Celcius, or 37 degrees Fahrenheit
    - *-18 degrees Celcius, or 0 degrees Fahrenheit*
5. The Earth is currently undergoing its first climate change.
    - True
    - *False*
6. Which of the following are natural causes of climate change?
    - The changing of seasons: autumn, winter, spring, summer
    - *Plate tectonics*
    - *Solar activity*
    - The tides
    - Cyclones
    - *Volcanic activity*
7. Fill in the blank space: Where humans influence climate change, the effect they have is generally ..... compared to natural climate change.
    - *Faster*
    - Slower
    - No different
    - Humans don't influence climate change
8. Which of the following human activities globally produce most greenhouse gas emissions?
    - Transport
    - Industry
    - Production and recycling of waste
    - *Heating and production of electricity*

9. Which of the following is a main direct consequence of climate change?
   - Lunar forces
   - *Rises in average temperature*
   - Air pollution
10. Which of the following cause a rise in the sea level?
   - *Melting glaciers and expanding oceans*
   - Melting sea ice and increasing rainfall
11. If we asked the above 10 questions to 100 random people from your country, how do you think you scored compared to them?
   - 0: I performed in the lowest 10
   - 1
   - 2
   - 3
   - 4
   - 5: I performed around average
   - 6
   - 7
   - 8
   - 9
   - 10: I performed in the top 10
12. Of the 10 questions, how many questions do you think you answered correctly?
   - 0: I answered 0 questions correctly
   - 1
   - 2
   - 3
   - 4
   - 5: I answered half of all questions correctly
   - 6
   - 7
   - 8
   - 9
   - 10: I answered all 10 questions correctly
13. What is your age?
   - [...]
   - I prefer not to answer
14. To which gender identity do you most identify?
   - [...]
   - I prefer not to answer
15. In which country do you currently live?
   [list]
16. Are you an English native speaker?
   - Yes
   - No
   - I prefer not to say
17. What is the highest level of school you have completed or highest degree you have received?

- Primary education
- Secondary education
- Bachelor degree
- Master degree
- Professional school degree or doctorate degree
- I prefer not to answer

17b. (Only for answer c - g) In what field did you study?
- Social sciences
- Arts and humanities
- Natural sciences
- Engineering
- Other [...]
- I prefer not to answer

17c. (Only for answer c and d) Do you consider yourself an earth/climate scientist?
- Yes
- No
- I prefer not to answer

18. Where would you place yourself on this political ideology scale?
- Extremely progressive
- Progressive
- Slightly progressive
- Moderate/middle
- Slightly conservative
- Conservative
- Extremely conservative
- I prefer not to answer

19. Where would you place yourself on this political ideology scale?
- Extremely left
- Left
- Slightly left
- Moderate/middle
- Slightly right
- Right
- Extremely right
- I prefer not to answer

20a. Did you vote during the last national election?
- Yes
- No
- I'm not eligible to vote
- I prefer not to answer

20b. (Only for answer a) What is the party you voted on in the last national election?
- [...]
- I prefer not to answer

21. Are you aware of the existence of the climate literacy quiz from the EU?
- Yes
- No

22. Which of these concepts are you familiar with?
    - Blockchain technology
    - Placebo effect
    - Financial crisis
    - Risk
    - Dunning-Kruger effect
    - Circular economy
    - Albedo effect
    - Internet of things
    - Genetically Modified Organisms
    - Unemployment rate
23. Do you have any final comments?
    [. . .]

Appendix B: Required sample size

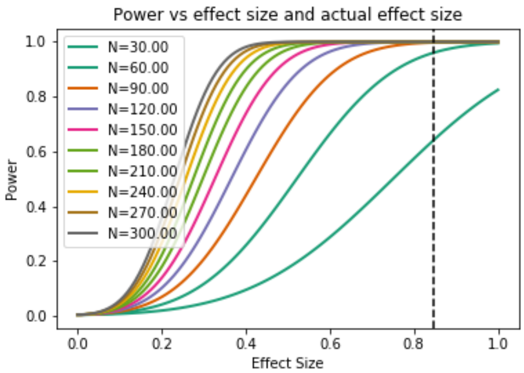

Figure 5.1: Power vs effect size and actual effect size (dashed line)

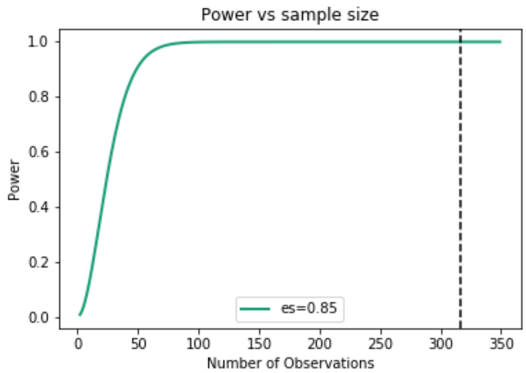

Figure 5.2: Power vs sample size and actual sample size (dashed line) for actual effect size 0.85

