Data Generation in the Discovery Sciences – Learning from the Practices in an Advanced Research Laboratory

Wolff-Michael Roth, University of Victoria

Abstract

General scientific literacy includes understanding the grounds on which scientific claims are based. The measurements scientists make and the data that they produce from them generally constitute these grounds. However, the nature of data generation has received relatively little attention from those interested in teaching science through inquiry. To inform curriculum designers about the process of data generation and its relation to the understanding of patterns as these may arise from graphs, this five-year ethnographic study in one advanced research laboratory was designed to investigate how natural scientists make decisions about the inclusion / exclusion of certain measurements in / from their data sources. The study shows that scientists exclude measurements from their data sources even before attempting to mathematize and interpret the data. The excluded measurements therefore never even enter the ground from and against which the scientific phenomenon emerges and therefore remain invisible to it. I conclude by encouraging science educators to squarely address this aspect of the discovery sciences in their teaching, which has both methodological and ethical implications.

Keywords Inquiry • Discovery Sciences • Data Generation • Data Interpretation • Graphs • Graphing

Roth, W.-M. (2013). Data generation in the discovery sciences–learning from the practices in an advanced research laboratory. *Research in Science Education, 43*, 1617–1644.

> The science education literature indicates that students who are involved in collecting their own data often do not understand the fundamental reasons for doing so and are often more concerned with following laboratory protocols and getting "the right" data. As a consequence, "hands on" activities are often not "minds on" activities. (Cobb & Tzou, 2009, p. 169)

Inquiry – both in more constrained, planned and structured investigations and in more open, unstructured real-world settings – has long been a trademark of science education generally (e.g., Jordan, Ruibai-Villasenor, Hmelo-Silver, & Etkina, 2011; McElhaney & Linn, 2011) and of reform-based science in particular (e.g., Carlone, Haun-Frank, & Webb, 2011). It is not surprising, therefore, to find (as of February 1, 2012) 592 out of the 6,294 articles in six major science education journals included in the ISI Web of Science database using the term "inquiry" as identifier, in the title, or in the abstract. Inquiry is a trademark of science education even though in practice it often is pushed to the margins in the face of high-stakes testing and even though teachers often believe that inquiry should be encouraged (Nargund-Joshi, Park Rogers, & Akerson, 2011). But there are studies that do in fact report tremendous achievement gains when students engage in inquiry; and these gains are larger when teachers are more experienced in teaching in the inquiry mode (e.g., Fogleman, McNeill, & Krajcik, 2011). Even beginning teachers may find themselves surprised by the positive outcomes from their supervision of extended experimental investigations (Ritchie et al., in press). There is further evidence that scaffolding promotes teachers' competencies to guide students through open-inquiry projects, "especially the ability to know when and how to give students a well-balanced combination of 'structure' for open-inquiry learning and sufficient 'space' for that" (van der Valk & de Jong, 2009, p. 829). Others outright reject inquiry – at least the forms in which there is little guidance (Kirschner, Sweller, & Clark, 2006). Defenders of inquiry list – among its ideal benefits – that students not only learn how science operates from designing experiments to generating data and to the ultimate reporting of results but also outperform guided-inquiry students on a number of variables (e.g., Russell & Weaver, 2011; Sadeh & Zion, 2009). Even the youngest students learn from open inquiry, as shown in a study related to authentic inquiry at the primary (K–2) level (Akerson & Donnelly, 2010). Defenders of open ("authentic") inquiry further contend that inquiry works precisely because of the high levels of control students have over the task and task definition (e.g., Feldman & Pirog, 2011). In a study that drew on adapted primary scientific literature, teacher education students did in fact learn by engaging in and talking about scientific inquiry, their pedagogical content knowledge, and their subsequent curriculum designs (Falk, Brill, & Yarden, 2008).

Although science educators have shown interest in inquiry approaches to science, the question of the nature of data generation is less frequently raised: only 29 of the 592 articles on inquiry also show up when the search term is "open inquiry." That is, data generation appears to be of less interest even though science educators have noted the "sophisticated coordination among theories, phenomena, data, and data collection events" (Apedoe & Ford, 2010, p. 165) and even though others articulated the role of anomalous data in knowledge generation (Chinn & Brewer, 1993). Such coordination has also been observed among more advanced ("expert") undergraduate students but not among their "novice" peers (Jordan et al., 2011). Although Apedoe and Ford suggest that the complex interactions should be broken down and taught separately, several ethnographic studies among research scientists suggest that these interactions between theories, phenomena, data, and data collection events are irreducible (e.g., Roth, 2003, in press; Roth & Bowen, 1999b). One study did in fact report very different forms of behavior when students were provided with data versus when they decided about the goal of their research and generated data themselves (Roth & Barton, 2004). In the controlled context with data provided, students did not engage with the task, by and large suggesting that they "don't know how to do it." On the other hand, intense engagement and highly competent practice in plotting data were observed when the students were in complete control over their investigations and the way of representing their results. Similar differences were reported in another study, where eight-grade students, who had been collecting real data in investigations of their own design, outperformed pre-service teachers, who had already completed bachelors or Master's degrees in science, on data interpretation (Roth, McGinn, & Bowen, 1998). Something in the data generation process appears to allow students to know what to do with data and how to do it.

A recent review of studies in inquiry in simulation contexts suggests that there is little evidence for drawing conclusions about how the design of investigations generally and data collection specifically might advance science learning and understanding (Scalise et al., 2011). The present study was designed precisely to address this issue by seeking to understand the data generation process in the discovery sciences that precedes – but probably is integrated with or reflexively tied to – the interpretation of data for the ultimate goal of better informing science education practice about the data collection process typical of the discovery sciences.

Background

The media rarely provide information about the variations in the data, about laboratory contexts, or about what has not been included in the measurements or analysis. This is important because the "details of laboratory work, and of the visible products of such work, are largely organized around the practical task of constituting and 'framing' a phenomenon so that it *can* be measured and mathematically described" (Lynch, 1990, p. 170). It may therefore not surprise to find scientists who critique the kind of representations with which students are presented in their courses. Thus, upon seeing a graph of ideal birth rates and death rates to model the temporal dynamics of a population – as can be found in any introductory university textbook on ecology – an internationally known marine ecologist suggested: "You're never gonna find a data set that looks like this. This is a theoretical model, it's based on, you know, nice mathematics and equations, and it's the way we think the world probably works" (Roth, 2001, p. 14). He further suggested never having seen a data set that would contain a perfect relation, because "in the real world, [there] is a constant fluctuation" (p. 14). That is, to understand the claims made in the scientific literature or in the popular media, we need to know how the laboratory contexts *might have shaped* the data collection to understand what is included in and what has been excluded from the data mobilized in support of the scientific claims. Without such knowledge, even professors lecturing undergraduate classes may erroneously relate graphs and the phenomena in the world that these are intended to represent (Roth & Bowen, 1999a).

An understanding of the process of data generation generally and that of the data that underlie scientific claims should be of interest not only to those science educators interested in producing more scientists but also to those who focus on general scientific literacy. Across the media, we are confronted daily with the results of yet another medical study suggesting that eating more rolled oats, kale, or fish (oil) diminishes the incidence of certain medical conditions. Being able to understand such reporting is an important goal of science education (e.g., Aberg-Bengtsson & Ottosson, 2006; Garli & Rule, 2009). For example, while writing these words, I was directed to an "Infographics" with the subtitle "Sitting is Killing You" (Medical Billing and Coding, 2010). One of the panels reads: "Sitting increases risk of death up to 40%," specifying that "Sitting 6+ hours per day makes you up to 40% likelier to die within 15 years than someone who sits less than 3." Should I be alarmed giving the fact that I am sitting at my desk for 8+ hours, sometimes up to 14 hours? On what data are such claims based? More important is this question: "What was *not considered* as data in stating these claims?" To understand such claims generally and the limitations thereof more specifically, one actually needs to know more, for example, what data distributions look like (Cobb & Tzou, 2009). Take the example in Figure 1a. It shows that there are many individuals sitting less than 3 hours a day who die before individuals (all those shaded dark grey, representing the overlap of the two distributions) who sit 6+ hours a day. If the same graph represented the efficiency of a drug, the distributions might show the relative benefits of a placebo (left) and a drug (right). In this case, there would be many for whom the placebo worked just as well as the drug. A study among eighth-grade students suggests that students better learn to interpret when such distributions are included (Cobb & Tzou, 2009). They might ask, "What is the percentage of people sitting 6+ hours who have a higher risk?" Thus, participating in the construction of data will "ensure that the students [are] actually analyzing data rather than merely manipulating numerical values" (p. 167). Other forms of information, too, appear in the media and yet are frequently misunderstood (e.g., Figure 1b). For example, a psychologist might find that the correlation between *IQ* and science achievement (*Ach*) is given by the equation $Ach = m \cdot IQ + k$. How much should a science teacher be concerned? Teachers need to know why a student underperformed so that they could address his/her needs – and these needs are treated in the statistical model as error. That is, what the teacher needs to address a student's need *has been deleted* from this scientific model of achievement.

A typical instance of a lack of understanding or misrepresentation could be observed during a recent election for a Canadian parliament. The electoral system is based on a "first-past-the-goal-post" system, that is, the totality of individuals who have had the most votes in their riding constitute the parliament. Prior to Election Day, journalists were talking about the surge of the New Democratic Party and the possibility of it forming the new government. In such presentations, it is neglected (or misunderstood) that a party may have the most votes across the nation and yet not be the ruling party; or a party with less than 40% of the popular vote may have most seats. That is, the parliament is not *representative* of the Canadian population in a strict sense. But to understand this, the underlying mathematical relations between actual votes (data) and representation (number of seats) need to be understood. In this instance, there are nonlinear relationships. What voting results do not represent are all those who have cast spoiled ballots or who have not voted at all. Being literate about data, variations within the data, and just what constitutes non-data ought to be an important facet of a citizen in a knowledge-based society.

««««« Insert Figure 1 about here »»»»»

One of the shortcomings of the Cobb and Tzou (2009) study may be that their students did not collect the data themselves but rather learned to reason about data that others provided. Cognitive science research shows, however, that familiarity with the data significantly influences interpretation (Shah & Freedman, 2011). There is evidence to show that highly successful practicing scientists have difficulties interpreting graphs when they are not familiar with the source of the data, the nature of the instruments, or the laboratory contexts (Roth, 2003); or, given graphs, they might hypothesize possible data sources that in fact could not have led to the graphs (Roth & Bowen, 1999a). Thus, what the graphs in Figure 1 show are data against the background of inevitable variations – figure (difference, correlation) against ground. These representations do not show what is not included and therefore remains invisible. To use an analogy, not knowing which measurements have been excluded is like not knowing what a reporter's camera shot leaves out – e.g., the fact that the benches of the House of Parliaments not shown are actually empty.

Previous research shows that even science and mathematics graduates from university have difficulties interpreting data when these do not fall into some unequivocal continuous relationship (Roth, McGinn, & Bowen, 1998). Lack of familiarity in dealing with anomalous data may lead even advanced undergraduate students in honors programs into "creative solutions" and "fibbing results" (Roth & Bowen, 2001). In fact, students often are taught graphing using clean data which "enculturates students to an expectation that natural phenomena are inherently mathematical" (Roth, 2001, p. 12). As a result, students experience difficulties when interpreting data. Moreover, they seek other factors that might mediate the assumed perfect relationship. Although reviewers in the scientific community often take graphs without actual data as "lazy attempts at demonstration" (Myers, 1990, p. 244), these attempts apparently do not assist students in understanding the graphs. In fact, a think-aloud study with experienced research scientists showed that especially those working outside the university had trouble interpreting graphs from first- and second-year university courses of their own field (Roth & Bowen, 2003).

