W241 - Final Project - Icons, Charts and Denominators

Joshua Noble, Ryan T Orton, Sandip Panesar

4/8/2021

Introduction

Reading and understanding numerical ratios may present difficulties for people regardless of their mathematical abilities[1]. Much of the booming data visualization market has been built around the premise that humans are able to count objects and interpret ratios better when they are presented graphically, as opposed to numerically. Consequently, peoples' understanding of important concepts involving numerical ratios, such as vaccine efficacy might be related to how information is presented to them, with certain modalities being more likely to elicit understanding. This issue has become particularly relevant due to recent global events. We therefore decided to test the concept of denominator neglect in the context of the COVID-19 pandemic as it relates to vaccine efficacy.

Vaccine efficacy is measured by calculating the incidence of disease among vaccinated and unvaccinated persons, and determining the percentage reduction in incidence among vaccinated persons compared to unvaccinated persons. The greater the percentage reduction of illness in the vaccinated group, the greater the vaccine efficacy. The basic formula is:

$$\text{Vaccine Efficacy} = \frac{\text{Incidence among unvaccinated group} - \text{Incidence among vaccinated group}}{\text{Incidence among unvaccinated group}} \times 100$$

According to Garcia-Retamero et al. (2010), denominator neglect is the focus on the number of times a target event has happened (e.g. the number of treated and untreated patients who are affected, respectively) without considering the overall number of opportunities for it to happen (e.g. the overall number of treated and untreated patients). The authors' studied denominator neglect relating to problems involving treatment risk reduction, asking subjects to interpret medical data presented as numerical ratios. They found that subjects indeed were more likely to pay disproportionate attention to the numerator of a ratio, compared to the denominator. Nevertheless, they demonstrated that icon arrays were effective at mitigating this phenomenon somewhat.

Though visual aids like icon arrays have been shown to enhance understanding of mathematical data, such as ratios, data presented via pie charts have been demonstrated as particularly difficult to understand[3]**.** In information visualization, when presented with several pie charts, subjects are often unable to differentiate between them[4]**.** Though it has been shown that icon arrays are superior to text-based representations of ratios, it is unknown if pie charts are still better at eliciting understanding of proportions compared to text.

Based upon the above, we devised a causal research question:

Are icon arrays (or other visual aids) more effective than text-based descriptions for interpretation of numerical ratios?

Our main hypothesis is based upon the aforementioned research, that visual aids improve interpretability of ratio data. More specifically, that a group presented with a visual aid (of any type) correctly interprets the numerical ratios presented in them at a greater rate, compared to a group presented with ratios in only text or number format. The primary null hypothesis ($h_{1,0}$) is that there is no difference in the proportion of correct answers between a group asked to interpret a vaccine efficacy ratio presented in text form and a group asked to interpret a vaccine efficacy ratio presented graphically. The primary alternative hypothesis ($h_{1,1}$) is that there is a significantly larger proportion of correct interpretations in the group who are presented with a graphical aid. Moreover, we also hypothesize, based upon the reference data, that icon arrays are better than pie charts at enabling understanding of numerical ratios. Therefore our secondary null hypothesis ($h_{2,0}$) is that there is no difference in the number of correct interpretations of ratio data for groups treated with pie chart ratios versus groups treated with icon array ratios. The secondary alternative hypothesis ($h_{2,1}$) is that the proportion of correct responses in the icon array treatment group is greater than that of the pie chart treatment group.

Methods

The original paper on which our study was based[2] examined numerical ratios and icon arrays. In order to test both of our hypotheses, we utilized both pie charts and icon arrays as treatment options. Our survey was built in Qualtrics and the experiment was conducted using the Berkeley XLab for subject recruitment and testing.

Research Design

We adopted an alternative approach to the classical pre-test/post-test ROXO and instead used a RXRXO post-test control group design. Our survey did not involve any observation prior to treatment or control randomization. The subjects were tested via a multi-factorial approach - all users were first presented with several “stem” questions about the COVID-19 pandemic and children returning to classrooms. These questions were the same for all participants and were intended to obscure the actual purpose of our survey so that participants wouldn’t immediately think of it as testing ratios or mathematical concepts. Participants were first presented with a question from the pie chart treatment stage - either a text-based description (control) or a pie chart (treatment): This question stated that a hypothetical vaccine should be 90% effective, and then asked the subject to differentiate between two ratios (either in text or as a pie chart), one of which was correct. Then the same subject was presented with a question from the icon array stage - either a text-based description (control) or an icon array (treatment). This format differed from the first set of questions, in that it actually presented two ratios (one for a placebo group and another for a treatment group) and asked subjects to determine whether the vaccine was better than the placebo based upon presented incidences in the hypothetical groups. Within each stage, subjects were assigned randomly (using Qualtrics' survey flow randomization algorithm) into either control or treatment groups (**Figure 1).**

