

## Science Fair Organizers' Guide

Thank you for organizing a science fair and using this system for judging. We would like to judge science fair projects accurately, consistently, and precisely so students can get the best possible feedback. We would also like the process of organizing all of the judge information to be as intuitive as possible for you. Everything you need to conduct handle judging is available at [PerplexingQuestions.org](http://PerplexingQuestions.org). Though you don't need to, I would appreciate a copy of your data file so that I can add it to our analyses of judging. I would appreciate any feedback you have after using this so that we can improve it from year to year.

### *Assigning Project Numbers*

For each of your science fair projects, give it a unique project number. It could be as simple as "3" or it could be some more elaborate code you use like "mp-03." Whatever you choose, please fill in the "project list" tab in the excel file with each project number and each students' name (e.g., "Kevin Grobman"). This can help you keep track of data in the future. It is also very important for science fair research because we would like to see how students improve across years and so we can connect science fair data to other surveys students complete. The excel file is prepared to handle 300 projects in each category. If you will be judging more projects, you can contact me and I will prepare it to hold more data for you. If you realize too late that the excel file is insufficient, you can always do it yourself. With the worksheet blank, highlight all of row 3 (by clicking the 3 to the left of the first cell), copy it. Now click and scroll down for as many projects as you have (e.g., if 500 projects, scroll to at least 503. Then paste. I limited the spreadsheet to 300 projects per category in order to keep the excel spreadsheet to a manageable size (kb).

### *Assigning Judge Numbers*

For each of your science fair judges, please collect some basic information and assign them a judge number. Within a category (e.g., chemistry, biology), number the judges starting from 1. It's an arbitrary number, so it does not matter what order you put judges in. If you consider awards separately by grade (e.g., middle school awards and high school awards), then you have two choices. You can have a separate copy of the excel sheet for each grade. In this case, a judge for the middle school division may also be a judge with a completely different number in the high school division. Alternatively, you can put both middle and high school together. In this case, judges should get multiple columns if they judge both grade levels. For example, I might be "judge 10" when I evaluate the middle school social science projects and "judge 11" when evaluating the high school social science projects. The excel formulas have already been set up to adjust for judge bias (z-scores) and when you give somebody multiple judge numbers, you are saying that their evaluations should be scaled separately. For example, you would not want to judge

2 high school students on the same scale as 10 middle school students evaluated by the same judge. Please fill in the “judge list” tab in the excel file with each judge number, category, and grade (which should only appear once), each judge’s name (which may appear multiple times), the judge’s prior years as a judge (e.g., 3), and the judge’s highest degree (e.g., Ph. D. Physics). The excel file is prepared to handle up to 28 judges per category.

### *Entering the Judges’ Evaluations*

On the left side of each judge worksheet (e.g., ctg01) is a place to enter the project number (arbitrary code you used), category (e.g., chemistry), and grade (e.g., middle, high). Once you type in a project number, notice how it automatically appears in the gray columns across the entire spreadsheet. This is to help you match projects and judges. Judges will provide you with 4 numbers about each project: *Project Idea*, *Present & Frame*, *Rigorous Method*, and *Interpret Results*. Notice how each judge has a big column (incorporating 4 little columns) on a judge worksheet. The little columns, I, F, M, R match the judge’s ratings in the same order as the judge’s sheet. Find the project’s row and the judge’s column (just like you would find the product using a multiplication table). Then enter the 4 numbers. The worksheet “ctg\_sample” illustrates what this should look like. Each worksheet can only hold 28 judges. In order to score science fair projects correctly, you need to put all judges for each particular project on the same worksheet. For example, you could *not* have judge 28 (on sheet “ctg01”) and judge 29 (on sheet “ctg2”) evaluating projects 102-113. Instead, leave judge 28 blank and make these two people judges 29 and 30.

### *Collating Judges’ Evaluations*

On each judge sheet, scroll all the way to the right (passed all the columns in gray) to the 7 columns with light purple headings. Excel has automatically tabulated results for you about each science fair project! Highlight all of this information (*excluding* the purple headings) and *copy* it. Do not *cut* it. Now go to the last worksheet titled “scoring sheet.” Notice how it has the same purple headings. *Paste Values* into the corresponding place. Do not simply press *paste* because that will copy formulas. To paste just the numbers, click the excel heading “edit”, choose “paste special ...”, click the dot for “values”, and then press “okay.” Go to the next judge page and repeat the same process but now pasting the values in the row following the previous ones. If you had to create another excel spreadsheet, open it up and paste those values into a single scoring sheet. You can now sort the information in all sorts of ways to help you choose science fair winners. For example, if you are going to give awards to the top 3 projects in each category and grade-level, sort by “category” and “grade” and finally by the scoring variable (descending so the top performers are on top). To sort, choose excel menu “data”, choose “sort”, and then you can sort by three things in a specific order. A later

section in this manual will explain what each composite score is so you can decide which way you would like to judge science fair projects. Sort by the column you choose.