Appendix C: Subgroup analysis

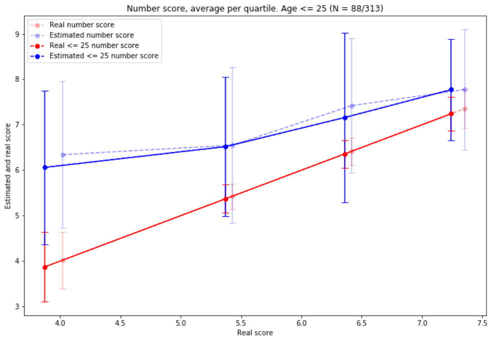

Figure 5.3: Subgroup Age ≤ 25 number scores

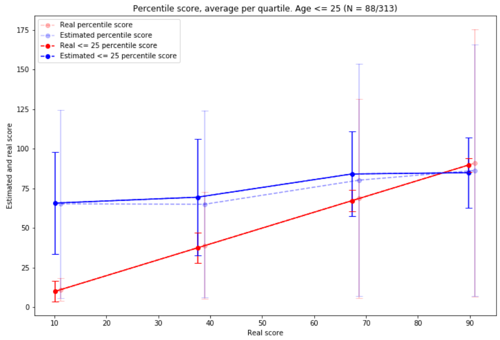

Figure 5.4: Subgroup Age ≤ 25 percentile scores

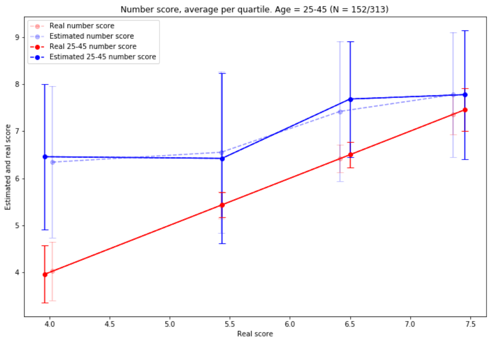

Figure 5.5: Subgroup Age 25-45 number scores

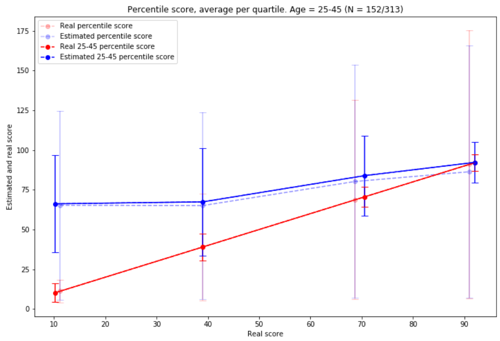

Figure 5.6: Subgroup Age 25-45 percentile scores

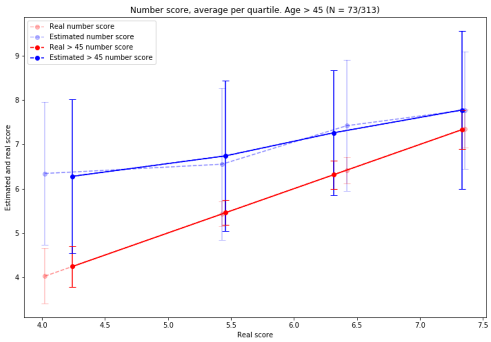

Figure 5.7: Subgroup Age > 45 number scores

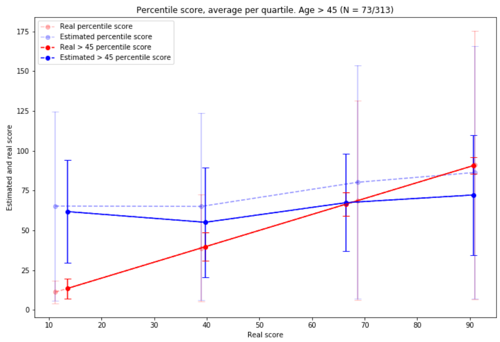

Figure 5.8: Subgroup Age > 45 percentile scores

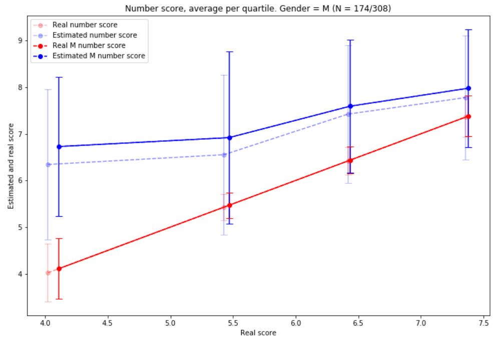

Figure 5.9: Subgroup Gender = M number scores

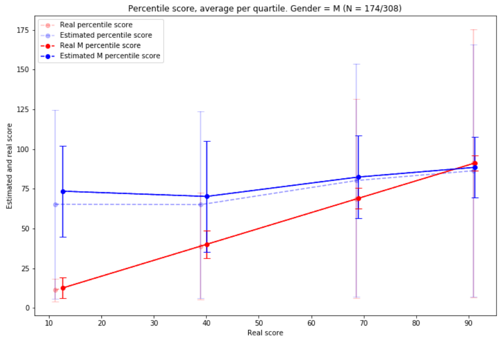

Figure 5.10: Subgroup Gender = M percentile scores

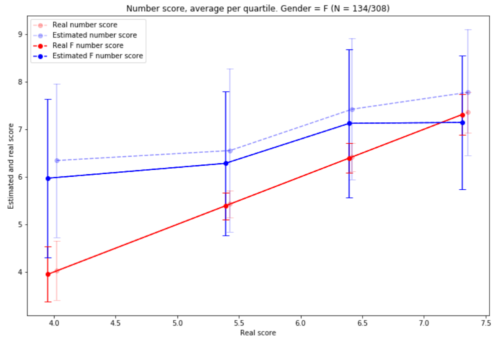

Figure 5.11: Subgroup Gender = F number scores