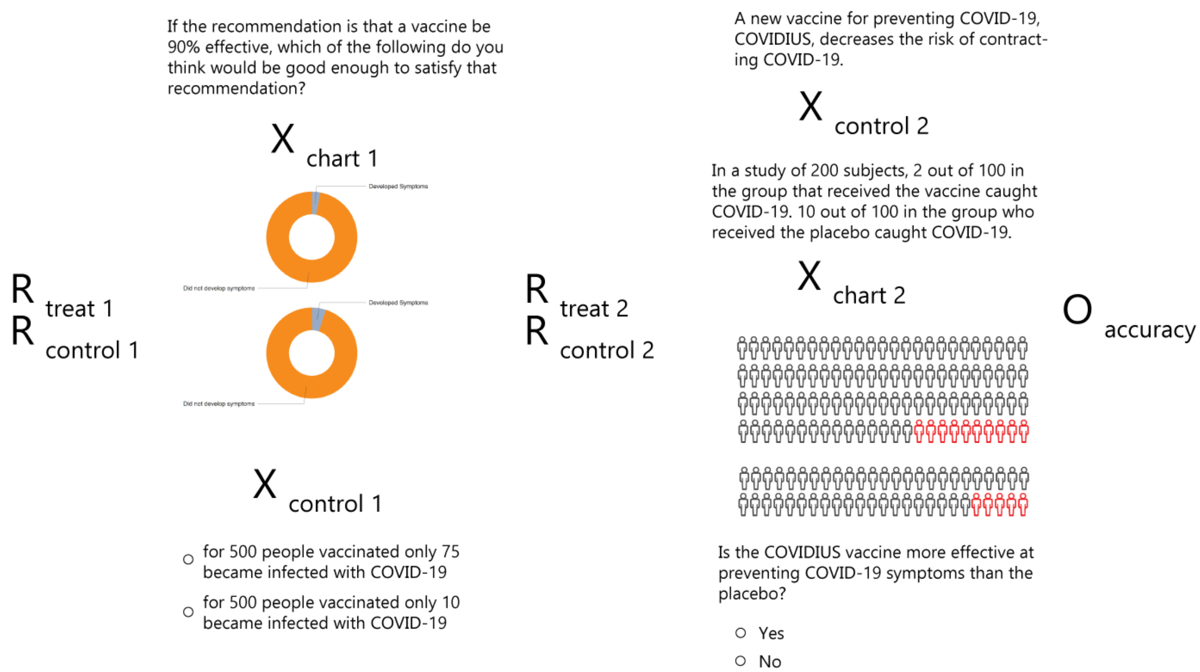

Figure 1 - A diagram demonstrating the overall RXRXO study design

Pie Chart Questions

In this arm, the pie chart was either right or wrong, depending upon the ratio presented. If a pie chart without labels is more effective than a numerical ratio at expressing efficacy then this provides us with alternative evidence to test the primary hypothesis that visual aids (regardless of format) were still more effective than either text or numerical ratios **(Figure 2)**.

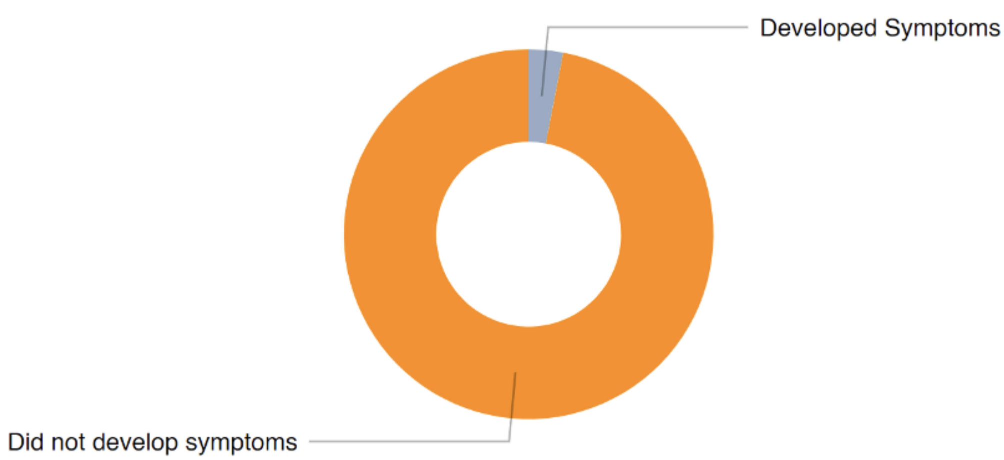

Figure 2a. - Demonstration of pie chart illustrating 97.5% efficacy.