This state of art suggests that we need a more general form of graphical literacy (i.e., "graphicacy") not just about reading the various forms of graphical representations but also a form of literacy with respect to the relationship between claims and the possible origin of the data (Shah & Freedman, 2011). What do scientists do during the generation of data? How are data distinguished from non-data? What are the criteria for what counts as data and, therefore, for what is included? What does not count as data and therefore is not even included in the analysis on which subsequent claims are based? In this study I (a) present analyses of the real-time data collection process in one advanced science laboratory and (b) use these results for opening and encouraging a debate on the design of science curriculum.

Research such as that by Cobb and Tzou (2009) provides initial indications about how students come to understand the relationship between claims, on the one hand, and the data on which such claims are built, on the other hand. However, this study provided students with data rather than allowing them to collect data themselves as this had been done in another study with eighth-grade students (Roth, 1996) or the fourth- and fifth-graders in another study of open inquiry (Metz, 2004). Students therefore may develop only a partial sense of the relationship between some phenomenon and the manner in which it presents itself in data and subsequent reports. Based on these studies and the stated findings by Shah and Freedman (2011), we may state a tentative hypothesis in this way: *(a) Students cannot assess data actually presented if they do not know how the data have been generated; and (b) students cannot arrive at sound conclusions if they do not know how real data differentiate themselves from non-data, that is, how signals are separated from noise.*

Methods

This study was designed to investigate the ways in which successful research scientists generate the data that they subsequently use in publications. The featured episodes of scientific discovery work were recorded as part of a five-year ethnographic study of an advanced biology laboratory focusing on fish vision.

Laboratory Ethnography: Environment, Research Focus

The research team was interested in better understanding various aspects of the life history of salmonid fishes on the Pacific West Coast of Canada. The team had specialized on the visual system of these fishes, its uses, and the changes it undergoes throughout the life cycle. An important aspect of salmonid fishes is their migration from the rivers where they hatched to the salt water feeding grounds and back to their spawning grounds in the same river systems where they were born. Research in the 1930s had found that just prior to seaward migration, salmon apparently change their pigment composition from a freshwater to a seawater form. The team was interested in measuring the changes in visual pigment over time as a possible indicator for the optimal time of releasing artificially raised juveniles into the wild. This would address the historical problem of unpredictable return rates of adult salmon with tremendous impacts on the local economy (Roth, Lee, & Bowyer, 2008). The team also developed new apparatus and new software for collecting data that exceeded the number of data points in previous studies by two orders of magnitudes. Rather than measuring light absorption in the retina for different wavelengths one point at a time, the new apparatus allowed measuring the light absorption across the entire spectrum in “one shot” taking around 500–1,500 milliseconds.

To maintain the excised retinal tissue in an active state, the fish has to be kept in a dark container for a minimum of two hours. Because of the light sensitivity of retinal tissue, the experiment has to be conducted at very low intensities of red light – which requires the researchers to dark-adapt their eyes for a period of 30–60 minutes. The fish is anaesthetized and, immediately before removal of the eyes, sacrificed by severing the spinal cord. After removal, the eyes are “hemisected” and the retina removed. Under the microscope, the researcher cuts one piece of the retina, which he mounts on a slide whereas the remainder is stored on ice in a saline solution. From here on, the retinal pieces are handled only under infrared illumination.

The piece of retina on the microscope slide is macerated. Adding some saline solution, covering the preparation with a cover slip, and sealing the preparation to prevent evaporation of the solution completes the mounting process. The slide is placed under a microscope fitted with two light sources, one for the stimulus (xenon) light beam the other, an infrared lamp, for providing the background illumination to search for the objects of interest.

Conceptually, the measurement unfolds like this: To obtain information about the photoreceptors in the retina, two measurements have to be made. In the first, a light pulse is made to traverse the slide at a spot where there are no cells (the person operating the microscope asks to take a “reference”). In the second, the pulse is made to go through the cell (the person operating the microscope asks to take a “scan”). Because more light (normally) is absorbed in the cell than in the surrounding saline solution, the intensity difference in the two light pulses is attributed to absorption in the photoreceptor cell. The absorption spectrum covers a range of frequencies, but depending on the type of photoreceptor cell, light in the ultraviolet, blue, green, or red part of the spectrum is maximally absorbed. Figure 2 presents the spectrum for a blue cone generated in the laboratory based on a large number of measurements. When a cell had been exposed to light before (“bleached”), no absorption spectrum is observed. The maximum of the absorption curve is called “lambda-max” (λ_{max}). It is used to calculate the ratio of the two vitamin-A-based chemicals that absorb light and were thought to characterize the different stages in the life cycle of the salmon (i.e., while living in salt vs. freshwater environments). That is, in the course of the life cycle, the absorption spectrum for a blue cone shifts depending on the relative amount of vitamin-A_1- and vitamin-A_2-based photoreceptor cells.

««««« Insert Figure 2 about here »»»»»

Research Team

A full professor in biology, with a publication record that spanned more than 30 years, headed the lab (Craig). He had been successful throughout his career in many respects and subsequent to this study obtained an endowed chair at another university. He had received a number of awards and fellowships, had obtained continuous, often multiple-concurrent funding from national agencies, and had a substantial publication record. Theo was a full-time research associate with a background in physics. Theo was responsible for the software, data storage, and data processing. He also participated in the collection of the data. A postdoctoral fellow (Elmar) contributed to the design of the experiments and was mostly responsible for the field settings where the specimens for the experiments were sourced. His PhD had focused on salmon. A doctoral student (Shelby) did most of the measurement together with one of the other team members. As part of a larger project on the interaction between scientists and society, the head of the laboratory and I had joined efforts to study salmon and the exchange of knowledge between a fish hatchery raising salmon and this laboratory. As a trained physicist, I was a member of the team participating in designing the studies, mathematical modeling of light absorption from source to detector, collecting data, modeling data, interpreting data, and publishing the results in the natural sciences (e.g., Temple et al., 2006, 2008) and education (e.g., Roth, Hawryshyn, Haimberger, & Welzel, 2001).

Data Collection and Analysis

This study used apprenticeship as ethnographic research method (Coy, 1989; Roth, 2005a) because I learned about the relevant biology and laboratory techniques while participating in the scientific work. A research assistant or I videotaped data collection sessions in the wet laboratory and recorded the 2–3-hour team meetings. I also kept field notes, collected PowerPoint slides used during presentations, and copied sample graphs produced during data processing. The research assistant periodically interviewed team members. The videotapes were transcribed verbatim, enhanced by images of the graphs presently being talked about that were copied from the videotape. The transcriptions were annotated while being prepared whenever something appeared to be salient because members themselves were pointing it out or when something out of the ordinary happened. During subsequent passes, further annotations were added. For example, the research team was scrambling when the equipment, which had worked the night before, no longer worked in the morning. The note "what scientists do when they do not know what they are doing?" was added to the transcription.

For the present study, the tapes were analyzed in a first-time-through approach: at no point during the analysis is it allowed to take something that happened later as a resource in the interpretation (Garfinkel, 1996). That is, each instant on the tapes was viewed through the lens of the unforeseeable nature of what happened subsequently. Thus, the mentioned episode of the equipment that did not work, the method allowed focusing on what scientists really do when they do not know when something does not function rather than on their explanations that they provide for the event once the issue is resolved (Roth, 2004). This form of analysis forces the anthropologist to abandon insights that come with and from hindsight. The transcripts were improved and enhanced during the analysis to bring them to the level featured here, including pauses, overlaps, and prosodic features (pitch, pitch contour, speech rate).

This study was informed by conversation analysis, an approach that assumes the speaking turn pair as the minimum unit of analysis that makes sense (e.g., ten Have, 1999). The effect of this approach is that it reveals the way in which members to the conversation hear what is being said rather than the analyst's interpretation. In the following example, Craig says, "Do you want me to bleach it" (turn 019); because the intonation (pitch) is rising toward the end, a question mark is placed. Rather than interpreting this locution, suggesting that Craig has asked a question, the role of the statement from *within* the conversation itself is brought out by following how the subsequent speaker takes it taken up (turn 021).

```
019  C: okay; save that. (0.27) do you want me to blEACH it?
020     (0.73)
021  T: i zINK we dont need thIS one.
022  C: okay.
```

Theo states, "I think we do not need this one" with a strongly falling pitch (indicated by the period after the last word) as this tends to be the case in constatives. Here, then, we find a question–answer pair: Whether the preceding statement functions as a question depends on the second statement, which, technically speaking, makes available the perlocutionary part of the speech act (i.e., the effect). But the question actually answered may not be about bleaching but about retaining a particular data point. The next turn pair (i.e., 21–22) constitutes a proposed ("I think") constative–acceptance ("okay") pair. The team goes on not "bleaching" the receptor and therefore not capturing the bleached data point. In the following section, a statement such as "Theo comments . . ." should be read as a short form of stating that the turn pair to which Theo's locution belongs has treated his contribution to the laboratory talk as a comment.

This approach to analysis, therefore, does not require special interpretive methods. Rather, it requires the analyst to hear the participants in the manner they hear (understand) each other (Garfinkel & Sacks, 1986). What the following analyses present is "shop floor talk" from a scientific laboratory. Because I had been a member of the research team for a five-year period, I am fluent in this shop floor discourse. When shop floor competency cannot be ascertained – such as when, as happened in one of my studies, a social psychologist without physics background listens to physicists – tremendous mishearing ("misinterpretation") may and does occur.

Constructing Data: Differentiating What-is-in from What-is-out

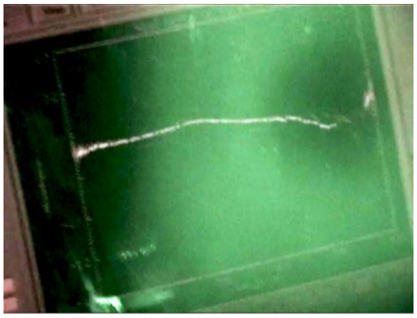
In school science, students are presented with tasks and task conditions that they have to address and for which they are held accountable. In everyday life situations, however, people also choose to abandon a problematic issue (Lave, 1988). That is, under certain conditions, those facing a problem abandon it rather than spending time and resources in the perhaps futile attempt of trying to solve it; and when the problem disappears, even scientists may not try to understand what had caused it in the first place (Roth, 2004). Watching the videotapes, we observe very similar situations in the scientific laboratory. The scientists decide, at different instances along the trajectory that takes them from living fish to the representation of retinal light absorption in a research article what to include and what to exclude from their data. This selection process begins in the laboratory, where the scientists make a first decision about whether to keep (saving it to the hard disk) or scrap a measurement.

"Rather Nice . . . Pretty 'Pretty' if you Ask Me."

Scientists are particularly pleased when a measurement exhibits features that resemble the ideal. This is shown in Episode 1, which begins with the noises made when Craig opens and closes the shutters that allow the sampling beam to fall onto the microscopic slide. (Transcription conventions can be found in the Appendix.) The participants in the lab hear these noises as specific transitions in the data collection process. Here, they precede the announcement that a reference measurement is to be taken. Theo responds by formulating that the measurement is under way (turn 003). Once the shutter-related opening and closing noises are heard again, Theo – as anyone else working in the lab – knows what has happened and formulates that the reference measurement has been taken. Any lab member present also knows that Craig now is aligning the photoreceptor with the beam. After a long pause, Craig utters the name of what he has located, “a single cone,” and he then announces the scan. There is another longish pause, after which the computer monitor displays the difference between the two intensity measurements. Theo is adjusting the scale, as the difference between the intensity distributions next to and through the photoreceptor is very small. “Looks pretty green to me, ” Theo says and adds, using the disjunctive conjunction “but,” that it is “rather nice, actually” (turn 009). That is, this turn acknowledges that the curve is a rather nice looking one, but there is a problem. This problem is apparent if one knows that the green photoreceptor cone is actually paired with a red cone. When Craig takes a measurement on one of these, he would announce a “double cone.” That is, we have an opposition here between the single cone that Craig has announced and the curve that Theo sees as resulting from the green member of a double cone.