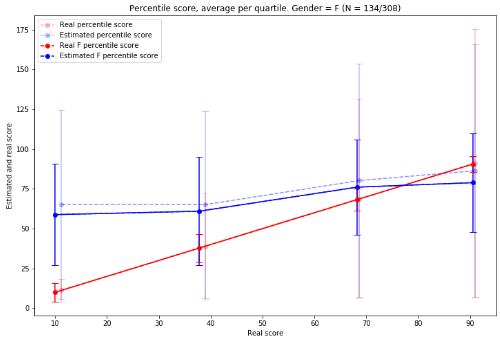

Figure 5.12: Subgroup Gender = F percentile scores

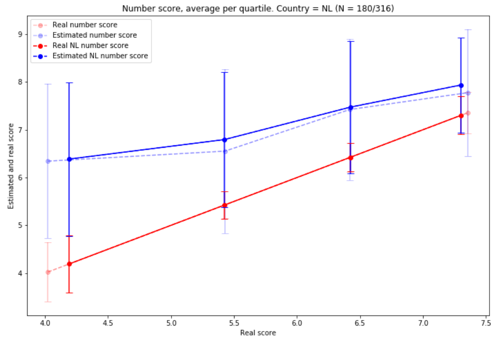

Figure 5.13: Subgroup Country of residence = NL number scores

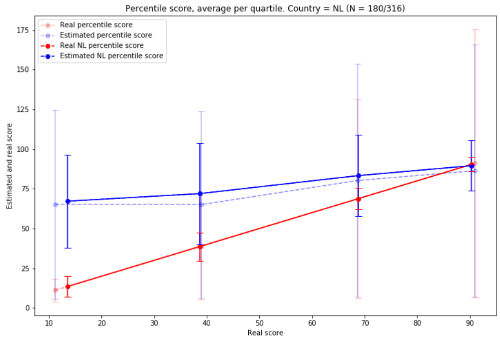

Figure 5.14: Subgroup Country of residence = NL percentile scores

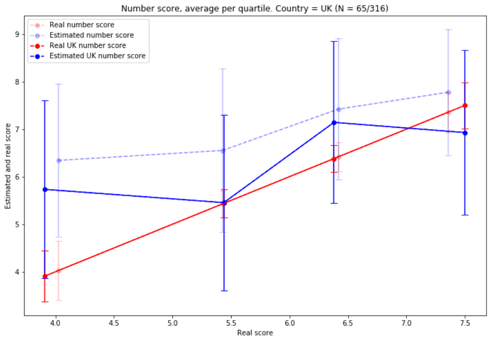

Figure 5.15: Subgroup Country of residence = UK number scores

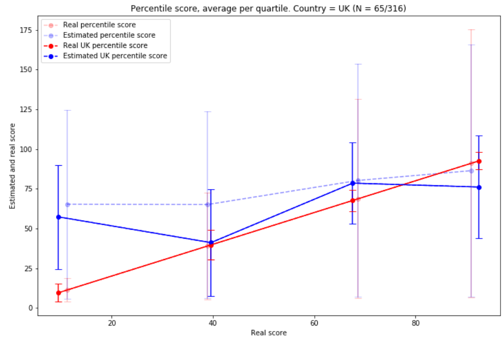

Figure 5.16: Subgroup Country of residence = UK percentile scores

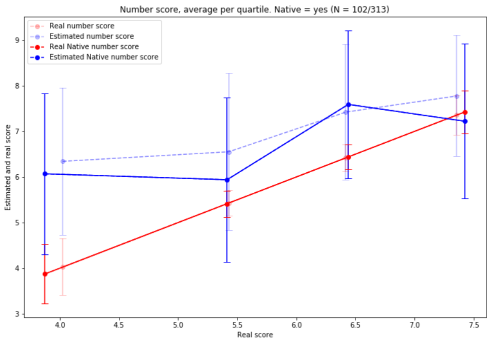

Figure 5.17: Subgroup Native English speaker = yes number scores

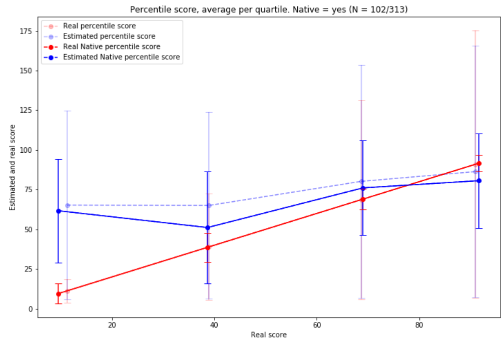

Figure 5.18: Subgroup Native English speaker = yes percentile scores

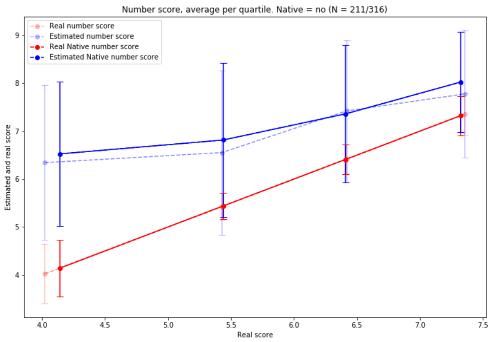

Figure 5.19: Subgroup Native English speaker = no number scores