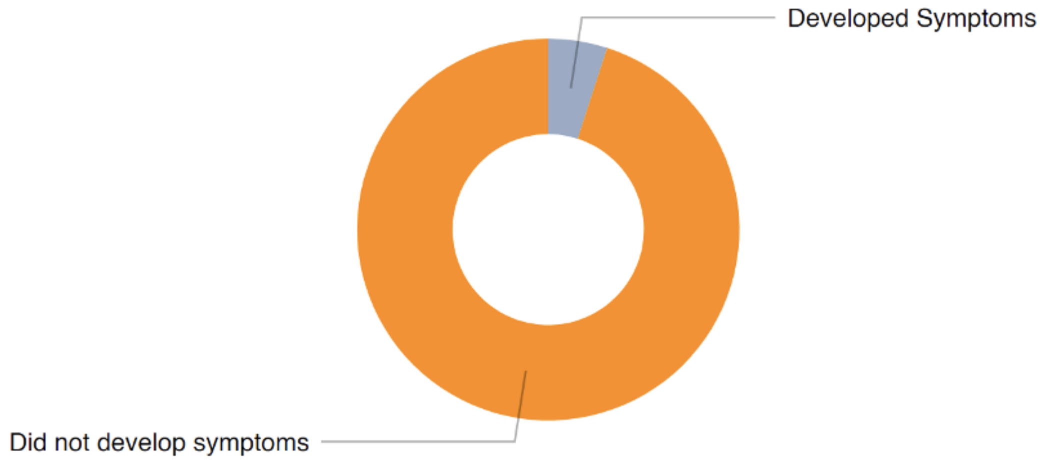

Figure 2b. - Demonstration of pie chart illustrating 95% efficacy.

Icon Array Questions

For the second stage, icon arrays were used as the primary visual aid. These icon arrays differ from the pie chart pairings in that they contain countable quantities of different colored icons and thus can be more precise than the pie charts **(Figure 3).**

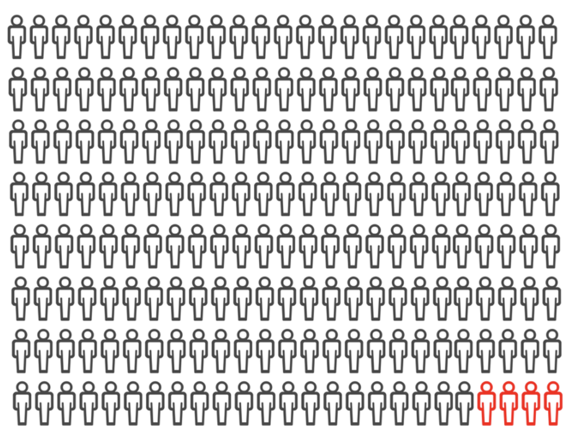

Figure 3a. A figure demonstrating the incidence ratio (red figures) of disease among subjects in the treatment group.

Figure 3b. A figure demonstrating the incidence ratio (red figures) of disease among subjects in the control group.

Control Variables

In addition to the study specific data, we also acquired a range of demographic data of participants. Numerous factors may affect the responses to our questions, namely how well subjects were able to understand them. These include age, level of educational attainment, and English as a second language (ESL) status. This data was used as control variables, and discussed further in the *descriptive statistics* section of this paper.

Approach to Analysis

Our design focuses on three core, and one ancillary type of analysis:

1. Descriptive statistics
2. Power Analysis → ancillary
3. Between group statistical comparison (Chi-Squared)
4. Regression analysis with and without potential control features

Pilot Data

We ran a pilot study using Qualtrics which received 18 valid responses. This survey was open for 10 days. The structure of the pilot survey contained two stages which tested a pie chart against a text-based numerical ratio representation followed by an icon array against a text-based numerical ratio. Demographic data was not collected in this stage.

In the pie chart control group, users answered correctly 60% of the time, while in the treatment group they answered correctly 83.3% of the time **(Table 1).**

	Right	Wrong	%Right
Control	3	2	60
Treat	10	2	83.3

Table 1 - Pie Chart Questions

For the icon array pilot, subjects answered the control question correctly 80% of the time. In the treatment group, they answered correctly approximately 86% of the time **(Table 2).**

	Right	Wrong	%Right
Control	8	2	80
Treat	6	1	85.7

Table 2 - Icon Array Questions

Based upon these pilot data, we determined that given adequate sample sizes, we could expect the proportion of correct responses in the icon array treatment group to be ~86%, while we might expect a proportion of correct responses in the control group to be ~0.80%. According to power calculations (using R's 'pwr.2p.test' function), at 80% power and an alpha of 0.05, we would require 611 subjects per group to detect a difference of this magnitude between groups (**Figure 4).** Moreover, based on data from the pilot, to detect the ~3% difference observed between the pie chart and icon array arms at the same levels as previously mentioned, we would need 2280 subjects (not plotted in the interests of brevity).