Episode 1

001 C: ((*click*)) (0.75) ((*cluck*)) ref
002 (0.82)
003 T: under way
004 (1.77) ((*click*)) (0.83) ((*cluck*)) (0.40)
005 T: reff ´done
006 (16.49) ((*lab members know that Craig is aligning the photoreceptor with the beam*))
→ 007 C: ((*click*)) (0.78) ((*cluck*)) ((*Craig opens the shutter for allowing the light to come through*)) (1.49) sINgle cONe (0.62) scA:N
008 (7.27)
→ 009 T: * looks prETty grEEN to mE; but e:h (1.21) rUZer nICE ACtually (0.31) huh
010 (1.12)
011 C: o:kAY (.) could be a double cone sidewa[ys].
012 T: [yea]
013 (5.61)
014 T: * looks ´prETty `prETty if you ask me
015 (0.56)



016 C: o(.)kAY=
→ 017 T: =<<dim>but i zink it is in the green region.>
018 (1.86)
019 C: okay; save that. (0.27) do you want me to blEACH it?
020 (0.73)
021 T: i zINK we dont need thIS one.

After some time has passed, Craig acknowledges this possibility by saying that what he is looking at what may be a double cone but seen from the side – in which case the sampling beam would have gone through both the red and the green member. The signal from this member does not show up, however; or, rather, the team members do not articulate the graph as exhibiting this second aspect. Theo acknowledges Craig's explanation. After a long silence, he describes the curve as something that looks "pretty 'pretty'." However, after Craig acknowledges this description (turn 016), Theo again uses the disjunctive "but" to introduce his assessment of the recording as consistent with a green cone – leaving open that it is a contrast to the single cone (blue, UV) that Craig suggested to have seen. Craig acknowledges, requests the data to be saved, and then asks whether they should "bleach" the cone. Bleaching means shining light on the cone for about 2 minutes until all of the light-sensitive molecules have changed chemically. A subsequent measurement would then show no longer an absorption spectrum. It constitutes a form of experiment where, after the procedure, the phenomenon as disappeared from hand. This change from presence to absence of the absorption curve is therefore proof that the observed absorption curve "was real" rather than artefactual. In the present instance, there is a proposal | acceptance turn pair sequence (turn 020 | 21, 21 | 22) as a result of which they do not need to bleach this one. This process would have enabled them to establish the absorption more clearly as the difference between the spectra before and after bleaching. It would have allowed them to compare two measurements through the photoreceptor rather than comparing the measurement with the reference, which has been taken next to the photoreceptor and therefore is not taking account of any absorption or effect from the cell walls and within cell fluids. In contrast to the astronomers featured in another study (Garfinkel, Lynch, & Livingston, 1981), however, the present scientists cannot repeatedly vary the phenomenon by shifting the telescope, thereby literally having their phenomenon "*in hand* at all times in the inquiry" (p. 137). As the present example shows, the scientists frequently leave out the bleaching part when they are convinced that they have their phenomenon in hand. Thus, "Do you want me to bleach this one?" and "I don't think we need this one" may constitute a proper gloss of what the scientists were saying.

The scientists tend to scale the data so that the phenomenon exposes itself, which allows them to make a rapid decision whether to retain the measurement, which then becomes data, or whether to "chuck" the measurement, so it cannot become data. Scaling is a graphing-related practice that scientists often enact but that high school students are unfamiliar with (Roth & McGinn, 1997). Sometimes they keep a measurement and decide later whether it should be discarded and not taken into consideration (see below). In this laboratory, certain expressions allowed me to recognize when the data belonged to the ideal type, when they looked the way the scientists wished the data would look all of the time. The expressions included "pretty," "nice peak," "it's [looks] pretty good," "pretty 'pretty'," and "beauty." It is with reference to these ideal types that scientists excluded other measurements as irrelevant. Characteristic expressions marking the appearance include "I struck out on this one," "quite a bit of absorption," "flat liner," "bleached," "very hard to read," "too much in here that I want to look at," or "photo products." In each case where such a descriptor occurred, the measurement was discarded.

"*Alright, I'll Venture [On]*"

In my field notes and transcriptions, I repeatedly entered comments of the type "there is a potential graph, but C discards it. The novice would probably not know at all what to do with this graph." As research on concept formation showed, humans learn about the nature of a concept from the contrast of instances and non-instances (Lakoff, 1987). In Episode 2, the scientists eventually discard the measurement without providing explicit reasons for doing so. They read the measurement as possibly being consistent with some object that they know. But because it does not fit the phenomenon they are after, the measurement is not retained. The episode begins with Craig's announcement – around the time he formulates the scan – that he is looking at "either a single cone or a broken rod" (turn 011). When the first amplified images of the absorption (difference) spectrum shows up on the monitor, Theo comments, using the disjunctive conjunction "but" that he is looking at (the signal of) what looks like a "bleached rod" (turn 019). This description picks out one of the two possibilities that Craig has formulated, modifying it by the adjective "bleached," which means, a signal in the region where the rod would be expected but much weaker. He follows up his description by producing an extended chuckle. After a pause, however, Craig points toward a peak on the left part of the screen, querying, "what is *this,*" and ends the statement with the disjunctive conjunction "though" (turn 023). Theo acknowledges the statement and presence of the feature and, following a long pause, names the possible peak: "UV a [alpha]), that is, the peak they were after and would be reporting in the article to which these data contributed.

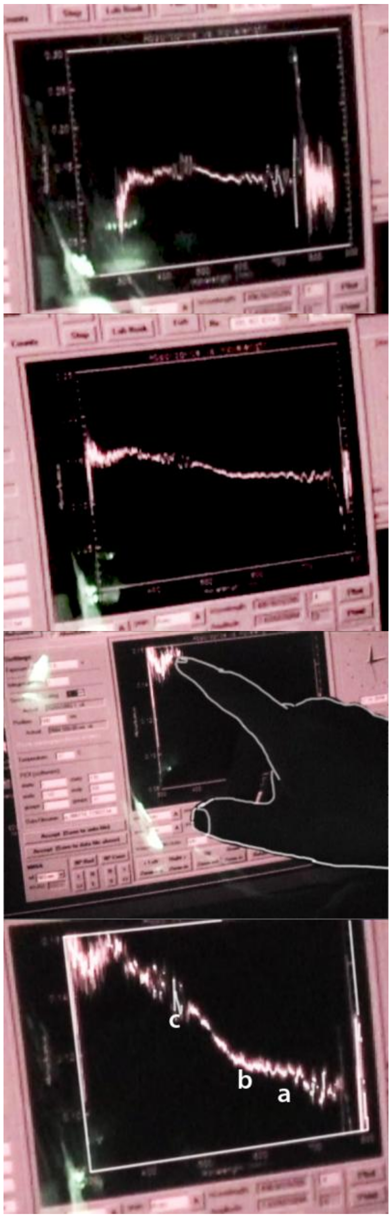
Episode 2

  007 C: scan
  008   (0.71)
  009 T: <<pp>under way>
  010   (2.92)
→ 011 C: now its EIther a sINGle cONE or a brOKen
         rOD.
  012   (0.40)
  013 T: * <<pp>alright> ((*modifies graph,*
         magnifies difference))

  014   (10.29)
  015 M: is that something that looks similarly?
  016   (1.38)
  017 C: yea. (0.18) it cAN look ˘similar.
  018   (3.85)
→ 019 T: * but now looks like a bleached rod.
  020   (0.56)
  021   <<dim>hu hu hu hu hu> .hhfs

  022   (1.26)
→ 023 C: * <<f>well> whats thIS though. ((*points*
         to middle, "fuzzy peak"))
  024   (0.45)
  025 T: yea.
  026   (5.49)
→ 027 T: yea ze ze [u: vee a]

  028 M: [would bASe]line be dOWn here *
         ((*points to "a"*))
  029   (0.58)
  030 T: yes. baseline would (.) would be * down
         here. ((*moves back and forth around "b"*))
  031   (1.03)
  032 T: id be one pOSsibility. (0.53) dis *
         ((*points to "c"*)) could be a (.) blEACHed
         <<f>rOD.> (0.19) dere is a little remem
         remnant with the photoproduct [right]



033 C: [yea]
034 (0.40)
035 T: dats one way of reading it.
036 (2.33)
037 T: ze other one is to read dis ze whole branch from <<dim>here down is
something which> ((*from upper left to "a"*)) (1.04) <<f>caused by (.)
in ze reference;
038 (0.65)
039 M: uh hm
040 T: and dat we have somezing else really going on here ((*left "peak"*))
041 (1.57)
042 T: <<p>but dat (??) in the positions here
043 C: <<p>alright i=ll, (0.99) venture (2.49) ((*clack*))

I then ask a question about the location of the baseline (turn 015) – upon which the curve is "grafted" and which would be subtracted by an algorithm that Theo has written – to which Theo responds by providing a more extensive (than normal) reading of the possible things that might have caused the features of the graph. He locates the right end of the baseline in the graph marked by the letter "b" and then points to the area marked by the letter "c" suggesting it could have been caused by a broken rod. Finally, he suggests that the entire "branch" could be the result of something in the "reference" (measurement) so that it is not caused by the photoreceptor at all (turn 037). Craig then announces that he is "venturing" on, which concludes the episode and starts the search for a new cone. They have not saved the data. Theo, in not challenging Craig's "decision," and by not taking the initiative to save the data on his own, de facto accepts the decision to discard these data. Subsequent to the transcribed part of the episode and in response to my question, Theo comments: "We don't think we can use it."

In this situation, the scientists "venture on." A spectrum where there is only a faint hint of a Gaussian-shaped absorption curve is removed from the data set. During the discussion of the first several months of my presence – amounting to over 3,000 data points – the lead scientist repeatedly suggested removing some of the measurements. However, we do not know what the relation is between these 3,000 data points that the scientists discuss and all those instances that they have not included while in the laboratory. In the subsequently published studies from this work, there is no hint about the relation. But in other situations, they did actually retain measurements. This may be driven by the needs to have sufficient data points for a particular phenomenon. Thus, in some data retained, the peak hardly showed up at all and was not very different from the one they discarded in this episode: the signal is of the same order as the noise (Figure 3). As the scientists wanted to extract the location of the maximum of the peak from the data that they have retained, they need to clean these up – a phenomenon enabled by the nature of inscriptions themselves (e.g., Latour, 1987). But the graphs themselves tended to exhibit noise (e.g., "we have some noise on top here that is the problem"). For example, they looked at the curves and saw them as approximately Gaussian-shaped that "sit" on an incline. Because the incline is considered an artifact, they "subtracted" it from the actual measurement. The resulting curve is "cleaned up" or fitted in one of a number of ways. Thus, the scientists ultimately noted in one of their publications:

««««« Insert Figure 3 about here »»»»»

> Each record was linear detrended if necessary (Harosi, 1987). A nine-point adjacent averaging function was used for line smoothing, and the smoothed curve was normalized to zero at baseline on the long wavelength arm and to one at the centre of the α-band. The

> fit of the normalized curve was compared with a nonlinear least-squares routine to the upper 20% of the weighted A_1/A_2 averaged Govardovskii et al. template (Govardovskii et al., 2000) (based on the centre of the α-peak ±40 nm). (Temple, Veldhoen, Phelan, Veldhoen, & Hawryshyn, 2008, p. 3880)

It is from the fit that they extracted the wavelength at which the absorption curve has its maximum. In essence, the scientists got rid of the variation in the measurements to extract what they de facto take to be the real data. Because the team wanted to get rid of unwanted detail in their data, they used a Fourier transformation procedure ("FFT" and "inverse FFT"). The basic idea underlying this procedure is that any mathematical function can be represented as a sum of sine curves. This sum, which may consist of an infinite number of terms, is called a Fourier series. Once represented as a Fourier series, the scientists "lop off" the higher-order frequency terms, which corresponds to getting rid of the high frequency "noise" in the curve. The scientists then retransform the series into a curve, which then looks similar to the original but excluding the "noise." That is, they include the measurement but exclude the variation in it.