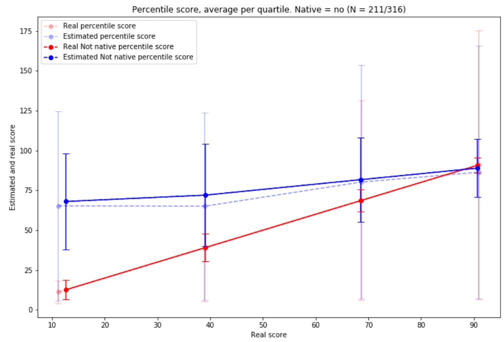

Figure 5.20: Subgroup Native English speaker = no percentile scores

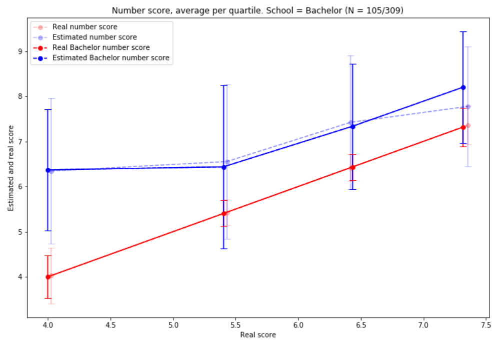

Figure 5.21: Subgroup Highest attained degree = Bachelor number scores

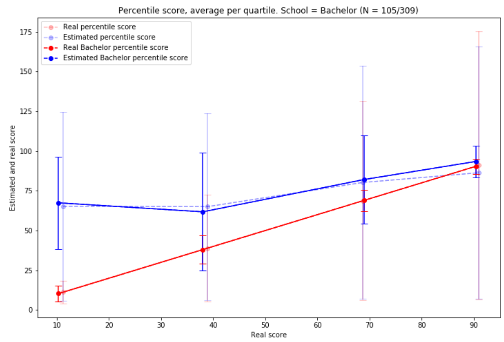

Figure 5.22: Subgroup Highest attained degree = Bachelor percentile scores

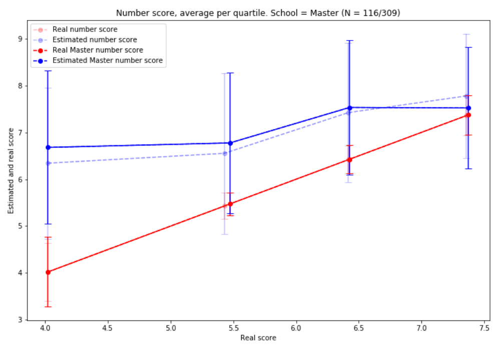

Figure 5.23: Subgroup Highest attained degree = Master number scores

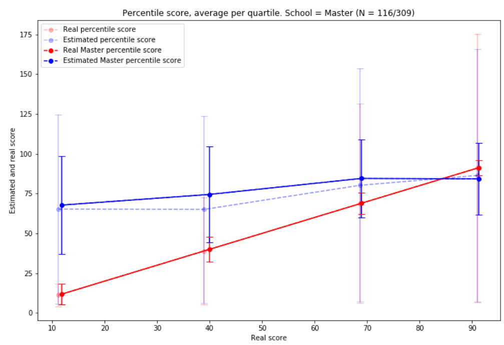

Figure 5.24: Subgroup Highest attained degree = Master percentile scores

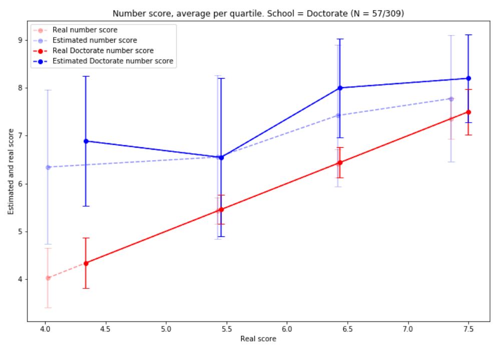

Figure 5.25: Subgroup Highest attained degree = Doctorate / Professional school number scores

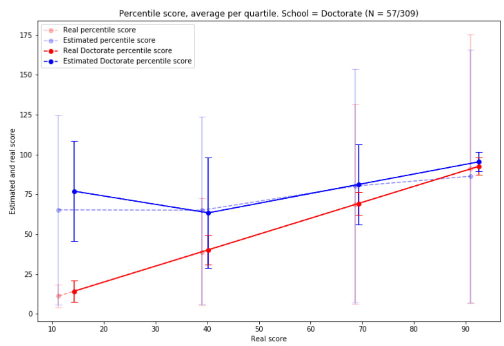

Figure 5.26: Subgroup Highest attained degree = Doctorate / Professional school percentile scores