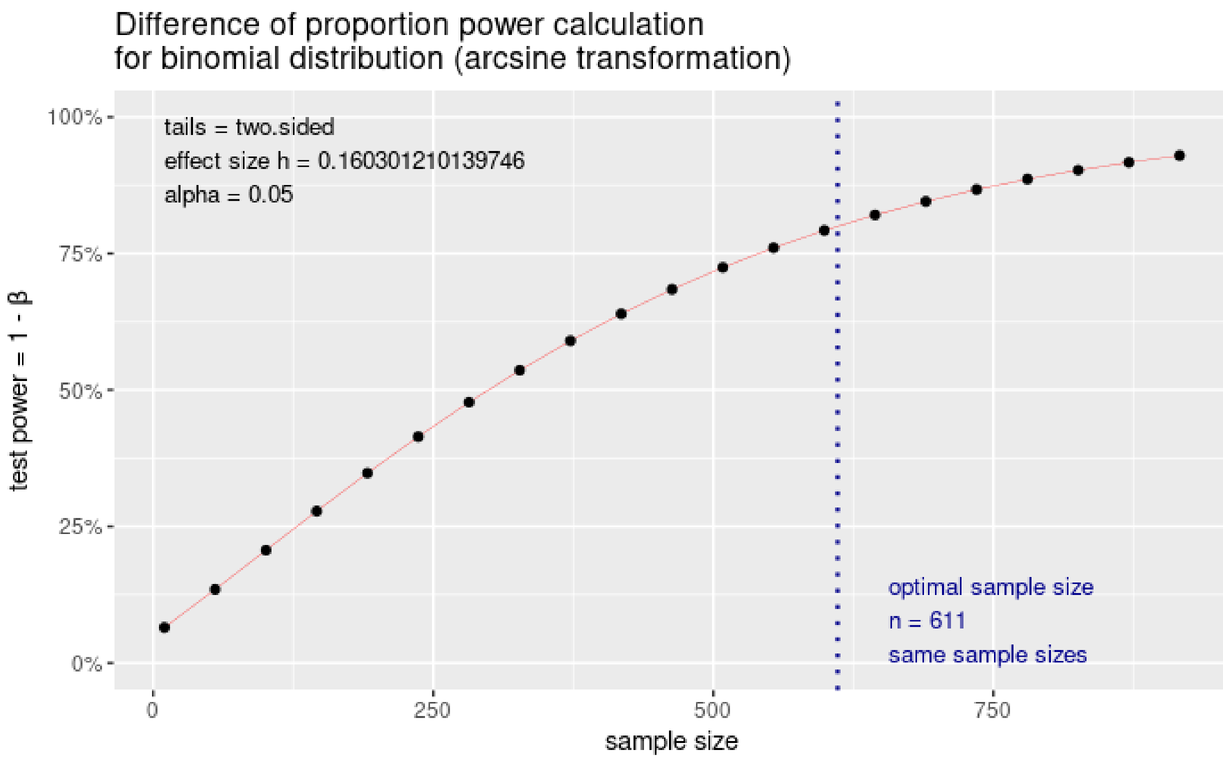

Figure 4 - A plot of the sample size and expected power of the experiment.

Results

The overall breakdown of our study is provided in **Figure 5**. In all there were 309 respondents. 45 subjects were lost to attrition (test-stage surveys, failure to complete). For the first question, 131 subjects were presented with the pie chart (treatment group) and 133 subjects were presented with the text control question. For the second question, 128 subjects were presented with the icon array (treatment group) and 135 subjects were presented with the text control question.