Can Some of the Data Legitimately Be Excluded?

In the preceding section, we see how the scientists exclude data along the way even before they get to the analysis of the location of the peaks of the spectra. What the nature of the measurements included is therefore depends on the nature of the measurements not included. The nature of the data does not derive from the measurements themselves: even if unacknowledged, there is always an irreducible figure|ground relation. Distinguishing this figure (data) from ground (noise, background) – where ground is necessary for the figure to appear – is part of the data collection process that allows scientists subsequently to make sense. Here, they do a first selection in the laboratory. If there is too little evidence that the data "meet inclusion criteria" (Craig), then these are not even saved. Later, as discussed in the session analyzed below, further exclusion criteria are made operative. Thus, Craig suggests excluding all the data that are below 503 nm, that is, less than what previous research has reported to be expected for the vitamin A_1-based chromophore (absorbing chemical) and everything above 527 nm, which is more than the expected vitamin A_2-based chromophore. In this way, the data included would be selected based on the results of previous research. However, if the true range of the wavelengths were to be different, then the scientists would have eliminated data that could have been used for revising the accepted range of the maxima for the absorption curves.

In the present situation, Theo suggests retaining these data until after they have a better sense about the quality of their data. This would be consistent with an orientation of retaining judgment until a better understanding has been arrived; it also keeps open the possibility for revising the scientific canon with respect to the range of possible λ_{max} values. At the moment, the decision which data to retain is based on the curves that other researchers have published, whereas Theo proposes getting the quality of the data so high that they themselves can decide which λ_{max} ("lambda max," wavelength where absorption is maximum) to take, and, therefore, to establish their own scale for the A_1/A_2 ratios. This part of the meeting begins when, following a comment about the variability of the data, Shelby presents the results to the other team members (Figure 4). It is a series of 5 histograms. As Shelby explains, these are "batches" of data collected in two-week intervals from fish that the laboratory received in this case from the Kispiox First Nation fish hatchery.

««««« Insert Figure 4 about here »»»»»

The videotapes show that relevant to the interpretive work of the scientists is that they know where the data are coming from, that is, where from which river system or hatchery they derive. Without knowing where the data come from and how these were collected, scientists struggle (Roth, in press). This is so because, as seen in another study, even the identification of the species of a specimen depends on contextual information: a group of scientists could not distinguish between the young of three or four species unless they knew where in the river these young have been caught (Roth, 2005b). Thus, it is not surprising that Carl would be asking, by proposing a possible answer with rising intonation, whether the data presented derive from the Kispiox First Nation hatchery.

Episode 3, Fragment 1a

```
  001  S: so ive just kind of quickly summarIZed this um we have um hIStograms
          for what ive call bATches ((Figure 4)) um so you have the first
          batch which came in well and two weeks lAter and two weeks lAter, so
          EAch bAtch is two weeks apart.
  002  C: this is kispiOX?
  003  S: kispiox. um ALl fISh from the kIspiOX so our fIRSt batch april
          thIRteenth this was our distribUTion for ALl fish um individual rods
          s histogram um all plotted on the same scALe so by the sECond batch
          and what we sEEM to be gETting here is
→ 004  C: kay so bt whY: thAT spike of A:ONe there?
  005  S: this is where i say that perhaps we=re getting these individual
          [rODs]
  006  T: [kkm ] the problem is that we we ARe wORking with them on the
          munznbEAtty dAta
  007  S: oh no; this is; this is not both.
  008  T: this goes from fIVehundredzrEE to fIVehundredtwENtysEVen
          nANo[mEters     ]
  009  S:     [<<p>alright>]
  010  T: and so if we have sOMething lets say at say fOUrninetyEIght or
          fourninetySEVen that would be in the tALL bAR in there
  011  S: yea
  012  T: so,
  013  C: uh um
  014  T: so what we really have to decIDe is what ze rANge is [of the ]
→ 015  C:                                                      [well if] its
          lOWer than fIVehundredandthrEE then it doesnt belONG in there.
```

Shelby confirms the origin of the slide as coming from the Kispiox, that he received the fish on April 13, and that the graph presents the distribution for all individual rods from which measurements were taken. Not articulated because available to all gazing at the graphs is the abscissa bearing the label “percA2” (i.e., % A_2) and the ordinate being labeled “counts.” The members to the setting also understand that the second graph entitled “2.00” is the “second batch” because Shelby had moved the cursor to this graph while naming it (turn 003). Craig asks about the “spike of a=one,” which, as can be seen from the plot, is a high count for the first bin of the histogram corresponding to 0% A_2 and, therefore, equivalently to 100% A_1 (turn 004). Shelby and Theo, who had collected and processed the data, respectively, take turns to explicate.

Shelby begins by making an attribution to the effect from individual rods, but Theo points out that they are working with the “Munz and Beatty” data, that is, with the algorithm for determining the A_1/A_2 ratios from a given polynomial regression equation with the specific λ_{max} that they determined from the data. This determination itself requires one of several possible mathematical procedures for approximating the absorption curve – a published seventh-order polynomial, curve smoothing, or a process of removing high-frequency parts of the curve through Fourier and reverse Fourier transformation. Theo points out that the Munz and Beatty regression curve is based on the 503 nm and 527 nm as the λ_{max} for vitamin A_1 and A_2, respectively. But in their data, he points out (turn 010), there are curves with λ_{max} = 498 nm or λ_{max} = 497 nm. These data “would be in the tall bar in there” (turn 010). He also suggests that they have to make a decision about the range of, but does not succeed in completing his statement as Craig interrupts him with the categorical statement that “if its lower than 503 [nm], then it doesn’t belong in there” (turn 015). As the Munz and Beatty (1965) study had been conducted, among others, on the five Pacific salmon (*Oncorhynchus*) species, Craig takes this as a strong reason for excluding data that appear to suggest maximum absorption lower and higher than the range set in this 35-year-old study.

Although Shelby and Theo appear to accede, they also articulate further reasons for retaining the measurements under discussion. Theo, who is not a biologist by training, wonders whether the range Munz and Beatty offered is “exclusive” for the coho, thereby implying that there might actually be a different range for the coho. He suggests not knowing what the range of the λ_{max} would be for the coho (turn 020) and that they do not yet have sufficient or “sufficiently good” data to “decide [them]selves” (turn 022) and, therefore, where to expect these to lie (turn 024).

Episode 3, Fragment 1b

```
016  T: <<p>yea right>
017  C: its
018  T: you see what I dONt know is this is fIVehundredthrEE t
        fivehundredtwentysEVen is the um exclUSive rANge which which is
        pOSSible for ze for ze coho
019  S: yea
020  T: i dont know what the cOHo curves where from where to where they do
        go,
021  S: yea; thats right a [good point]
022  T:                    [and our   ] data isnt gOOD enOUgh yet to decIDe
        oursELVes which
023  S: yea
024  T: where we expECt it to [be]
025  S:                       [i:] think I think that thats a rEElly good
        pOInt that if we were to have it; if we were to get rID of thAT say
        no NO more zeros ye actually can go below fiveothrEE then weed start
        to sEE that the cURves are a bETter shAPed; ((gestures an inverse
        parabola in the air))
026  T: <<p>yea>
027  S: becOS that thats a rEALly good pOINt cos we dont know for sure that
        the protein in coho is the sAMe as what <<dim>munznbeatty hve done
        it for other fish like rainbow trout and salmon and so it> mAY BE
        slightly different
028  T: so I wouldnt trust the first ten to five percent of the and from
        ninety <<p>five plus or some[thing like  ]>
029  S:                             [yea no thATs] a really good pOINt
030  T: <<p>zats really out of range rEAlly>
031  S: so that thats
032  M: ee if you take a bIN size you have fIVe now?
033  S: uh bin sIZe of yea [fIVe]
034  M:                    [fIVe] if you took tEN would the the cURves come
        out clEARly and shift alONG ah as we go through tIMe?
```

```
  035  S: u::m well they they dO kind of now anyways; so well, the first ones,
          first ones our very first day so i=m you know not; this one here
          maybe not but; these ones hERe seem to be slOWly shIFting towards
          the rIGht u lEFt rather; towards more a. and thIS one hERe is
          looking pretty nICe almost
→ 036  C: yea thATs thats nice. i mean i i=m hA:Ppy with what i:ve see there;
          thats nICe data.
```

Shelby supports Theo. He suggests that if they were to plot the measurements below the 503 nanometer minimum wavelength then the curves (histograms) would exhibit better shapes (turn 025). He accompanies this suggestion by an iconic gesture that outlines a slightly skewed Gaussian curve, as visible in the histograms displayed. He further elaborates suggesting that the coho may in fact have some differences in their protein that is part of the absorbing molecule so that the λ_{max} of the associated visual pigment might change with respect to the data that Munz and Beatty (1965) provided and approximated with their regression equation for determining A_1/A_2 ratios. Theo says that he does not trust the first and last one or two bins of the histogram (turn 028), and Shelby affirms, “that’s a really good point” (turn 029).

My own question in turn 032 pertains to changes to the curves that might be observed if the bin size were enlarged to 10% and whether the expected shifts in the A_1/A_2 ratios that are expected over time would be better visible. In response, Shelby asserts that the distributions were already visibly shifting to the left, which means, a shift toward less A_2 (porphyropsin) and more A_1 (rhodopsin) as would be expected from fish at that time of the year just prior to migration.

It is perhaps his status of an outsider to the community of biologists that makes Theo less susceptible to the strong disciplinary constraints of the reigning paradigm in the field of biology. In Craig’s case, whose 30-year career to that point has been entirely within the field of salmonid fish vision and its paradigm, deviating from the paradigmatic canon may be more difficult. The effect of this canon may have been particularly strong, because based on it its founder, George Wald, had been a co-recipient of the 1967 Nobel Prize in Physiology or Medizine in 1967. This same pattern can be observed in the videotapes when the team interprets the data of changing A_1/A_2 ratios in the course of the life history of the fish, as these unfolded during the research project.

In the end, we find out what good data are to look like, and measurements that otherwise are excluded, first in the laboratory when the scientists are assessing the absorption spectrum and discard those that do not fit what they want. Here again, Craig suggests that certain data points “do not belong here” and what should be included because it constitutes “nice data” (turn 036). There is an inner contradiction not made salient by the members to the meeting is the fact that the team already has excluded many other measurements that might make the results look even less nice, and that Craig further suggests to remove all those data that lead to the high peak for the bin in which the amount of A_2 in the photoreceptor is between 0 and 5%. That is, the data look nice because they have been made to look nice, not because they are inherently nice.

Theo and Craig seem to be acquiescing to Craig, who categorically excludes the data points below a cut-off point suggested by other, much older research. Craig then listens to the discussion. But it is not just an arbitrary decision to drop data. Because the team ultimately has to defend its decision – when attempting to publish the study – it needs to be able to articulate a set of reasonable inclusion/exclusion criteria.