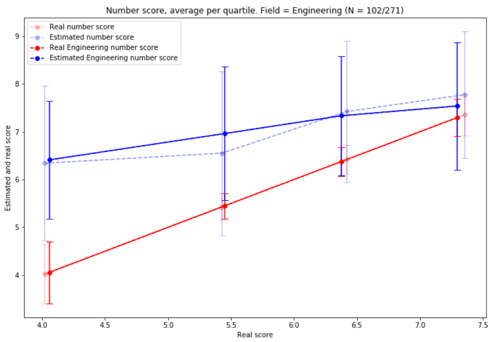

Figure 5.27: Subgroup Field = Engineering number scores

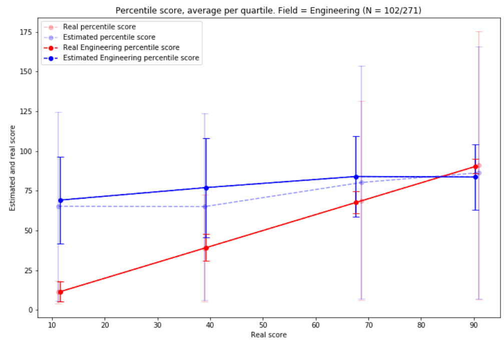

Figure 5.28: Subgroup Field = Engineering percentile scores

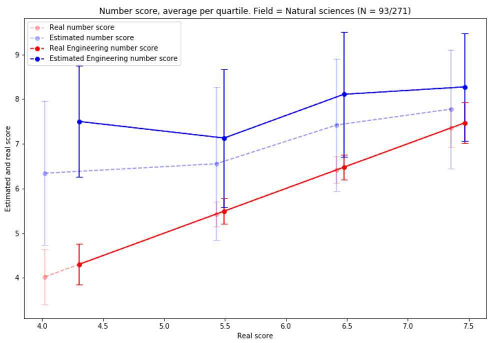

Figure 5.29: Subgroup Field = Natural sciences number scores

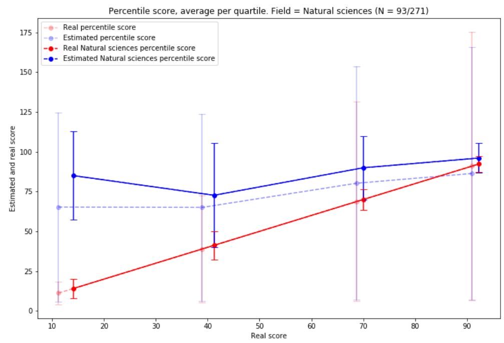

Figure 5.30: Subgroup Field = Natural sciences percentile scores

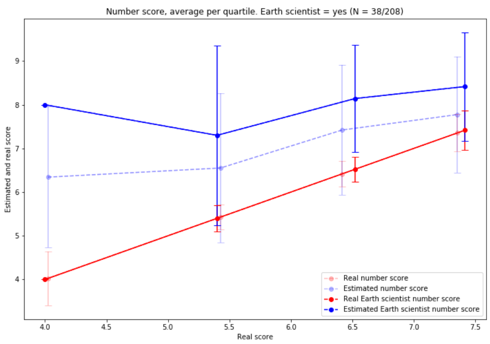

Figure 5.31: Subgroup Earth scientist number scores

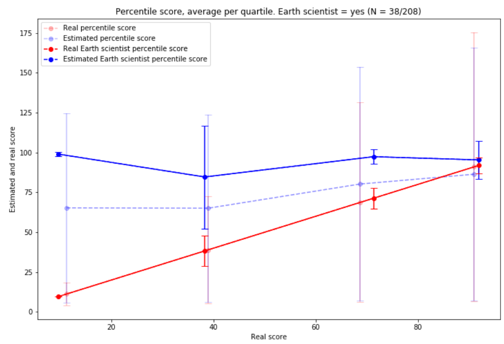

Figure 5.32: Subgroup Earth scientist percentile scores

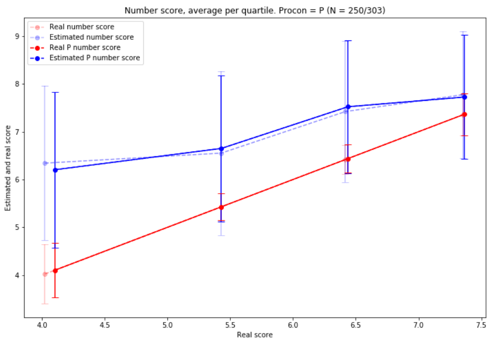

Figure 5.33: Subgroup Ideology = Progressive number scores

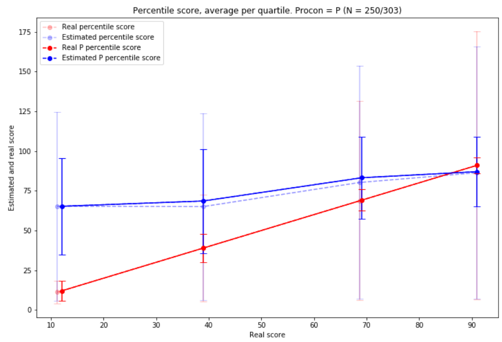

Figure 5.34: Subgroup Ideology = Progressive percentile scores

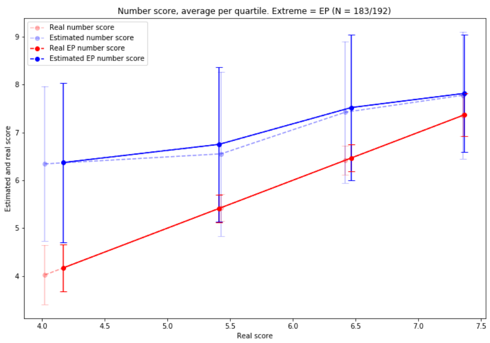

Figure 5.35: Subgroup Ideology = Extra progressive number scores

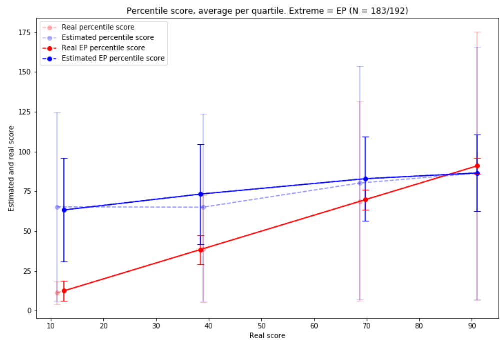

Figure 5.36: Subgroup Ideology = Extra progressive percentile scores

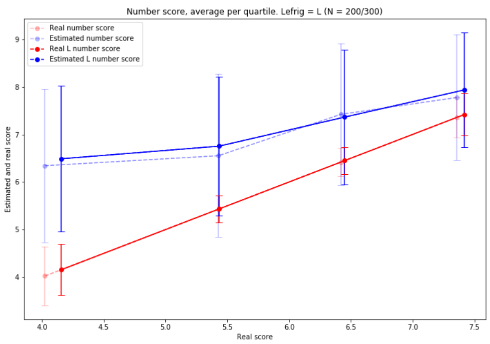

Figure 5.37: Subgroup Ideology = Left number scores

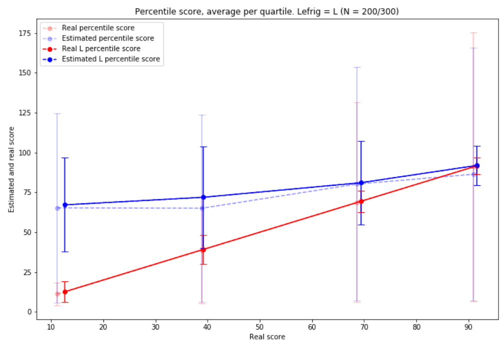

Figure 5.38: Subgroup Ideology = Left percentile scores

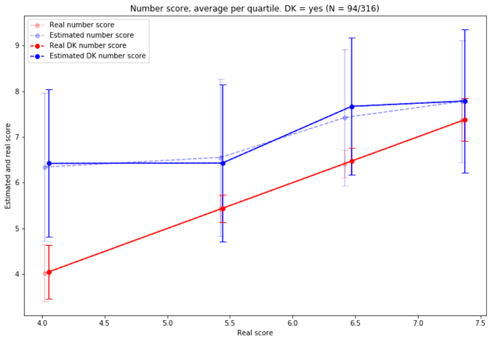

Figure 5.39: Subgroup Dunning-Kruger number scores

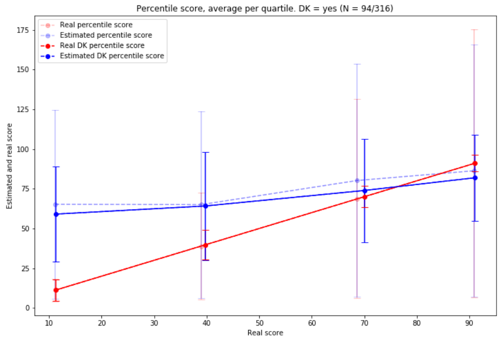

Figure 5.40: Subgroup Dunning-Kruger percentile scores

Appendix D: Real-estimated scores p-values

Table 5.1: Comparing real and estimated number scores

N	Real - estimated			
	Q1	**Q2**	**Q3**	**Q4**
All	<0.001†	<0.001†	<0.001†	0.016
Age ≤ 25	<0.001†	<0.001†	<0.001†	<0.001
Age 25-45	<0.001†	<0.001†	<0.001†	0.079
Age >45	<0.001†	<0.001†	<0.001†	0.063
Gender = M	<0.001†	<0.001†	<0.001†	<0.001