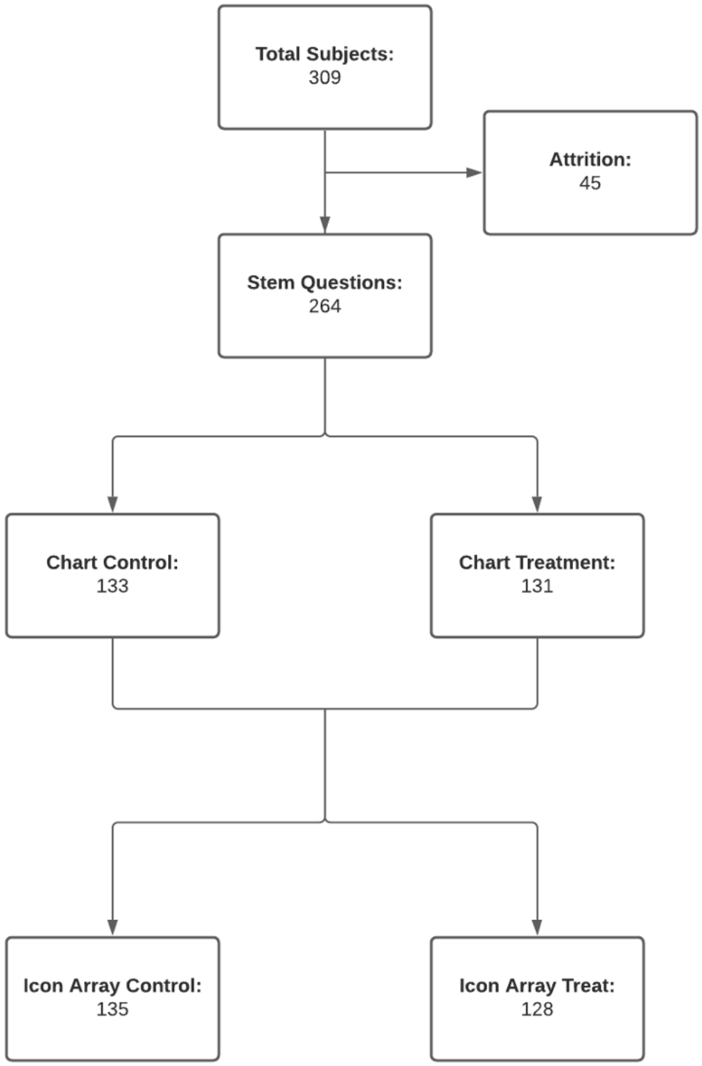

Figure 5. A diagram demonstrating the study flow and numbers of respondents for each respective control and treatment stage.

Stem Questions

Stem questions were asked at the beginning of the survey as a distractor for subjects, prior to being presented with the questions of interest **(Table 3)**.

	Question	n
1	Should Elementary Schools be Opened?	
2	Yes	114
3	No	150
4	How Many Days Should Schools Open?	
5	None	110
6	<5 Weekdays	116
7	Mon-Fri	38
8	At What Capacity Should Schools Open?	
9	25%	95
10	50%	125
11	75%	28
12	100	16
13	How Effective Should the Vaccine Be?	
14	85%-90%	98
15	90%-95%	96
16	>95%	69

Table 3 - Stem Question Responses

Descriptive Statistics

Regarding relevant demographics, participant age ranged from 18 to 54, with a mean age of 23.5 (standard deviation 5.5). Relevant demographics of the test subjects are presented, grouped by education level in **Table 4.**

	Educational Level	n	%Total	Mean Age	%ESL
1	Bachelor's degree	67	25.4	25.4	68.7
2	Advanced degree (Master's, Doctorate)	21	8.0	29.0	81.0
3	Some college	134	50.8	22.4	79.1
4	No college	42	15.9	21.4	71.4

Table 4 - Demographic Data

Approximately half of the respondents had completed some college (likely undergraduate students), a quarter had completed their undergraduate studies, while only 8% progressed onto further education, and 16% of respondents had no college. Interestingly, the majority of respondents in all educational groups were English as a second language (ESL) students, with over 4/5 of postgraduate respondents reporting ESL status.

Between Group Statistical Comparison

Table 5 summarizes the responses received for the first set of questions, where the subjects were exposed to either the text or the pie chart. 83.5% of those in the control answered the question correctly, while only 68.9% in the treatment did. As these data were categorical, consisted of counts between groups, and were generated through random sampling we used a chi-square test to determine if the proportion of right and wrong answers was significantly different between groups. The value was 7.42 and had a p-value of 0.006, indicating a statistically significant difference at the 5% level, between the control and treatment groups.

	Right	Wrong	%Right
Control	111	22	83.5
Treat	82	37	68.9

Table 5 - Outcomes for the Pie Chart Questions

Table 6 summarizes the responses received for the second set of questions, where the subjects were exposed to either the text or the icon array. 84.4% of those in the control answered the question correctly, while 93.8% in the treatment answered correctly. As these data were categorical, consisted of counts between groups, and were generated through random sampling we used a chi-square test to determine if the proportion of right and wrong answers was significantly different between groups. The value was 5.80 and had a p-value of 0.0160, indicating a significant difference at the 5% level.

	Right	Wrong	%Right
Control	114	21	84.4
Treat	120	8	93.8

Table 6 - Outcomes for the Icon Array Questions

The pie chart treatment group got a lower proportion of questions correct relative to control than the text or icon array treatment group, relative to their respective control group. These differences were significant at the 5% level, confirming our original hypothesis and the cited work concluding that pie chart data was more difficult to interpret.

We also tested the difference between the proportions of correct answers for the chart control and icon array treatment arms using a chi-squared test. The value was 25.54 and had a p-value ~0, indicating very high significance, beyond the 5% level. Nevertheless, the chi-squared hypothesis test cannot control for the additional demographic factors that we identified as potentially affecting outcomes, so analysis using regression may offer further insight by allowing us to include covariates.

Power Analysis

Given that we had 264 respondents, which is less than half of the required response rate to detect a significant effect of the predetermined magnitude at 80% power, we conducted an analysis to determine how well powered our experiments actually were. According to our analysis (conducted using the R 'pwr.2p2n.test' function for groups of unequal size), given our actual sample sizes and desired effect size (see **Figure 4.**), our experiment's power was only ~25% for the tests pertaining to the primary hypothesis, and ~10% for the secondary hypothesis.