The issue about inclusion and exclusion of measurements was not settled but came up repeatedly during the 2-hour laboratory meeting. Thus, for example, some 30 minutes after Fragment 1 (Episode 3), the issue becomes again the topic of talk. The fragment begins just after

Elmar has asked about the last data point for the Kispiox hatchery, something of special interest to him because he has done a lot of research on fish in this geographical area. Shelby notes that the last batch had arrived on June 12 and that he had already shown it prior to Elmar’s (late) arrival (turn 037). Shelby notes that there are a number of measurements with $\lambda_{max} < 503$ nm and, as before, suggests that these data “should not necessarily [be] discounted” (turn 039). He begins by pointing out that especially with the data they do have, one would get some particular result, but Craig insists again on the point that anything that does not meet the criterion of being “within the window of λ_{max}” “ha[s] to be rejected” (turn 42).

Episode 3, Fragment 2

037 S: <<pp>which is, this (.) this is one> batch five was the twelfth of
June; yes so that was our last batch, so the one I showed you was
indeed the last one I guess
038 E: and it was?
039 S: so what; what I would deal with eventually what was really
interesting was that um this is a matter of fact to the fact that
there is some readings that are below fiveothree; and that we
shouldnt necessarily discount those. and that they may not actually
bE fiveothree; but youre gonna obviously get some ((*gestures
triangular shape*)), specially with that kind of anal[ysis]
040 E: [yea] yea
041 S: <<dim> you get some of [the]>
042 C: [so] if if they dont if they dont meet the
cRItErion of bEing within the wINdow of a lambda max then they have
to be rejected
043 T: yea, no doubt, a few weeks ago i am not sure that fiveothree to
fivetwentyseven is the correct range. i would say its fournINEtyfive
to fivezirtyfIVe is the correct range
044 C: but but did you go to mUNznbEAttys coho?
045 S: um not yet do they, do they have coho? munznbeatty
046 C: well i mean did you gO look at the literature and see whether
the[re is]
047 S: [we hav]ent been out of the lab; so we havent gone thAT deep into
it yet.
048 C: um
049 S: but we obviously need to goo
050 C: maybe look at alexANders papers too
051 S: yea they they just go to the standard fiveotwo. but that could be
the technique they are using too and they use a raw average mean so
theyre not using the actual is small point for the individual rods
00:52:15
→ 052 C: yea which maybe,
053 S: <<p>which isnt bad>
054 C: no I I knOW what you are saying
055 T: yea
056 C: theodore there maybe that maybe um variance withIN the OPsin; the
perfORmance within the OPsin. um molecule.
057 S: and it could also be the experiment
058 C: and its could. also. it could it may have something to do with
chromophore binding within the uh um the counterion pocket.((*Throws
up hand, as if saying “who knows?”*)))) um. ((*As if thinking*)) there
could be a variety of things explaining those short wave length
lambda maxes but I think for pURposes of analyses that has to be

remOVed.
→ 059 T: I would prefer to drop five percent and above the five percent to
fiveninetyfive and the rest. <<pp>id prefer that do this>
060 S: yea; because is the same with the other end too. because there
is [that one point yeah]
061 T: [where it is higher]
062 S: its just tapering ((*gesture far toward his right*)) off at either
end. yea. if we got rid of the if we get rid of thOSe two ends it
starts looking very normally and distributed. its slightly skewed.

Theo responds, arguing for retaining the data. He suggests not being sure that the range is the correct one and proposes to use a different range for retaining measurements: 495 nm < λ_{max} < 535 nm, which in fact extends the heretofore accepted range by 8 points above and below (turn 043). Craig insists: had they checked the data for coho salmon in the Munz and Beatty (1965) article. (Coho [*Oncorhynchus kisutch*] is one of five major salmon species of the *Oncorhynchus* genus in the Salmonidae family.) Shelby asks whether this study has in fact coho data, to which Craig responds asking whether they have gone to the literature to find out. That is, Craig does not insist on asserting that Munz and Beatty *actually* have coho data but asks whether they (Shelby, Theo) have looked into the literature more generally. Shelby says that they have been in the lab and therefore "not gone so deep into it yet" (turn 047). Craig also suggests that they go to the paper by Alexander, Sweeting, and McKeown (1994), which is the one on which this entire research project is based.

Shelby notes that "they just go to the standard, five-o-two that could be the technique they are using too and they use a raw average mean" (turn 051). Not only is the number 502 different from the number 503 that the previous speakers had articulated, but also the Alexander et al. paper does not at all mention such a number. The paper refers to the same Munz and Beatty study that already has been discussed in this meeting. Craig then accedes in the sense that he gives a reason why Theo might be correct that there are maximum wavelength peaks below 503 nm (turns 054, 056). He refers to the possibility that there could be variations arising from difference within the chromophore-binding counterion pocket (turn 058). Changes in the amino acid sequences near this pocket may result in changes of λ_{max} of the visual pigment. That is, he articulates a detailed understanding of the chemistry associated with the rhodopsin (A_1)-associated vision processes that has been a central research issue of recent decade. But he then insists on removing those measurements from the present analysis, giving a particular emphasis on the center part of the verb. Theo, in turn, insists on his preference for retaining curves with λ_{max} values being 5 nm above or below the currently accepted range (turn 059), and Shelby – in using the confirmative "yea" followed by the conjunctive "because" that is followed by a reason – apparently supports this position (turn 060). What we do not see here is a discussion of the fact that the Munz and Beatty study provides an algorithm for establishing the A_1/A_2 ratio given that pure A_1 has a λ_{max} = 503 nm and A_2 pure has a λ_{max} = 527 nm. What would it mean for A_1/A_2 ratio if λ_{max} < 503 nm or λ_{max} > 527 nm? This question cannot be answered unless the group is to establish a different range of values with an adjusted regression equation to estimate the appropriate A_1/A_2 ratio.

Shelby elaborates that there is a similar issue "at the other end" (turn 060). A closer inspection of the histograms (Figure 4) shows that in each of the five "batches" the very last bin is indeed higher than those to the left. He suggests that removing the two ends would make the curves "start looking very normally and distributed" though they remain "slightly skewed" (turn 062). Shelby does not say that the measurements should be removed. It remains open whether the team should use an extension and the resulting change in the shape of the histogram, because the data are "tapering off on either end." Extending the acceptable range would produce normal curves that are slightly skewed – toward longer wavelengths, as the histograms show.

The publication resulting from this study will show that the team is going to retain the lower wavelength limit of λ_{max} = 503 nm but accept longer wavelength maxima on the other end (Temple et al., 2006). The team describes using a different than the heretofore-used Munz-and-Beatty algorithm for estimating the A_1/A_2 ratios (i.e., Govardovskii et al., 2000). The article states using the Munz and Beatty algorithm as a second estimate and using the average of the two for deriving the relative amount of A_2 present (in %). The more recent paper had not done measurements on salmonids but published a general algorithm based on the observation that across a broad range of animal species, the shape of the absorption curves is independent of the λ_{max}/λ ratio. Based on the data Munz and Beatty (1965) had published for coho salmon, the team derived, using a least square regression, a third-order polynomial for the determining the A_1/A_2 ratio. Biologically, the explanation given in Shelby's dissertation and the associated scientific journal article is in terms of the broadening of the spectra towards *longer wavelengths*, consistent with observing $\lambda_{max} > 527$ nm, whereas there are no processes that would explain the observation of $\lambda_{max} < 503 \pm 1$ nm.

In summary, as exemplified in this meeting, the researchers address the issue of the abnormally high counts in the first and last bin of the histogram. These are the result of the fact that Theo and Shelby have counted all data with $\lambda_{max} < 503$ nm as indicating the presence of 100% A_1 and counted all data with $\lambda_{max} > 527$ nm as indicating the presence of A_2 even though the previously established curve maxima for the two chromophores are 503 and 527 nm, respectively. According to Craig, this means that the corresponding absorption curves, even though they might look "nice" and fall into the category of "beauties," they do not meet inclusion criteria and therefore should be excluded. They should be excluded even though there is a possibility that the maxima shift because of chemical processes or because of some other reason. For the purposes of their present analysis, they should be excluded. The two individuals less enculturated and invested in the canon (Theo, Shelby) oppose this recommendation and express the preference of retaining the data.

Discussion

This study was designed to better understand how scientists construct their data by including or excluding some but not other measurements. As studies of scientists at work shows, scientific claims are the end result of transformations that begin with pieces of natural matter (Latour, 1993; Roth & Bowen, 1999b). If a real understanding of the graphs requires familiarity with the original phenomenon and the transformations through which it is turned into a scientific fact, then scientific literacy with respect to graphing (i.e., graphicacy) means something like being able to make the symbolic ascent from the claim to the original setting in which measurements have been produced. This study was designed as an investigation into the scientific practice of data generation for the purpose of reflecting on the design of science (and mathematics) education. Although there are attempts to explicate data generation drawing on observations among undergraduate students (Brewer & Chinn, 2001), the theory is overly rationalist and does not explicate the actual *course of data generation* observed in the social studies of science where radical uncertainty leads to a dialectical tension between the natural world and its representation (e.g., Latour, 1993; Roth, 2009).

The first episode shows how scientists retain measurements even when there is a contradiction between the visual assessment (single cone) and the graph (green member of double cone). The scientists do retain measurements when they believe that they can make use of it in an ensuing publication. In the featured episode, they decide to conduct only part of the measurement because they already have sufficient information with respect to this object. They have their phenomenon in hand even if they do not bleach and thereby destroy the photoreceptor to see whether the signal disappears – as the astronomers in another study did (Garfinkel et al., 1981). In the second episode, we observe the scientists discarding a measurement even though there is evidence that they elsewhere retain data of a very similar quality. In discarding this run, the measurement does not even enter the consideration of shaping the data used to support the research claims about the phenomenon at hand. In this choice, the scientists shape the data that they ultimately work with in a way that differs from merely dealing with error variance. Therefore, the data ultimately made visible – including both true and error variance – is set against the "non-data" that are in fact invisible. The phenomenon, therefore, rises as figure against the ground in what scientists present, which itself is set against an invisible ground of all the possible responses that the scientists obtain when they probe nature.

In the third episode, the scientists are confronted with the results of their earlier selection. The episode shows how in the face of existing experimental results, the chief scientist requests chucking out all those of the remaining data that do not fit the paradigm – here locations of the peaks below 503 nm and above 527 nm. Although the team members who collected and processed the measurements suggest retaining these until they know more, the lead scientist argues in favor of excluding them based on the scientific canon at the time. There is a tension, however, because this very project, in its totality, ultimately overthrew the Nobel Prize-winning canon on the variations in the composition of the photoreceptor molecules (between rhodopsin and porphyropsin). It would eventually turn out that they slackened the requirement for the upper boundary – without providing information as to how this decision affected their assessments of the A_1/A_2 ratios – but did not change the lower boundary of λ_{max}.

It is evident from the analyses that scientists make their selection based on both an intimate knowledge of the laboratory equipment and the entire process by means of which retinal tissue comes to be transformed into %-A_2 distributions. The scientists also exhibit orientation to their field in attempting to adhere to the canon even when the data themselves appear to contradict it. Of course, the scientists adhere to the canon by using legitimate equipment or by extracting A_2 ratios from the λ_{max} determined by templates rather than, for example, by a best-fit polynomial grounded in the measurement points themselves. At various stages in the process, measurements are dropped and thereby become invisible in and to the construction of the phenomenon – which is always based on the measurements retained rather than those that are excluded from consideration. Within the retained measurements, the phenomenon comes to stand as figure against the ground (unexplained variation). This study shows that even to the actual transformations of the measurements, scientists do what they can so that "order is not simply constituted" but it is "*exposed, seized upon, clarified, extended, coded, compared, measured*" so that it can in fact be "*subjected to mathematical operations*" (Lynch, 1990, p. 163). Measurements that can be anticipated to resist the processes of order generation are simply excluded as unsuitable because, for one or another reason, "they do not meet criteria for inclusion." Moreover, this study shows that mathematical operations – e.g., curve fitting, FFT, inverse FFT – are used to make the measurements suitable for subsequent modeling.