Gender = F	<0.001†	<0.001†	<0.001†	0.375*
Country = NL	<0.001†	<0.001†	<0.001†	<0.001†
Country = UK	<0.001†	0.933	<0.001	0.015*
Native = yes	<0.001†	<0.001	<0.001†	0.365*
Native = no	<0.001†	<0.001†	<0.001†	<0.001†
School = Bachelor	<0.001†	<0.001†	<0.001†	<0.001†
School = Master	<0.001†	<0.001†	<0.001†	0.392
School = Doctorate	<0.001†	<0.001†	<0.001†	<0.001†
Field = Engineering	<0.001†	<0.001†	<0.001†	0.185
Field = Natural sci.	<0.001†	<0.001†	<0.001†	<0.001†
Earth scientist = no	<0.001†	<0.001†	<0.001†	0.066
PC = Progressive	<0.001†	<0.001†	<0.001†	0.045
PC = Extra pro.	<0.001†	<0.001†	<0.001†	0.027
LR = Left	<0.001†	<0.001†	<0.001†	<0.001
DK = yes	<0.001†	<0.001†	<0.001†	0.053

* = estimate lower than real score, † = significant for Bonferroni corrected $\alpha = 0.00025$

Table 5.2: Comparing real and estimated percentile scores

P	**Real - estimated**			
	Q1	**Q2**	**Q3**	**Q4**
All	<0.001†	<0.001†	0.152	0.658*
Age ≤ 25	<0.001†	<0.001†	<0.001†	0.104*
Age 25-45	<0.001†	<0.001†	<0.001†	0.915
Age >45	<0.001†	<0.001†	0.779	<0.001*
Gender = M	<0.001†	<0.001†	<0.001†	0.280*
Gender = F	<0.001†	<0.001†	0.021	<0.001*
Country = NL	<0.001†	<0.001†	<0.001†	0.691*
Country = UK	<0.001†	0.661	<0.001†	<0.001*
Native = yes	<0.001†	<0.001	0.033	<0.001*
Native = no	<0.001†	<0.001†	<0.001†	0.471*
School = Bachelor	<0.001†	<0.001†	<0.001†	0.023
School = Master	<0.001†	<0.001†	<0.001†	0.022*
School = Doctorate	<0.001†	<0.001†	<0.001†	<0.001
Field = Engineering	<0.001†	<0.001†	<0.001†	<0.001*
Field = Natural sci.	<0.001†	<0.001†	<0.001†	<0.001
Earth scientist = no	<0.001†	<0.001†	<0.001†	0.192*
PC = Progressive	<0.001†	<0.001†	<0.001†	0.183*
PC = Extra pro.	<0.001†	<0.001†	<0.001	0.245*
LR = Left	<0.001†	<0.001†	<0.001†	0.874
DK = yes	<0.001†	<0.001†	0.276	0.013*