Regression Analysis

The responses from pie chart and array groups were analyzed separately to study for an effect, as even though participants had to answer one question from each group, the randomization process meant that participants could either be presented with control or treatment group questions for either phase. In these models the proportion of correct answers for each phase (i.e. pie chart, icon array) was regressed against the proportion presented with treatment or control questions for each phase **(Table 7)**.

	Dependent variable:	
	q1_right	q2_right
	(1)	(2)
q1_control	0.835^{***}	
	(0.032)	
q1_treat	0.689^{***}	
	(0.043)	
q2_control		0.844^{***}
		(0.031)
q2_treat		0.938^{***}
		(0.022)
Constant	0.000	0.000
		(0.00000)
Observations	264	264
R^2	0.155	0.051
Adjusted R^2	0.149	0.044
Residual Std. Error (df = 261)	0.410	0.311
F Statistic (df = 2; 261)	23.949^{***}	7.021^{**}
Note:	*p<0.05; **p<0.01; ***p<0.001	

Table 7 - Base regression model

Table 7 demonstrates the results of the baseline regression model. The results are congruent with the results presented in the *Between Groups Comparison*. Here we can see that the coefficients for both the treatment and control questions, for each phase are significant and align with the proportion of correct answers we observed in the previous results section. Moreover, we can see that the coefficients are all associated with suitably small robust standard errors to render them highly significant. From this, we can calculate the average treatment effect (ATE) for pie charts as -0.146 (treat - control; 0.689 - 0.835) and for icon arrays as 0.094 (0.938 - 0.844). Nevertheless, other covariates may also affect these figures.

The second model is based on the first, but with categorical education level added as an additional covariate, in order to determine whether education level could further explain some of the variance in the dependent variable **(Table 8)**. From this model, we can study further the effects of education on the average treatment effect. Nevertheless, for the pie chart array control and treatment questions, it appears that education level did not exert a significant impact upon getting either of the control or treatment questions correct, as exemplified by the large robust standard errors and lack of significance for any of the coefficients aside from those from q1_control and q1_treat. Moreover, we cannot find any evidence of significant interaction between the control or treatment questions for the pie chart phase, with education level and it appears that the interaction term for the pie chart treatment with having no college was automatically dropped due to multicollinearity. Regarding the icon array phase, firstly, it appears that the interaction terms for all educational levels with q2_treat have been dropped due to multicollinearity, meaning that we are unable to fully interpret the effect of education on the treatment effect. The coefficients of 1.000 for both q2_control and q2_treat, (together with a constant of ~0.000, representing having an advanced degree due to R's dummy variable coding idiosyncrasies) for both the control and treatment questions reflect that those with advanced degrees got 100% of either control or treatment questions correct in the icon array phase, which we can confirm from the data table. There also appears to be a significant interaction between getting q2_control correct and having some college or a bachelor's degree, with coefficients less than those for the baseline (i.e. those with advanced degrees): The coefficient for those with no college is -0.211 ((-0.000) + (-0.174) + (-0.037)/constant + coefficient for education level + interaction term) lower than the baseline for q2_control. For those with some college it is -0.200 ((-0.000) + (-0.058) + (-0.142)) lower than the baseline for q2_control, and for those with a bachelor's degree it is -0.103 ((-0.000) + (-0.000) + (-0.103)). However, it is not possible to estimate ATE's in this model due to the dropped interaction terms for q1_treat and q2_treat, and due to the uncertainty of the calculated coefficients. The increased adjusted R^2 values (pie chart model: 0.154 vs. 0.149, icon array model: 0.061 vs. 0.044) indicate that the more complex model has better goodness of fit relative to the baseline regression model, however.

	Dependent variable:	
	q1_right	q2_right
	(1)	(2)
q1_control	0.818^{***}	
	(0.119)	
q1_treat	0.444^{**}	
	(0.169)	
q2_control		1.000
q2_treat		1.000
as.factor(Edu_Level)Bachelor's degree	-0.000	0.000
	(0.00000)	
as.factor(Edu_Level)No college	0.144	-0.174^{*}
	(0.209)	(0.080)
as.factor(Edu_Level)Some college	-0.000	-0.058^{*}
	(0.00000)	(0.029)
q1_control:as.factor(Edu_Level)Bachelor's degree	-0.047	
	(0.139)	
q1_control:as.factor(Edu_Level)No college	-0.162	
	(0.254)	
q1_control:as.factor(Edu_Level)Some college	0.069	
	(0.126)	
q1_treat:as.factor(Edu_Level)Bachelor's degree	0.185	
	(0.194)	
q1_treat:as.factor(Edu_Level)No college		
q1_treat:as.factor(Edu_Level)Some college	0.328	
	(0.177)	
q2_control:as.factor(Edu_Level)Bachelor's degree		-0.103^{*}
		(0.049)
q2_control:as.factor(Edu_Level)No college		-0.037
		(0.125)
q2_control:as.factor(Edu_Level)Some college		-0.142^{*}
		(0.058)
q2_treat:as.factor(Edu_Level)Bachelor's degree		
q2_treat:as.factor(Edu_Level)No college		
q2_treat:as.factor(Edu_Level)Some college		
Constant	0.000	-0.000
	(0.00000)	
Observations	264	264
R^2	0.186	0.089
Adjusted R^2	0.154	0.061
Residual Std. Error	0.409 (df = 253)	0.308 (df = 255)
F Statistic	5.786^{***} (df = 10; 253)	3.120^{**} (df = 8; 255)
Note:		*p<0.05; **p<0.01; ***p<0.001