In contrast to much of school science, where students do what they are told to do, the scientists are in control over what to do and which measurements to retain for the analyses that they ultimately report. (Even many science teachers, in part as a result of deprofessionalization, have to cope “with a top-down, assessment-drive curriculum” [Levinson, 2011, p. 113].) Scientists’ inclusion and exclusion criteria are grounded in their familiarity with all those instances that do not even qualify for entry into the data sources. They literally constitute the frame that allows only some measurements to enter into consideration. This frame therefore reduces the original messiness, which then permits the phenomenon to appear more clearly against the ground then it would if everything were included.

Based on the present results, I strongly suggest allowing students to make decisions about which measurements to include or exclude from subsequent interpretations and claims. I also make this suggestion because an important dimension in learning appears to be the level of control that students have over framing the questions that are to be answered through the inquiry (e.g., Chin & Chia, 2004). In this form, learning is student-centered and satisfies the students’ needs to seek and find answers to their own questions. Student question-based learning environments “[afford] many possibilities for transforming classrooms into active learning environments where there is a dynamic interplay of questioning, explanation, argumentation, design of investigations, communication of ideas and findings, collaboration, and reflection” (p. 725). In advocating data generation as an integral aspect of students’ science experience, I do not however abandon the idea that teachers are inessential, for the mere introduction of some scientific tools does not necessarily lead to inquiry (Waight & Abd-El-Khalick, 2011). Rather, if Vygotsky (1989) is right in stating that *all* higher psychological functions are societal relations first, then arrangements in which science students interact with other individuals that represent current lay or professional scientific practices is essential to the development of interpretive practices.

Without an integral knowledge of where the data come from, how they are generated, possible problems in the production of data, and how data differ from non-data, even scientists would be hard pressed to make conclusions and support claims. Making such decisions is important, for example, in democratic decision-making processes. This became evident to me when the mayor, town council, and town engineers in my hometown based a decision on constructing a water main to supply people with running water on the report of a particular scientist who only collected data on a single day and in only one-sixth of the homes concerned. (The most extensive presentation of all issues involved in this case can be found in an article for municipal engineers on the construction of community health and safety, Roth, 2008.) In the ensuing public debate, some savvy citizens, however, did point out both aspects of the data collection as problematic issues. However, the mayor, town council, and town engineers not only disregarded the critique of the data collection and quality but also failed to take into account, and even omitted from entry into the data sources, more than 30 years of information that locals had collected about the water. That is, these municipal officials could perceive a phenomenon emerging from their data rather than a different phenomenon that would have emerged if *all* the information had been considered that was available at the time. The citizens displayed exactly the kind of scientific literacy that science educators might want to foster: Rather than simply accepting scientists’ claims, we want to develop a scientifically savvy citizenry that raises questions about the data collection, demands public articulation not only about how claims had been produced on the data presented but also about data not retained, or engages in queries about how the framing of the nature of data collection is related to the sociopolitical agendas in play. It was just such forms of scientific literacy that AIDS activists displayed and that led to changes in the scientific protocols for collecting data on the efficacy of new drugs (e.g., Epstein, 1995 &

1997). It would lead us toward a citizenry engaged in "more and more diverse trajectories of [scientific] fact construction and closure in controversies" as much as learning about the myth of "science as clean and elegant" (Roth & Désautels, 2004, p. 154).

This study shows that scientists do not just interpret decontextualized data. They require familiarity with the natural setting and with the measurement process and the criteria that include or exclude some of these. The resulting graphical representations are integral part of the entire research process, and familiarity with it is a requirement for interpreting them. Thus, the graphs have a part|whole function to the entire research – or, more technically speaking, they are synecdoches of the research process, that is, parts of the research process that point to the entirety of the process. There is convergent evidence from at least two studies at the middle school level (Cobb & Tzou, 2009; Roth, 1996) and one at the elementary level (Metz, 2004), where students, confronted with some data, persisted in asking questions about the context within which the data were collected that they had been asked to analyze. In the absence of background information requested by students, the teacher "eventually abandoned the data analysis that was planned for this class session because the students continued to ask questions about the situation from which the data were generated" (Cobb & Tzou, 2009, p. 162). The students did not want to engage in the data analysis until after being familiar with the context of the data collection.

Cobb and Tzou's episode points us to an important aspect of problem-solving practices in the everyday world where "[p]ersons-acting are free to transform, solve or resolve a problem, or abandon it in favor of other options. In the parlance of the [Adult Math Project], they 'own' their own problems" (Lave, 1988, p. 156); and Metz (2004), even though she studies children, takes a similar perspective. In an equivalent manner, the scientists shown here own their problems and their data. They make a decision whether they want to include or exclude measurements that they have made for supporting the claims that they intend to reporting in a research article. Without understanding the relation between graphing (a social practice) and the setting (of research) we have little understanding of how "[c]ognition is constituted in dialectical relations among people acting, the contexts of their activity, and the activity itself" (Lave, 1988, p. 148). Dialectical relation here means that there is a unity to the activity as a whole, which is the minimal unit to understand the sense of any of its parts, including data and their graphical representation.

Learning about data generation is important because the very nature of a scientific phenomenon depends on it. For example, if the scientists in this study had included all data, then their very phenomenon might have been lost in the variance caused by the data actually excluded from analysis. If students are to become more savvy about the nature of science and to take å more critical stance towards the results of scientific research, they need to learn both to interpret the data that are included and to make judgments about the quality of evidence that includes considerations of data not retained for interpretation. Thus, it is only under specific condition that Galileo's inclined plane experiment yields the data that support his claims about the quadratic increase of distance traveled with time (or linear increase of velocity with time) (Garfinkel, 2002). Students do not generally have experiences in learning to differentiate the conditions under which a scientific phenomenon appears und under which conditions it will be lost. In traditional laboratory exercises, students are held to produce data such that these support the scientific theory. Even in extended experimental investigations, which are premised on the ideas that students learn to conduct independent research, teachers may disallow an experiment so that students get data that confirm some existing theory (S. M. Ritchie, unpublished data). Knowledgeably, reasonably, and accountably making distinctions between conditions that produce versus those that lose a phenomenon should be an integral aspect of scientific literacy. There appears to be no better place to learn making such distinctions than open inquiry school science.

The results of this study should encourage science educators to begin a debate concerning the experimental and ethical dimensions in data generation. Students need to learn to deal with a range of questions: Which measurements may be legitimately excluded from entering the data sources? On what basis are the decisions that distinguish between sources included and sources excluded? How does exclusion influence our understanding of nature? What ethical implications are related to the question of excluding measurements from the data sources? From a nature of science perspective, students *ought to know* what scientists do and how their actions affect *what* we know about nature. Knowing what might affect the selection of data sources is as important for understanding nature as understanding the nature of the phenomenon (figure) against the overall variation within the data (ground). Future science education research, therefore, ought to investigate and (experimentally) tease out the role that familiarity with the entire inquiry process in general and the data generation process in particular plays in students' learning of science and understanding of the nature of science.

Acknowledgments

A joint grant from the Social Sciences and Humanities (SSHRC) and Natural Sciences and Engineering Research Councils of Canada (NSERC) funded both the natural science and social science parts on which this research is based. I am grateful to my team members Craig Hawryshyn, Theodore (von) Haimberger, Elmar Plate, and Shelby Temple for their contributions to our collaborative project.

References

Aberg-Bengtsson, L., & Ottosson, T. (2006). What lies behind graphicacy? Relating students' results on a test of graphically represented graphical information to formal academic achievement. *Journal of Research in Science Teaching, 43*, 43–62.

Akerson, V., & Donnelly, L A. (2010). Teaching nature of science to K–2 students: What understandings can they attain? *International Journal of Science Education, 32*, 97–124.

Alexander, G., Sweeting, R., & McKeown, B. (1994). The shift in visual pigment dominance in the retinae of juvenile coho salmon (Oncorhynchus kisutch): An indicator of smolt status. *Journal of Experimental Biology, 195*, 185–197.

Apedoe, X, & Ford, M. (2010). The empirical attitude, material practice and design activities. *Science & Education, 19*, 165–186.

Carlone, H. B., Haun-Frank, J., & Webb, A. (2011). Assessing equity beyond knowledge- and skills-based outcomes: A comparative ethnography of two fourth-grade reform-based science classrooms. *Journal of Research in Science Teaching, 48*, 459–485.

Chin, C., & Chia, L.-G. (2004). Problem-based learning: Using students' questions to drive knowledge construction. *Science Education, 88*, 707–727.

Chinn, C. A., & Brewer, W. F. (1993). The role of anomalous data in knowledge acquisition: A theoretical framework and implications for science instruction. *Review of Educational Research, 63*, 1-49.

Cobb, P., & Tzou, C. (2009). Supporting students' learning about data generation. In W.-M. Roth (Ed.), *Mathematical representation at the interface of body and culture* (pp. 135–170). Charlotte, NC: Information Age Publishing.

Coy, M. (1989). Being what we pretend to be: The usefulness of apprenticeship as a field method. In M.W. Coy (Ed.), *Apprenticeship: From theory to method and back again* (pp. 115–135). Albany, NY: State University of New York Press.

Epstein, S. (1995). The construction of lay expertise: AIDS activism and the forging of credibility in the reform of clinical trials. *Science, Technology, & Human Values*, *20*, 408–437.

Epstein, S. (1997). Activism, drug regulation, and the politics of therapeutic evaluation in the AIDS era: A case study of ddC and the 'Surrogate Markers' debate. *Social Studies of Science*, *27*, 691–726.

Falk, H., Brill, G., & Yarden, A. (2008). Teaching a biotechnology curriculum based on adapted primary literature. *International Journal of Science Education, 30*, 1841–1866.

Feldman, A., & Pirog, K. (2011). Authentic science research in elementary school after-school science clubs. *Journal of Science Education and Technology, 20*, 494–507.

Fogleman, J., McNeill, K. L., & Krajcik, J. (2011). Examining the effect of teachers' adaptations of a middle school science inquiry-oriented curriculum unit on student learning. *Journal of Research in Science Teaching, 48*, 149–169.

Garfinkel, H. (1996). Ethnomethodology's program. *Social Psychology Quarterly, 59*, 5–21.

Garfinkel, H. (2002). *Ethnomethodology's program: Working out Durkheim's aphorism*. Lanham, MD: Rowman & Littlefield.

Garfinkel, H., Lynch, M., & Livingston, E. (1981). The work of a discovering science construed with materials from the optically discovered pulsar. *Philosophy of the Social Sciences*, *11*, 131–158.

Garfinkel, H., & Sacks, H. (1986). On formal structures of practical action. In H. Garfinkel (Ed.), *Ethnomethodological studies of work* (pp. 160–193). London, UK: Routledge & Kegan Paul.

Garli, B., & Rule, A. C. (2009). Integrating social justice with mathematics and science: An analysis of student teacher lessons. *Teaching and Teacher Education, 25*, 490–499.

Govardovskii, V. I., Fyhrquist, N., Reuter, T., Kuzmin D. G., & Donner, K. (2000). In search of the visual pigment template. *Visual Neuroscience*, *17*, 509–528.