* = estimate lower than real score, † = significant for Bonferroni corrected $\alpha = 0.00025$

Appendix E: Subgroup-not subgroup scores p-values

Table 5.3: Comparing subgroup-not subgroup real and estimated number scores

N	Estimate sg-nsg					Real sg-nsg
	Q1	Q2	Q3	Q4	All	All
Age ≤ 25	0.173*	0.830*	0.207*	0.983	0.800*	0.020**
Age 25-45	0.540	0.467*	0.0572	0.992*	0.430	0.782
Age >45	0.757*	0.261	0.311*	0.994	0.422*	0.009**
Gender = M	0.036	0.048	0.287	0.235	<0.001†	<0.001†
Gender = F	0.071*	0.075*	0.091*	0.001*	<0.001*†	<0.001**†
Country = NL	0.818	0.084	0.738	0.228	<0.001†	<0.001†
Country = UK	<0.001*†	<0.001*†	0.140*	<0.001*	<0.001*	0.968
Native = yes	0.203*	<0.001*	0.345	0.016*	<0.001*†	<0.001**
Native = no	0.317	0.102	0.700*	0.076	<0.001†	<0.001†
School = Bachelor	0.878	0.511*	0.568*	0.012	0.956	0.547**
School = Master	0.100	0.134	0.494	0.138*	0.013	0.333
School = Doctorate	<0.001	0.966*	<0.001†	<0.001	<0.001†	0.001
Field = Engineering	0.682	0.007	0.545*	0.196*	0.705	<0.001
Field = Natural sci.	<0.001†	<0.001	<0.001†	<0.001	<0.001†	<0.001†
Earth scientist = no	0.076	0.017	0.492	0.806*	<0.001†	<0.001†
PC = Progressive	0.496*	0.513	0.523	0.774*	<0.001†	<0.001†
PC = Extra pro.	0.915	0.338	0.645	0.838	0.004	<0.001†
LR = Left	0.449	0.165	0.719*	0.336	<0.001†	0.318
DK = yes	0.686	0.460*	0.139	0.973	0.086	<0.001

* = lower than average estimate, ** = lower than average real score

† = significant for Bonferroni corrected $\alpha = 0.00025$

Table 5.4: Comparing subgroup-not subgroup real and estimated percentile scores

P	**Estimate sg-nsg**					**Real sg-nsg**
	Q1	**Q2**	**Q3**	**Q4**	**All**	**All**
Age ≤ 25	0.895	0.219	0.190	0.625*	0.051*	0.001**
Age 25-45	0.800	0.478	0.191	0.001	<0.001	0.581
Age >45	0.376*	<0.001*	<0.001*	<0.001*	<0.001*	0.001**
Gender = M	0.021	0.139	0.450	0.401	<0.001†	<0.001†
Gender = F	0.071*	0.075*	0.091*	0.001*	<0.001*†	<0.001**†
Country = NL	0.584	0.028	0.278	0.131	<0.001†	<0.001
Country = UK	0.045*	<0.001*†	0.547*	0.019*	<0.001*†	0.970**
Native = yes	0.367*	<0.001*†	0.213*	0.149*	<0.001*†	<0.001**
Native = no	0.450	0.030	0.613	0.288	<0.001†	<0.001†
School = Bachelor	0.548	0.371*	0.550	0.109	0.673	0.750**
School = Master	0.510	0.002	0.120	0.485*	0.001	0.129
School = Doctorate	<0.001	0.630*	0.724	<0.001†	0.054	0.022
Field = Engineering	0.298	<0.001	0.210	0.340*	0.749	<0.001†
Field = Natural sci.	<0.001†	0.029	<0.001†	<0.001†	<0.001†	<0.001†
Earth scientist = no	0.131	0.036	0.526	0.623	<0.001†	<0.001†
PC = Progressive	0.998*	0.273	0.204	0.822	<0.001†	<0.001†
PC = Extra Pro.	0.708*	0.042	0.468	0.966	0.045	<0.001†
LR = Left	0.611	0.030	0.785	0.002	<0.001†	0.859
DK = yes	0.088*	0.798*	0.082*	0.221*	0.430*	<0.001

* = lower than average estimate, ** = lower than average real score

† = significant for Bonferroni corrected $\alpha = 0.00025$

Bibliography

[1] Liars and Fools: Climate change deniers | NBR. URL https://www.nbr.co.nz/opinion/liars-and-fools-climate-change-deniers.

[2] Module 1: Introduction to Climate Change. URL https://www.climate-literacy.eu/en/elearning/modules/170-module-1.

[3] Environmental Law Prof Blog, 2011. URL https://lawprofessors.typepad.com/environmental{_}law/2011/09/the-tea-party-the-dunning-kruger-effect-and-climate-change-too-sure-of-themselves-.html.

[4] John Cook. The Dunning-Kruger effect and the climate debate, 2010. URL https://skepticalscience.com/The-Dunning-Kruger-effect-and-the-climate-debate.html.

[5] Katelyn M Cooper, Anna Krieg, and Sara E Brownell. Who perceives they are smarter? exploring the influence of student characteristics on student academic self-concept in physiology. *Advances in physiology education*, 42(2):200–208, 2018.