Table 8 - Base regression with added education level factor

Based upon the paper by Garcia-Retamero et al. (2010), we anticipated those of increased age might perform worse when tasked with denominator interpretation compared to younger people. We decided to see if our results demonstrated this same phenomenon by including participant age (represented by 'Birthyear') in the regression model, together with an interaction term. The results are represented in **Table 9.**

	Dependent variable:	
	q1_right	q2_right
	(1)	(2)
q1_control	0.854^{***}	
	(0.121)	
q1_treat	0.779^{***}	
	(0.181)	
q2_control		0.942^{***}
		(0.140)
q2_treat		0.979^{***}
		(0.015)
Birthyear	0.000	0.006^{*}
	(0.000)	(0.003)
q1_control:Birthyear	-0.001	
	(0.005)	
q1_treat:Birthyear	-0.003	
	(0.008)	
q2_control:Birthyear		-0.002
		(0.005)
q2_treat:Birthyear		
Constant	-0.000	-0.186^{*}
	(0.00000)	(0.079)
Observations	257	257
R^2	0.142	0.057
Adjusted R^2	0.125	0.042
Residual Std. Error	0.410 (df = 251)	0.301 (df = 252)
F Statistic	8.292^{***} (df = 5; 251)	3.821^{**} (df = 4; 252)
Note:		*p<0.05; **p<0.01; ***p<0.001

Table 9 - Base regression with added age factor

By adding age as an independent co-variate in the regression (**Table 9**) the coefficients for the pie chart and icon array control and treatment questions remain highly significant. Birthyear does not seem to have a significant effect upon the proportion of right answers for the pie chart control or treatment questions, and there does not seem to be a significant interaction between the question or age. There is a significant effect of Birthyear on getting either of the icon array phase questions right (coefficient of 0.006, robust standard error of 0.003, $p < 0.05$), however the interaction coefficients between between age and q2_control and q2_treat are not significant. Moreover, even though the coefficient for Birthyear on the icon array questions was significant, it is substantially smaller than the coefficients for q2_control and q2_treat. Moreover, the adjusted R^2 values for both models are smaller than those of the baseline regression (pie chart model: 0.125 vs. 0.149, icon array model: 0.042 vs. 0.044).

Students who are ESL speakers may find reading English difficult, especially if the language is technical and regarding health topics[5]. Consequently, ESL students may naturally perform better on visual interpretations compared with interpreting text. We therefore decided to include whether the student was ESL in our regression model to see if it accounted for any of the variation in the dependent variable (**Table 10**). The coefficients associated with q1_control and q1_treat remain highly significant and in line with that of the baseline, though slightly different compared to the baseline model. We can also see that being ESL has no significant effect on getting either of the pie chart phase questions correct. Moreover, neither interaction term between being ESL and either q1_treat and q1_control is significant. For the icon array phase, similarly there is no significant effect of ESL on either getting the questions correct, nor is there any interaction between the independent variables. Moreover, it appears that there is some multicollinearity between the interaction term for q2_treat and another variable causing this coefficient to be dropped from the equation. The increased adjusted R^2 values were less than the baseline model (pie chart model: 0.139 vs. 0.149, icon array model: 0.041 vs. 0.044) indicating worse goodness of fit of the more complex model to the data.

	Dependent variable:	
	q1_right	q2_right
	(1)	(2)
q1_control	0.821***	
	(0.073)	
q1_treat	0.697***	
	(0.081)	
q2_control		0.762***
		(0.080)
q2_treat		0.929***
		(0.026)
as.factor(ESL == "Yes")	0.000	-0.038
	(0.00000)	(0.042)
q1_control:as.factor(ESL == "Yes")	0.017	
	(0.082)	
q1_treat:as.factor(ESL == "Yes")	-0.011	
	(0.095)	
q2_control:as.factor(ESL == "Yes")		0.098
		(0.088)
q2_treat:as.factor(ESL == "Yes")		
Constant	-0.000	0.038
	(0.00000)	(0.042)
Observations	264	264
R^2	0.155	0.056
Adjusted R^2	0.139	0.041
Residual Std. Error	0.412 (df = 258)	0.311 (df = 259)
F Statistic	9.482*** (df = 5; 258)	3.828** (df = 4; 259)
Note:		*p<0.05; **p<0.01; ***p<0.001

Table 10 - Base regression with added ESL factor

Due to the fact that education level was the only covariate that seemed to be meaningfully related to getting the proportion of icon array-phase questions one got right or wrong, we elected not to build a "fully saturated" model including the non-significant covariates.