### *Important Cautions when using Excel!*

★ *Do not* enter anything in cells when you have no information for them. *Do not*, for example, type “0” or “n/a” when a student did not show up for the fair. MS Excel calculations are finicky and they will interpret what you type as though it’s a real judge rating. Just leave cells with missing data blank.

★ *Never cut* and paste in the excel spreadsheet. You *can copy* and paste. MS Excel has an odd underlying logic so that when you *cut* and paste it will change the underlying calculations in bad ways.

### *What are the Composite Scores?*

Notice how each science fair projects has 5 composite scores. For all of them, 100 is the best possible score and 0 is the worst possible score. All of the scores give equal weight to each judge who rated that project. In most cases, the typical project will score around a 50. If you would like to skip reading this nitty-gritty, I recommend using the “overall” composite score for your awards. Below that are some issues to consider when choosing a scoring method.

*Should all 4 judge ratings be treated equally?* Judges gave 4 ratings: *Project Idea* and *Present & Frame*, *Rigorous Method*, and *Interpret Results*. One way to score is to say that each of these ratings is equally important. Judges will probably start to think that simply because they make four ratings. Another possibility is to say that the first two ratings are worth 1/3 less than the last two ratings. That is, *Project Idea* and *Present & Frame* are each worth 20% and *Rigorous Method*, and *Interpret Results* are each worth 30%. That would be more consistent with the feedback we got from judges because more judges mentioned the latter as important. If you would like to weight the 4 ratings equally, choose either *EQ* composite score but if you would like to weight them like the feedback we received, choose either *WE* composite score.

*Should judge scores be adjusted for bias?* Judges are supposed to follow the benchmarks provided, so they should all be thinking, for example, that a “7” means exactly the same thing. To the extent judges take the benchmarks seriously, you may want to use their “raw” ratings (numbers without being corrected / ‘cooked’). However, judges can have biases. Some judges might be very strict (giving lots of low ratings and few high ratings) and others might be lenient (vice-versa). You can especially see this if you have many judges rating the same projects. If you feel judges were taking benchmarks seriously and one judge giving higher rating than another is primarily because they

judged better projects, then choose either  $R$  composite score. If you feel each judge rated projects of similar quality and differences between judges were their bias, then choose either  $Z$  composite score.

*What is  $Z$  and how is it adjusting for bias.* We can statistically adjust for bias by assuming each judge rated projects of equal caliber and of an equal range of caliber. That means each judges' ratings should have the same average and the same standard deviation. Excel is shifting, stretching, or crunching each judges' scores so that they give an average rating of 5 and so that a score 2.5 standard deviations away is a 0 or 10. Ratings even further from the norm are made either 0 or 10 to avoid one judge overly influencing the results. For those with a background in statistics, this is done with a linear transformation of the z-score (thus the  $Z$ ).

*What is the overall composite score?* The overall composite score is the average of the 4 other scores. If you answer the 2 previous questions by saying, "yes and no", then you should use this score. It's saying that there is reason to consider the 4 dimensions equally but also a reason to weight them. This score compromises by saying the first 2 dimensions should be worth 22.5% and the last 2 dimensions should be worth 27.5%. It's also saying that some judges really are stricter but judges also got projects of different caliber. This score makes half of each rating the 'raw' score and half from the z-score.

### *Giving Feedback to Students*

Students might like to know why they missed getting an award and how they can improve in the future. Some may like to know what set their project apart and gave them among the best scores. To give them feedback, find their row in the original judge sheet. Scroll all the way over to the right. Just before the purple headings are two "average judge" headings: one for raw scores and one for average z-scores. Look at their lowest number of the 4 numbers under "average judge." If it is below about 3.5, tell them that this dimension is one that they have the most room to improve. Their score is below what judges felt was the minimum standard. If their lowest score is above about 4.5, tell them that they did reasonably well and it's just the high quality of other projects that stopped them from getting an award. Nevertheless, tell them that the place they can improve the most is on their lowest dimension. If students score above about 6.5 on any dimension, highlight that to the student as a place that judges were especially impressed and felt they did better than they would have expected from a science fair project. One goal of this new judging system is so that, in the next few years, we will be able to make calculations from this excel form automatically turned into qualitative feedback on a letter for each student to see their strengths and weaknesses.