[6] Philip M. Fernbach, Nicholas Light, Sydney E. Scott, Yoel Inbar, and Paul Rozin. Extreme opponents of genetically modified foods kno the least but think they know the most. *Nature Human Behaviour*, 2019.

[7] Helen Fischer, Dorothee Amelung, and Nadia Said. their climate change knowledge. *Nature Climate Change*, 2019. ISSN 1758-6798. doi: 10.1038/s41558-019-0563-0. URL http://dx.doi.org/10.1038/s41558-019-0563-0.

[8] Angela Fritz. What's behind the confidence of the incompetent? This suddenly popular psychological phenomenon. - The Washington Post, 2019. URL https://www.washingtonpost.com/science/2019/01/07/whats-behind-confidence-incompetent-this-suddenly-popular-psychological-phenomenon/.

[9] Andrew Gelman and Eric Loken. The garden of forking paths: Why multiple comparisons can be a problem, even when there is no "fishing expedition" or "p-hacking" and the research hypothesis was posited ahead of time. *Department of Statistics, Columbia University*, 2013.

[10] Andrew Gelman, Jennifer Hill, and Masanao Yajima. Why we (usually) don't have to worry about multiple comparisons. *Journal of Research on Educational Effectiveness*, 5(2):189–211, 2012.

[11] P Sol Hart and Erik C Nisbet. Boomerang Effects in Science Communication : How Motivated Reasoning and Identity Cues Amplify Opinion Polarization About Climate Mitigation Policies. 2012. doi: 10.1177/0093650211416646. URL https://doi.org/10.1177/0093650211416646.

[12] Rolf Hut, Anne M Land-zandstra, Ionica Smeets, and Cathelijne R Stoof. Geoscience on television : a review of science communication literature in the context of geosciences. pages 2507–2518, 2016. doi: 10.5194/hess-20-2507-2016.

[13] Malte Jansen, Ulrich Schroeders, and Oliver Lüdtke. Academic self-concept in science: Multidimensionality, relations to achievement measures, and gender differences. *Learning and Individual Differences*, 30:11–21, 2014.

[14] Dan M Kahan. Climate-Science Communication and the Measurement Problem. 36, 2015. doi: 10.1111/pops.12244.

[15] Meg Kennedy, Brian; Hefferon. Rising U.S. concern about climate change is mostly among Democrats | Pew Research Center, 2019. URL https://www.pewresearch.org/fact-tank/2019/08/28/u-s-concern-about-climate-change-is-rising-but-mainly-among-democrats/.

[16] Justin Kruger and David Dunning. Unskilled and Unaware of It : How Difficulties in Recognizing One ' s Own Incompetence Lead to Inflated. (January 2000), 1999. doi: 10.1037//0022-3514.77.6.1121.

[17] Erich Leo Lehmann and Joseph P Romano. Generalizations of the familywise error rate. In *Selected Works of EL Lehmann*, pages 719–735. Springer, 2012.

[18] Mark S Mccaffrey and Susan M Buhr. Clarifying Climate Confusion : Addressing Systemic Holes , Cognitive Gaps , and Misconceptions Through Climate Literacy CLIMATE LITERACY. (October 2014):37–41, 2013. doi: 10.2747/0272-3646.29.6.512.

[19] Matthew Motta, Timothy Callaghan, and Steven Sylvester. Social Science & Medicine Knowing less but presuming more : Dunning-Kruger e ff ects and the endorsement of anti-vaccine policy attitudes. *Social Science & Medicine*, 211(January):274–281, 2018. ISSN 0277-9536. doi: 10.1016/j.socscimed.2018.06.032. URL https://doi.org/10.1016/j.socscimed.2018.06.032.

[20] Warren Pearce, Brian Brown, Brigitte Nerlich, and Nelya Koteyko. Communicating climate change : conduits , content , and consensus. 6(December):613–626, 2015. doi: 10.1002/wcc.366.

[21] Gordon Pennycook, Robert M Ross, Derek J Koehler, and Jonathan A Fugelsang. Dunning – Kruger effects in reasoning : Theoretical implications of the failure to recognize incompetence. pages 1774–1784, 2017. doi: 10.3758/s13423-017-1242-7.

[22] Kenneth J Rothman. No adjustments are needed for multiple comparisons. *Epidemiology*, pages 43–46, 1990.

[23] Marina Sawdon and Gabrielle Finn. world setting. pages 279–285, 2014. doi: 10.1111/joa.12072.

[24] Marshall Shepherd. No, Earth's Climate Is Not In A 10-Year Cooling Trend, 2016. URL https://www.forbes.com/sites/marshallshepherd/2016/09/30/no-climate-is-not-in-a-10-year-cooling-trend-an-example-of-the-dunning-kruger-effect/{#}534a9b052c70.

[25] Molly J Simis, Haley Madden, Michael A Cacciatore, and Sara K Yeo. The lure of rationality : Why does the deficit model persist in science communication ? (X), 2016. doi: 10.1177/0963662516629749. URL https://doi.org/10.1177/0963662516629749.

[26] Arek Sinanian. The Rejection of Climate Change - Fair Observer, 2016. URL https://www.fairobserver.com/more/environment/the-rejection-of-climate-change-31394/.

[27] Monika Taddicken, Anne Reif, and Imke Hoppe. What do people know about climate change — and how confident are they ? On measurements and analyses of science related knowledge. 17(03):1–26, 2018.

[28] Michael F Treisse, Debbie and Weigold. Advancing Science Communication. 23(3): 310–322, 2002.