Science Fair Judges' Guide

Guide to Accurate & Precise Rating

Thank you for being a judge in the science fair. We would like to judge science fair projects accurately, consistently, and precisely so students get the best possible feedback. We would like the judging process to be as intuitive as possible, so you can fully express the quality of the science fair projects without becoming "bogged down" in recording too many numbers. It would be prohibitively time-consuming for every judge to interview every student thoroughly, so we will have a first round of judging based on materials submitted by the student.

In the first round of judging you rate projects on 4 dimensions (idea, presentation, method, & results). Within each dimension you rate with 4 benchmarks of quality: 0 (needs improvement), 1 (developing skills), 2 (solid achievement), or 3 (exceptional). Thinking like a teacher you might say 0 is C or lower, 1 is B, 2 is A, and 3 is A+. I interviewed dozens of judges from diverse disciplines about how projects excel or fall short. The consensus formed our 4 dimensions. Benchmarks similarly represent how scientists from diverse disciplines described varying degrees of quality. Given this judging rubric emerged from scientists' intuitions, I hypothesized, and found evidence for, greater inter-rater reliability than with previous commonly used rubrics.

Though there is no precise rule for how often you apply each quality benchmark, a regional science fair typically has about 17% needs improvement (0), 33% starting to develop skills (1), 33% solid achievement (2), and 17% exceptional (3). If your ratings are far off, please step back and reconsider your approach. For example, if you rate one-third as poor, you may be overly strict and if you rate one-third as exceptional, you may be overly lenient. If you rate two-thirds with solid, you may be missing important qualities worth differentiating.

Each of the 4 dimensions is an independent rating. It is possible for you rate a project as "exceptional" in one dimension and "needs improvement" in another. Please do not try to adjust for circumstances. For example, it's unlikely middle school students have the same mathematical skill as high school students. But you should not adjust the "results" dimension to students' grade level; we already determine winners separately by grade level and it's helpful to see the growth in students' abilities through your ratings.

Within each dimension, you may feel something particularly important within your scientific field is not explicitly mentioned. Conversely, you may feel a description does not precisely describe work within your field. We do not expect you to take the rubric literally. Instead, consider the spirit of the rubric and adjust accordingly. For example, though the sciences of Physics, Geology, and Psychology share the same foundations as sciences (e.g., empiricism), their research looks quite different (e.g., tools, mathematical models). Even within fields, you may need to adjust the rubric for more applied or pure projects. For example, sometimes you may substitute "hypothesis" with "problem to solve" (e.g., engineering) or "proof" (e.g., mathematics). You are a judge because of your expertise; you know best what scientific rigor means in your discipline.

Below we describe the 4 dimensions with benchmarks in detail. Second, we describe some common challenges to judging well and offer suggestions.

Project Idea

What is the overall quality of the project idea: the research question, the hypothesis to be tested, the problem to solve, or the proof or computer program to be constructed?

Needs Improvement (0; C): The project lacks a clearly stated project idea or has an idea inconsistent with science (e.g., unfalsifiable). Conversely the project idea may be so commonplace it appears the student lacks independent motivation to think scientifically and instead simply copied it from a website (e.g., ScienceBuddies.org).

Developing (1; B): The project has a clearly stated project idea. While it's a commonplace hypothesis and possible to find similar ones online, the student did try (e.g., small twist on something from ScienceBuddies.org).

Achievement (2; A): The project has a clearly stated project idea demonstrating independent effort and decision-making from the student. The hypothesis may not be especially creative, but comes out of the student genuinely asking, "why?" or "how?" about something he or she cares about.

Exceptional (3; A+): The student was independently motivated and clearly thought deeply about the hypothesis. The hypothesis might be especially creative. The hypothesis may have begun with a passion to solve a pressing real-world problem. The hypothesis may have begun as a vague idea of the student's, but then he or she found mentors to help refine it to address a genuinely meaningful question within the field.

Presentation within the Scientific Context

What does the student know about the relevant science and about how to present science? How well did he or she execute this knowledge by finding background research (e.g., literature review) and creating a poster?

Needs Improvement (0; C): The poster was sloppy and disorganized. No consideration is made for what scientists know about the topic.

Developing (1; B): The poster was put together with care and shows reasonable organization on a "global" level (e.g., introduction, method, results, discussion). However, the student did not carefully plan how to best convey information within each section (e.g., table vs. figure). The project references relevant scientific concepts, but only in ways you would expect to be in standard school curriculum.

Achievement (2; A): The poster was put together with care, is well organized, and care was taken within each section to convey information well. The student clearly learned more about the relevant scientific concepts than would be expected in standard school curriculum and the student presents their study in the context of prior scientific work (e.g., by citing secondary sources).

Exceptional (3; A+): The student learned about the relevant science to a depth clearly beyond standard school curriculum. He or she has understood sophisticated lay person sources (e.g., Scientific American). He or she "muddled through" some primary sources enough to cite within the poster. The poster showcases the scientific process at its best with attention to detail. The poster is well organized (e.g., introduction, method, results, conclusion) and each section shows careful consideration (e.g., clear rationale behind the design of figures).

Research Methodology

How did the student find out if his or her hypothesis is supported by the evidence? Was it an appropriate method? Was care taken to avoid common methodological mistakes? Did the student consider different methodological approaches? Was sufficient data collected?

Needs Improvement (0; C): The method does not align with the hypothesis. There are serious methodological errors (e.g., confounds, non-random assignment). Few measurements are made. Measurements taken lack sufficient care.

Developing (1; B): While the student does not understand the nuance about the many choices made, the research methodology is appropriate for testing the hypothesis. The operational definitions are adequate, if obvious. The data collected may or may not be sufficient to test the hypothesis, but the amount of data collected demonstrates a genuine effort to test the hypothesis.