Jordan, R. C., Ruibai-Villasenor, M., Hmelo-Silver, C. E., & Etkina, E. (2011). Laboratory materials: Affordances or constraints? *Journal of Research in Science Teaching*, *48*, 1010–1025.

Kirschner, P. A., Sweller, J., & Clark, R. E. (2006). Why minimal guidance during instruction does not work: An analysis of the failure of constructivist, discovery, problem-based, experiential, and inquiry-based teaching. *Educational Psychologist*, *41*, 75–86.

Lakoff, G. (1987). *Women, fire, and dangerous things: What categories reveal about the mind*. Chicago, IL: University of Chicago Press.

Latour, B. (1987). *Science in action: How to follow scientists and engineers through society*. Milton Keynes, UK: Open University Press.

Latour, B. (1993). *La clef de Berlin et autres leçons d'un amateur de sciences* [The key to Berlin and other lessons of a science lover]. Paris, France: Éditions la Découverte.

Lave, J. (1988). *Cognition in practice: Mind, mathematics and culture in everyday life*. Cambridge, UK: Cambridge University Press.

Levinson, R. (2011). Science education from people for people: taking a standpoint. *Studies in Science Education*, *47*, 109–117.

Lynch, M. (1990). The externalized retina: Selection and mathematization in the visual documentation of objects in the life sciences. In M. Lynch & S. Woolgar (Eds.), *Representation in scientific practice* (pp. 153–186). Cambridge, MA: MIT Press.

McElhaney, K. W., & Linn, M. C. (2011). Investigations of a complex, realistic task: Intentional, unsystematic, and exhaustive experiments. *Journal of Research in Science Teaching*, *48*, 745–770.

Medical Billing and Coding. (2010). Sitting is killing you. Accessed February 1, 2012 at http://www.medicalbillingandcoding.org/sitting-kills/

Metz, K. E. (2004). Children's understanding of scientific inquiry: Their conceptualization of uncertainty in investigations of their own design. *Cognition and Instruction*, *22*, 219–290.

Munz, F. W., & Beatty, D. D. (1965). A critical analysis of the visual pigments of salmon and trout. *Vision Research*, *5*, 1–17.

Myers, G. (1990). Every picture tells a story: Illustrations in E. O. Wilson's Sociobiology. In M. Lynch & S. Woolgar (Eds.), *Representation in scientific practice* (pp. 231–265). Cambridge, MA: MIT Press.

Nargund-Joshi, V., Park Rogers, M. A., & Akerson, V. L. (2011). Exploring Indian secondary teachers' orientations and practice for teaching science in an era of reform. *Journal of Research in Science Teaching, 48*, 624–647.

Ritchie, S. M., Sandhu, M, Sandhu, S., Tobin, K., Henderson, S., & Roth, W.-M. (in press). Emotional arousal of beginning physics teachers during extended experimental investigations. *Journal of Research in Science Teaching.*

Roth, W.-M. (1996). Where is the context in contextual word problems?: Mathematical practices and products in Grade 8 students' answers to story problems. *Cognition and Instruction, 14*, 487–527.

Roth, W.-M. (2001). "Authentic science": Enculturation into the conceptual blind spots of a discipline. *British Educational Research Journal, 27*, 5–27.

Roth, W.-M. (2003). *Towards an anthropology of graphing: Activity theoretic and semiotic perspectives*. Dordrecht, The Netherlands: Kluwer Academic Publishers.

Roth, W.-M. (2004). "Tappen Im Dunkeln". Der Umgang mit Unsicherheiten und Unwägbarkeiten während des Forschungsprozesses [Groping in the dark. Dealing with uncertainties during scientific research]. *Zeitschrift für Qualitative Bildungs-, Beratungs-, und Sozialforschung, 5* (2), 155–178.

Roth, W.-M. (2005a). *Doing qualitative research: Praxis of methods*. Rotterdam: SensePublishers.

Roth, W.-M. (2005b). Making classifications (at) work: Ordering practices in science. *Social Studies of Science, 35*, 581-621.

Roth, W.-M. (2008). Constructing community health and safety. *Municipal Engineer, 161*, 83–92.

Roth, W.-M. (2009). Radical uncertainty in scientific discovery work. *Science, Technology & Human Values, 34*, 313–336.

Roth, W.-M. (in press). Undoing decontextualization or how scientists come to understand their own data/graphs. *Science Education.*

Roth, W.-M., & Barton, A. C. (2004). *Rethinking scientific literacy*. New York, NY: Routledge.

Roth, W.-M., & Bowen, G. M. (1999a). Complexities of graphical representations during ecology lectures: An analysis rooted in semiotics and hermeneutic phenomenology. *Learning and Instruction, 9*, 235–255.

Roth, W.-M., & Bowen, G. M. (1999b). Digitizing lizards or the topology of vision in ecological fieldwork. *Social Studies of Science, 29*, 719–764.

Roth, W.-M., & Bowen, G. M. (2001). "Creative solutions" and "fibbing results": Enculturation in field ecology. *Social Studies of Science, 31*, 533–556.

Roth, W.-M., & Bowen, G. M. (2003). When are graphs ten thousand words worth? An expert/expert study. *Cognition and Instruction, 21*, 429–473.

Roth, W.-M., Hawryshyn, C., Haimberger, T., & Welzel, M. (2001, August). *Visual perception: More than meets the eye*. Paper presented at the bi-annual meeting of the European Association for Research on Learning and nstruction, Fribourg, Switzerland.

Roth, W.-M., Lee, Y. J., & Boyer, L. (2008). *The eternal return: Reproduction and change in complex activity systems. The case of salmon enhancement*. Berlin: Lehmanns Media.

Roth, W.-M., & McGinn, M. K. (1997). Graphing: A cognitive ability or cultural practice? *Science Education, 81*, 91–106.

Roth, W.-M., McGinn, M. K., & Bowen, G. M. (1998). How prepared are preservice teachers to teach scientific inquiry? Levels of performance in scientific representation practices. *Journal of Science Teacher Education, 9,* 25–48.

Russell, C. B., & Weaver, G. C. (2011). A comparative study of traditional, inquiry-based, and research-based laboratory curricula: Impacts on understanding of the nature of science. *Chemistry Education Research and Practice, 12*, 57–67.

Sadeh, I., & Zion, M. (2009). The development of dynamic inquiry performances within an open inquiry setting: A comparison to guided inquiry setting. *Journal of Research in Science Teaching, 46*, 1137–1160.

Scalise, K., Timms, M., Moorjani, A., Clark, L., Holtermann, K., & Irvin, P. S. (2011). Student learning in science simulations: Design features that promote learning gains. *Journal of Research in Science*

Teaching, 48, 1050–1078.

Shah, P., & Freedman, E. G. (2011). Bar and line graph comprehension: An interaction of top-down and bottom-up processes. *Topics in Cognitive Science, 3*, 560–578.

Temple, S. E., Ramsden, S. D., Haimberger, T. J., Veldhoen, K. M., Veldhoen, N. J., Carter, N. L., Roth, W.-M., & Hawryshyn, C. W. (2008). Effects of exogenous thyroid hormones on visual pigment composition in coho salmon (Oncorhynchus kisutch). *Journal of Experimental Biology, 211*, 2134–2143.

Temple, S. E., Plate, E. M., Ramsden, S., Haimberger, T. J., Roth, W.-M., & Hawryshyn, C. W. (2006). Seasonal cycle in vitamin A1/A2-based visual pigment composition during the life history of coho salmon (Oncorhynchus kisutch). *Journal of Comparative Physiology A: Sensory, Neural, and Behavioral Physiology, 192*, 301–313.

Temple, S. E., Veldhoen, K. M., Phelan, J. T., Veldhoen, N. J., & Hawryshyn, C. W. (2008). Ontogenetic changes in photoreceptor opsin gene expression in coho salmon (Oncorhynchus kisutch, Walbaum). *Journal of Experimental Biology, 211*, 3879–3888.

ten Have, P. (1999). *Doing conversation analysis: A practical guide*. London, UK: Sage.

van der Valk, T., & de Jong, O. (2009). Scaffolding science teachers in open-inquiry teaching. *International Journal of Science Education, 31*, 829–850.

Vygotsky, L. S. (1989). Concrete human psychology. *Soviet Psychology, 27*(2), 53–77.

Waight, N., & Abd-El-Khalick, F. (2011). From scientific practice to high school science classrooms: Transfer of scientific technologies and realizations of authentic inquiry. *Journal of Research in Science Teaching, 48,* 37–70.

Captions

Figure 1. a. Hypothetical distribution of the achievement of boys and girls in science. b. Hypothetical relationship between school achievement and IQ.

Figure 2. This average absorption graph from the laboratory's database represents the results of many measurements with blue cones.

Figure 3. Measurement retained and fitted with a polynomial. The signal is barely noticeable against the variations (noise).

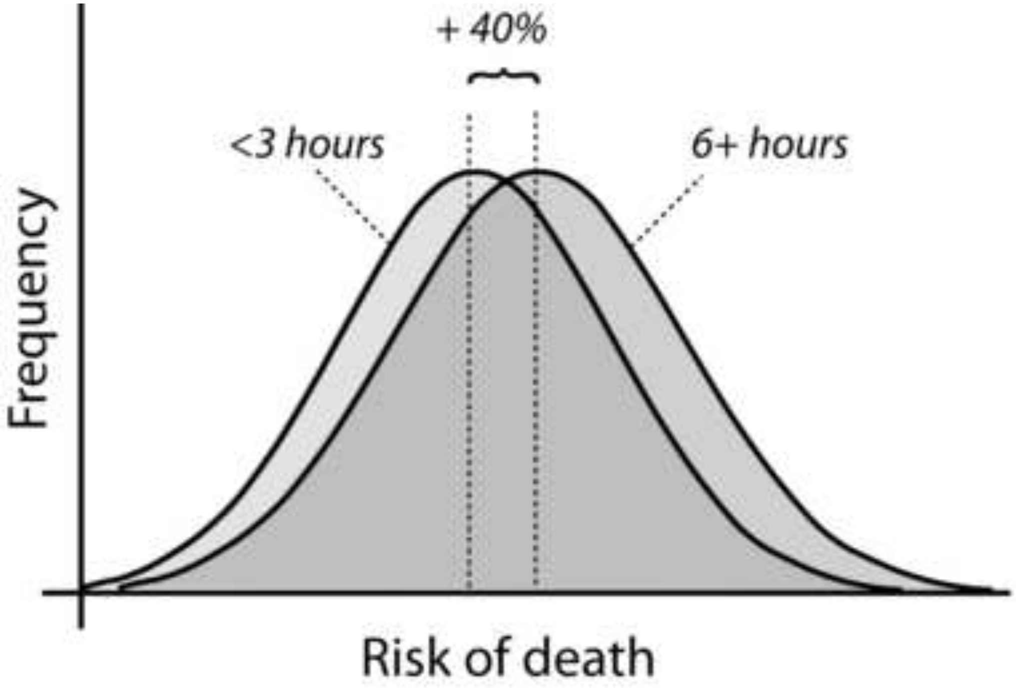
Figure 4. Data from a scientific project on the distribution of porphyropsin and rhodopsin in fish retina – here the number of cells with a certain amount of the vitamin-A_2-based chromophore (%A2).