Discussion

We conducted a randomized experiment whereby participants were asked to answer two questions successively pertaining to vaccine efficacy. They were either presented with control or treatment questions at each stage. We found that subjects performed significantly worse compared to controls when treated with pie charts, but performed significantly better compared to controls when treated with icon arrays. Based upon this, we cannot reject the primary null hypothesis ($h_{1,0}$), which specifically stated that *any* visual aid was better than a numerical ratio, as subjects performed worse when interpreting the pie charts compared to text-based ratios. We can however say that icon arrays are significantly better than text-based ratios when it comes to interpretation of numerical ratios. We can also conclude that icon arrays are significantly better than pie charts as a visual aid for ratio interpretation, thus enabling us to reject the secondary null ($h_{2,0}$) hypothesis in favor of the alternative ($h_{2,1}$).

Our regression results reinforce these findings. The simplest model matched the results we calculated from the initial analysis and chi-square tests. They demonstrated that coefficients for both control and treatment questions, for both the pie chart and icon array phases were highly significant allowing us to calculate the ATE with confidence. The fact that the proportion of correct answers was approximately the same for the control arms of both phases, despite being of a different format, indicates that the text questions were of similar complexity. Altogether, these findings point towards a potentially meaningful difference between the two treatment groups which should be further explored. Interestingly, the directionality of ATE for the pie chart phase actually defied that from our pilot data (which suggested that pie charts were better than text, which we based our effect size estimations on), indicate that study may in fact be adequately powered in spite of the results of our power calculations.

Only educational level seemed to have a significant effect and interaction with the two questions, but only with the icon array phase. Moreover, due to multicollinearity, we could not assess the interaction terms between educational level and q2_treat. One thing that we did observe was that certain groups answered both the icon array control (advanced degree students) and treatment (advanced degree students, bachelor's degree holders) questions correctly 100% of the time. This exerted a substantial effect on the coefficients for q2_control and q2_treat, which reflected answering the question correctly with 100% probability if one had an advanced degree. Altogether this observed phenomenon might mean that our study lacked enough respondents or perhaps the icon array control and treatment questions might have been too simple. It might also mean that the XLab participant sample, due to their relatively advanced education, may be non-representative of the same study-population used by Garcia-Retamero et al. (2010).

Age did not explain any of the variance in interpretation of either visual or text-based numerical ratios. The study by Garcia-Retamero et al. (2010) found that older people were more likely to demonstrate denominator neglect. Our results do not demonstrate this, but this could be because the overall mean participant age was early 20s, and no group had a mean age > 30. Therefore the sample of participants was not representative of a group that would be expected to demonstrate denominator neglect. Similarly, we did not find that ESL was significantly associated with getting either treatment or control questions correct or incorrect. We might expect that ESL students would have a harder time interpreting text written in English. The lack of association might be explained by the fact that the majority of XLab participants are in some way affiliated with UC Berkeley, an elite institute of higher education. Consequently, their English level might be good enough to negate any potential effect that ESL might have if we conducted this experiment in a different population.

Conclusions

We have successfully demonstrated that icon arrays are beneficial in helping people to interpret numerical ratios, and confirmed the work of cited authors. Icon arrays may also be better than pie charts, however we cannot say this with as great a degree of confidence. More work is needed to further confirm this postulation, as it may ultimately benefit the public education and health initiatives.

References

1. Garcia-Retamero R., Galesic M., Dhami M.K. (2012) Reducing Denominator Neglect. In: Garcia-Retamero R., Galesic M. (eds) Transparent Communication of Health Risks. Springer, New York, NY.
2. Garcia-Retamero R, Galesic M, Gigerenzer G. Do Icon Arrays Help Reduce Denominator Neglect? Medical Decision Making. 2010;30(6):672-684.
3. Cleveland W, The Elements of Graphing Data, Hobart Press, 1994
4. Tufte E, The Visual Display of Quantitative Information, Graphics Press, 1983, p. 178.
5. Rocio Garcia-Retamero PhD* and Mandeep K. Dhami PhD (2011) Pictures speak louder than numbers: on communicating medical risks to immigrants with limited non-native language proficiency. *Health Expectations*. Blackwell Publishing Ltd , Oxford, UK.