Achievement (2; A): The student conducted a study carefully following the scientific method with good methodological choices, good measures for operational definitions, and sufficient data. The method is well executed, if perfunctory. The student can explain methodological choices but may not realize other choices were possible.

Exceptional (3; A+): The student repeatedly measured the phenomena so sufficient data was collected to test the hypothesis. The research method addresses the hypothesis while avoiding common mistakes (e.g., confound, non-random assignment). The student can articulate a rationale for why he or she chose one research method over another (e.g., observation vs. experiment; between- vs. within-subject). The method is either impressively sophisticated or creative. For example, the student may have used specialized equipment to measure an operational construct or found a creative way around a methodological stumbling block.

Results & Interpretation

How does the student summarize his or her results? How does he or she connect the results to the hypothesis? How does he or she interpret the results in the larger scientific context and how does he or she see limitations in how much we can interpret the results?

Needs Improvement (0; C): The project either does not present results or simply provides all the raw data without summary. Whatever summary a student presents, it is either inappropriate to test the hypothesis or incorrect (e.g., math mistake when calculating average).

Developing (1; B): The project includes a summary with basic mathematics (e.g., mean) though the student may have trouble seeing how the summary connects with the hypothesis. The student does not necessarily understand what the results mean for the hypothesis (e.g., over-interpretation of chance).

Achievement (2; A): The student carefully considered how to summarize data with basic mathematics (e.g., mean, mode, median, standard deviation). The student connects the results with the hypothesis reasonably, though possibly not perfectly, and has some understanding of the generalizability and limitations when interpreting the results.

Exceptional (3; A+): The student carefully considered how to summarize data to test his or her hypothesis. The students' analysis goes beyond the basics and uses more sophisticated mathematics (e.g., t-test, correlation, best-fit equation). The students knows if the results support or refute his or her hypothesis. The student recognizes how much you can generalize your interpretation of the results and how the interpretation may be limited (e.g., correlation vs. causality).

Stumbling Blocks while Judging

Despite your expertise, some situations make comparisons between projects challenging. Most of these stumbling blocks can be summarized with the cliché of comparing apples to oranges.

Projects on Topics You Know Well: Experts in a particular science rarely know every topic equally well. In fact, you may see a project on the area you're most passionate about. Ironically, judges sometimes inadvertently rate projects closer to their expertise more negatively than other projects. When you know a topic well, you detect subtle mistakes more easily. If a student presents something that is blatantly incorrect in a way you would detect in even topics you do not know especially well, the error should certainly lower the student's score. However, be mindful if an error you see is subtle in a way you might not notice in other projects. Errors should always lower your rating, but do not deduct as much for subtle errors as fundamental errors of science.

Different Amounts Adult Help: Sometimes students complete projects entirely on their own. Sometimes students get help from non-experts like their parents. And sometimes students work with experts, like professors. When rating projects, you need to look beyond adults' contributions and rate the child's work. When children figure things out on their own, their scores should rise for showing independence. When children find an adult who can help them with aspects beyond their ability, their scores should rise for showing skill at finding collaborators and learning something more advanced than their grade-level. In fact, as someone who has helped many students with their projects and watched several compete at the international fair (ISEF), almost every project that places at ISEF was by a student who had some adult expert help. In sharp contrast, when adults do things for the child, the child should not get credit for the adult's effort. Distinguishing good and bad help usually comes through most clearly during interviews. Here are some examples to help you differentiate "good" and "bad" help.

Bad Help from Adults (lower rating)	Good Help from Adults (raise rating)
An adult had something that interested him or her and told the child it would make a good project.	The student approached a professor studying something he or she likes and the professor helped refine the idea into a scientifically important hypothesis.
An adult took measurements because they had the skill to be most precise (aside from using dangerous tools that children should never be expected to use).	A child wanted to measure something and he or she found an adult who recommended a precise tool. The adult and child measured together with the child doing as much as possible even if that meant less precise measurement.
An adult took the student's raw data and created polished results beyond the child's ability to understand.	The student did analyses they knew about and sought adults with skills to produce better results. Even if the adult did the analysis, the student understands what they did and why.

The Human Element of Judging

Thus far we have focused on your role as a judge inherent in the word, "judge." One key role you have is accurately and fairly rating projects. But your other - and actually much more important role – is as inspiration. Remember students have invested multiple months in their projects – possibly longer than any single school assignment – and now they get to meet actual scientists who study the same thing. That's very exciting, and very intimidating! What you say and do can have a life-long impact. If you're excited to meet students, smile, ask encouraging questions, and give constructive feedback, then you might be helping a budding scientist grow! In contrast, if you're sarcastic, look bored, or show contempt, you might squash a career path they'd consider. While your ratings should always be honest and accurate, your interactions should always err on the side of being gentle.

Any time you can share something positive with a student, please do! You can compliment their effort, a beautiful poster board, or how well they explain a scientific concept. Share enthusiasm, like how one of your colleagues down the hall, and right here on this campus, studies the very same thing! Make generous comparisons (e.g., wow, I didn't learn that until my first year of college). Whenever possible, turn what could be negative feedback into an opportunity. For example, looking at a student's results, you could say, "you should have done … ." But notice a statement like this emphasizes where the student fell short. Instead you could say, "I like how you figured out …, and if you'd like to take your project to the next level next year, try … ." Notice how now you're inspiring them to keep pursuing excellence – just like we as scientists know our theories are our best estimate awaiting falsification and each study is a step toward a true understand.

You are truly the most important factor in the success of the science fair. Thank you for offering your time and expertise to help the next generation of scientists!

Kevin Grobman (kgrobman@csumb.edu)