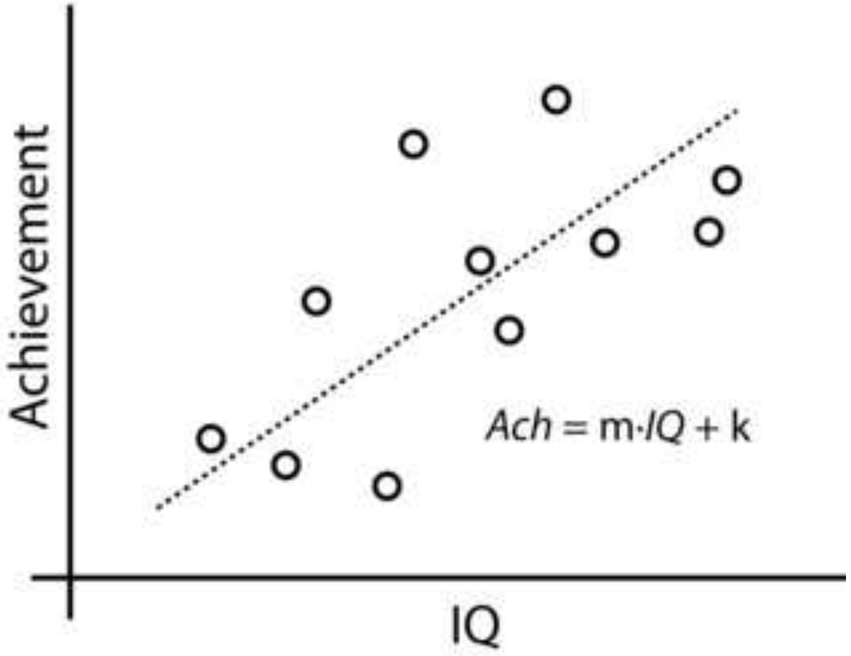
Appendix

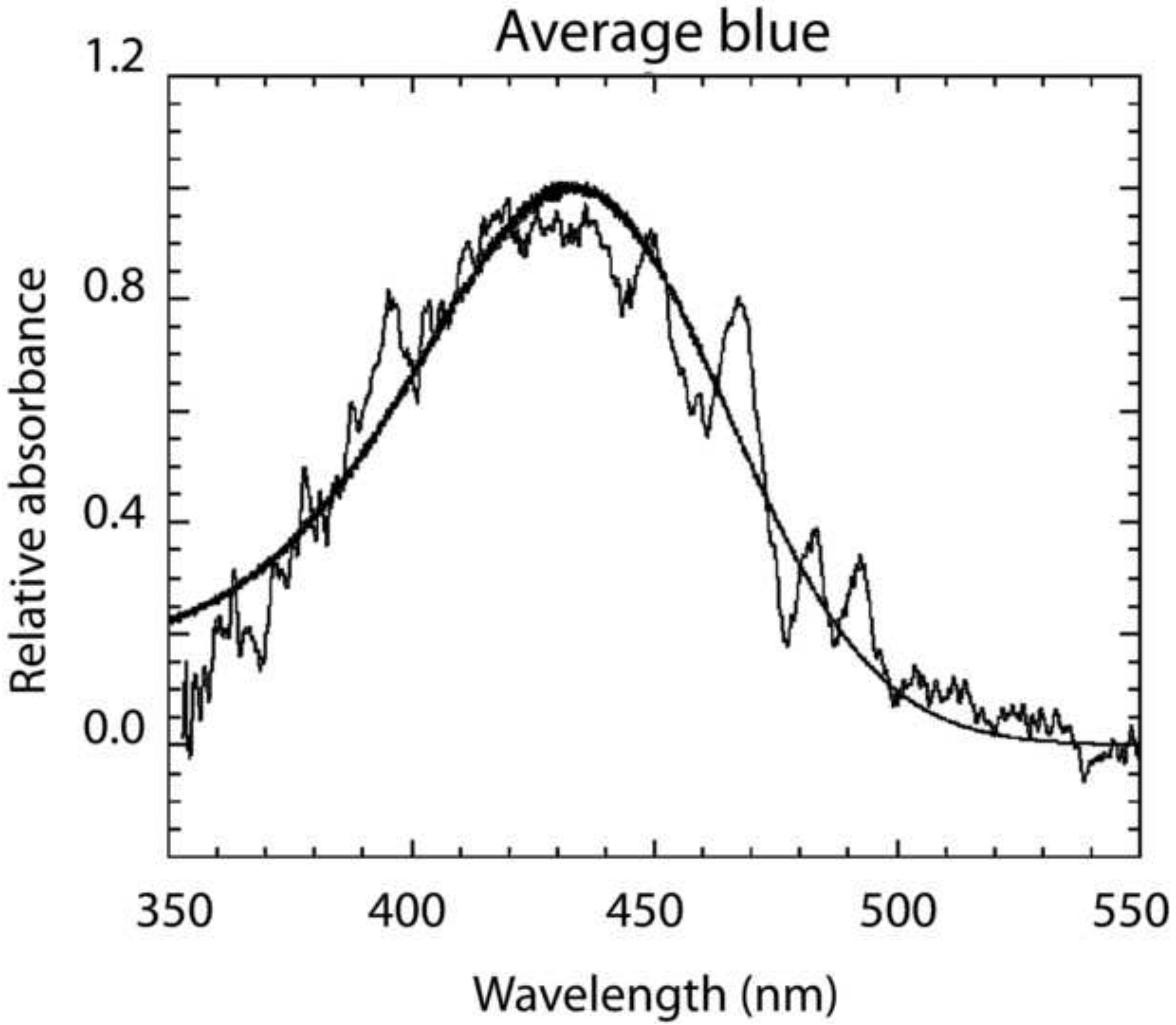
For the transcriptions, I follow a commonly used system based on conversation analysis adapted for the inclusion of prosodic features (Selting et al., 1998). In the rules implemented here, everything is written in small letters and sound words that run into each other are transcribed that way unless the run-in sign "=" is used when it would be difficult to distinguish pronunciation (e.g., "a=one"). The transcription is phonetic such that if a participant pronounces the words "this" or "that" in the way a French or German speaker often does, that is, with a soft "d" or "s," the transcription will read something like "ze other one is to read dis ze whole branch."

Notation	Description	Example
(0.14)	Time without talk, in seconds	more ideas. (1.03) just
(.)	Pause in speech less than 0.10 seconds	011 C: o:kAY (.) could be a double cone
((*turns*))	Verbs and descriptions in double parentheses and	((*modifies graph*))

	italics are transcriber's comments	
*	Asterisks marks the instant in speech that corresponds to the video image on the right	
(??)	Marks inaudible words, about one word per question mark	042 T: <<p>but dat (??) in the positions here
::	Colons indicate lengthening of phoneme, about 1/10 of a second per colon	si::ze
[ ]	Square brackets in consecutive lines indicate overlap	011 C: o:kAY (.) could be a double cone sidewa[ys]. 012 T: [yea]
<<f> >	Forte, words are uttered with louder than normal speech volume	<<f>um>
<<p> >	Piano, lower than normal speech volume	042 T: <<p>but dat (??)
<<pp> >	Pianissimo, much lower speech volume	009 T: <<pp>under way>
<<dim> >	Diminuendo, becoming weaker	<<dim>i donno>
prETty	Capital letters indicate louder than normal talk indicated in small letters.	looks prETty grEEN to mE
hh	Noticeable out-breath	
.h	Noticeable in-breath	021 T: <<dim>hu hu hu hu hu> .hhfs
–,?;.	Punctuation is used to mark movement of pitch (intonation) toward end of utterance, flat, slightly and strongly upward, and slightly and strongly downward, respectively	C: okay; save that. (0.27) do you want me to blEACH it?.
=	Equal sign indicates that the phonemes of different words are not clearly separated	i=ll
`, ´, ˘	Diacritic indicates movement of pitch within the word that follows – down, up, down up	˘similar.

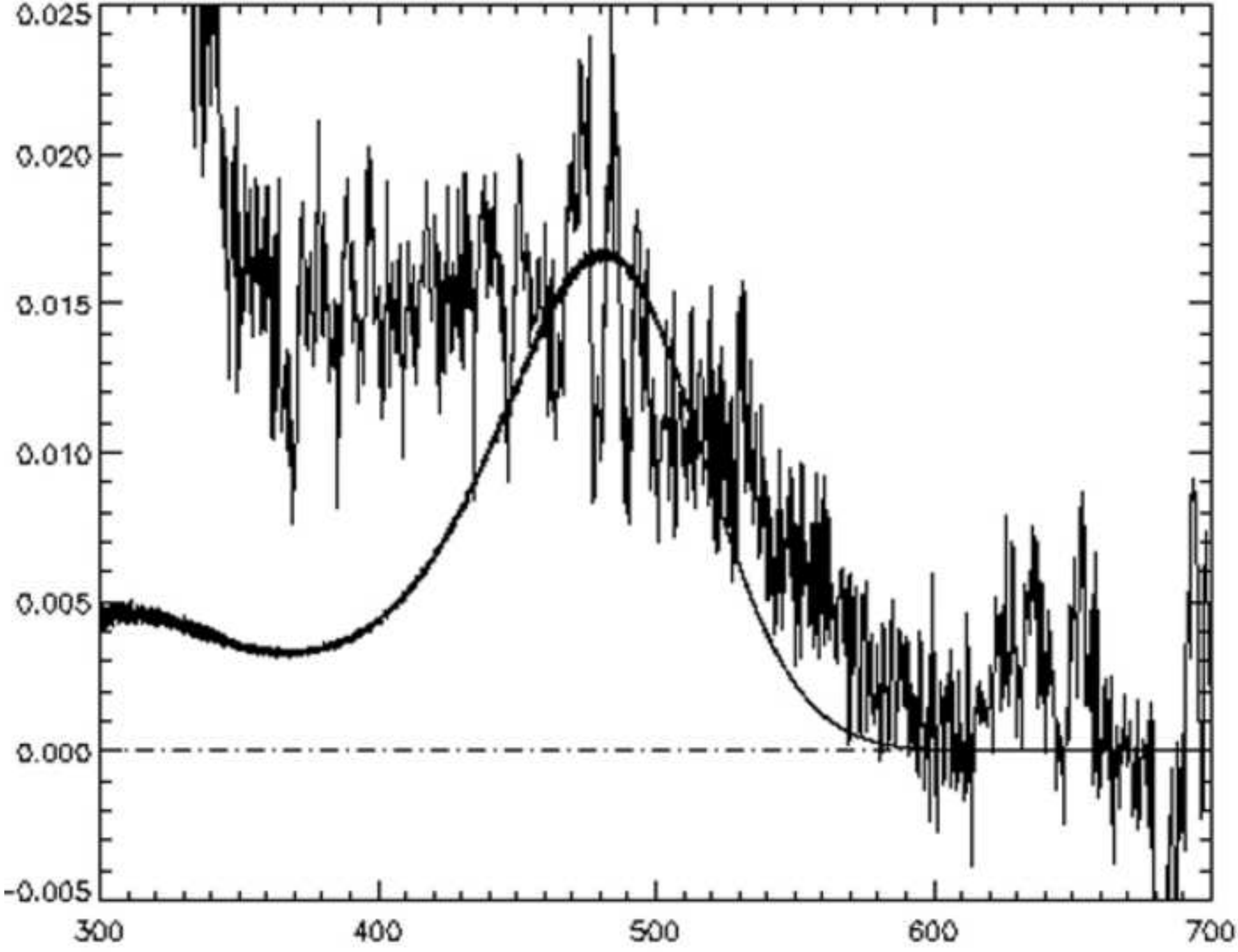

+ 40%
<3 hours
6+ hours
Frequency
Risk of death


Achievement
Ach = m·IQ + k
IQ


Average blue
1.2
0.8
0.4
0.0
Relative absorbance
350
400
450
500
550
Wavelength (nm)

line figure 3

Click here to download high resolution image



colour figure 4