Novice and Expert Characteristics in Teacher Professional Development with Astronomy Databases

Andria C. Schwortz

Andrea C. Burrows

University of Wyoming

Andria C. Schwortz, Department of Physics & Astronomy, University of Wyoming

Andrea C. Burrows, Department of Secondary Education, University of Wyoming

Correspondence regarding this article should be addressed to Andria Schwortz, Physics & Astronomy, 1000 E. University, Dept 3905, Laramie, WY 82071. Contact: aschwort@uwyo.edu

Novice and Expert Characteristics in Teacher Professional Development with Astronomy Databases

**Abstract**

This study characterizes the novice and expert behaviors of in-service K-12 teachers attending two astronomy-themed professional development workshops at a large research university in the Rocky Mountains. Fourteen unique individuals attended the workshops and participated in an activity involving a large astronomy dataset: six workshop A only, five workshop B only, and three attended both workshops A and B. Thirteen of these individuals participated in pre-/post-tests, and eleven participants contributed to more in-depth data collection including field notes during the activity and one-on-one interviews. Of the fourteen individuals, six were female and eight male. Pre-/post-test gains were calculated, and all materials were coded for themes. The authors found a great diversity in behaviors among the participants, with some possessing traits more similar to experts in astronomy, and some more similar to novices. All teachers demonstrated ability to pick out simpler trends and patterns in the data while some were able to observe more complex trends as well. Surprisingly, many teachers exhibited difficulty recognizing the need to put individual tasks into the context of the big picture, a recognition which is key to expert traits such as chunking information, organizing their knowledge based upon concepts, and contextualizing their knowledge. This implies the need for further content-based professional development opportunities for in-service teachers.

**Introduction**

One of the major open questions in science education research is that of how people transition from novices to experts. As educators it is our goal to help students (in classes), teachers (in professional development opportunities), and the general public (in informal education settings) to move a couple steps further along the scale from novice to expert. While a lot of research has been done on this in the fields of physics and math, less has been done in astronomy, and minimal amounts of research has been done regarding teachers working with datasets. The primary research question for this paper is “How do in-service K-12 STEM teachers fit into the framework of novices and experts?”

**Theoretical Framework**

This project is conducted using the theoretical framework of social constructivism (Vygotsky, 1978). Constructivism purports that learners come into the classroom with preconceptions, and they need to actively build new ideas (rather than passively absorbing them) and incorporate what they learn into their existing conceptions. The authors support that learning does not happen in isolation; people instead learn best with their peers and together construct meaning, and thus the authors have examined a group learning experience with big datasets in astronomy to assess the constructivist expert/novice aspects.

**Literature Review**

The Handbook of Research on Science Education (Abell and Lederman, eds. 2013), first published in 2007 and endorsed by the National Association for Research in Science Teaching (NARST), has served as a comprehensive literature review of the many fields of science education research, and as guidance for future work, since its first edition. One theme that occurs again and again throughout the Handbook, is that of the transition from novice to expert.

The differences between novices and experts in many subfields of science have already been studied extensively. Bransford, Brown, and Cocking (2000) described six key characteristics distinguishing experts and novices as follows.

1. Experts notice features and meaningful patterns of information that are not noticed by novices.
2. Experts have acquired a great deal of content knowledge that is organized in ways that reflect a deep understanding of their subject matter.
3. Experts’ knowledge cannot be reduced to sets of isolated facts or propositions but, instead, reflects contexts of applicability: that is, the knowledge is “conditionalized” on a set of circumstances.
4. Experts are able to flexibly retrieve important aspects of their knowledge with little attentional effort.
5. Though experts know their disciplines thoroughly, this does not guarantee that they are able to teach others.
6. Experts have varying levels of flexibility in their approach to new situations. (Bransford, et al., 2000)

These six traits can be shortened to (1) patterns, (2) organized knowledge, (3) contextualized knowledge, (4) knowledge retrieval, (5) pedagogical content knowledge and peer instruction, (6) adaptability and metacognition.

One idea common to the first three characteristics of experts is that they all require a working knowledge of the “big picture”. What makes experts more able to notice meaningful patterns is that they can chunk related ideas together so that they are more likely to understand when things are related, while novices will need to hold many unrelated ideas in their heads simultaneously. Experts not only know many isolated facts, but have a context and an organization to these facts to relate them to each other, allowing them to understand the overarching connections between these facts, while novices will not yet have this larger structure in place.

In the realm of physics, researchers such as Larkin, et al. (1980), have studied the differences between how experts and novices approach individual physics problems. For example, it is known that experts are able to distinguish when the shape of an object (i.e. rectangular vs. round) is relevant to the problem (such as when moving down a ramp with friction) as opposed to when it is irrelevant (such as when hanging from a rope) and will choose an approach making use of the relevant factors. On the other hand a novice is more likely to group all problems with round objects together, regardless of the best approach. At its root astronomy is a branch of applied physics, so one can assume that the differences between how experts and novices approach individual astronomy problems would be similar, however there is a lack of studies directly investigating novice/expert distinctions in solving individual astronomy problems. Astronomy is a key topic for teacher education, as approximately 40% of college students taking introductory astronomy intend to become teachers (Lawrenz, et al., 2005).

Research astronomy as it exists today is increasingly making use of large sets of data; for example, Schwortz, et al. (2015b) are working with a dataset with over 100,000 rows. Due to the large number of objects, these datasets must be analyzed as a whole rather than by examining one object at a time, requiring the use of computer programs written specifically for the question the researcher wishes to investigate. With the nation-wide push for STEM integration (for example, in the Next Generation Science Standards, or in NASA’s educational goals), big datasets become even more important. The skills that students learn from analyzing astronomical datasets can be applied to many other situations, such as other STEM fields, physical inventory, financial predictions, or human resources management. Because this need for handling big datasets is ubiquitous, it is important that education researchers study the transition from novices to experts in this field. Yet there exist few astronomy courses that directly address the use of large datasets, even at the post-secondary level.

There have also been a number of studies on differences in novice and expert teachers, particularly in the realm of pedagogical content knowledge (PCK) – that is, how teachers communicate science or other domain content to their students. Expert teachers possess a mental framework that allows them to interpret student behavior in a larger context (Westerman, 1991), and thus they are better able to respond appropriately than are novice teachers. However, these studies seem to be predominantly focusing on teachers’ PCK and not on the other aspects of their novice/expert transition, such as their metacognitive skills or adaptability to new situations.

**Methods**

**Participants**

Two astronomy-themed workshops for in-service K-12 teachers were held in the summer of 2014 at a large research university in the Rocky Mountains. All workshop attendees were invited to participate in this study. Fourteen unique individuals attended the workshops and participated in an activity involving a large set of astronomy data: six workshop A only, five workshop B only, and three attended both workshops A and B. Thirteen of these individuals participated in pre-/post-tests, and eleven participants contributed to more in-depth data collection including field notes during the activity and one-on-one interviews. Of the fourteen individuals, six were female and eight male. For the quantitative analysis, there are 16 pre-/post-tests (ten men and six women).

**Methodology**

This is a mixed methods study. The primary quantitative measurement used is scores on pre-/post-tests. Qualitative methods, such as interviews, artifacts from the activity, and field notes and recordings made during the activity (e.g., how the teachers explained concepts to each other), were coded for themes to explore the process by which subjects construct meaning from large datasets, and to characterize participants’ actions in a novice/expert context.

**Methods**

The authors developed an activity wherein subjects analyze data approximating that of accompanying science research on active galaxies (Schwortz, et al., 2015b), but of a scope appropriate for non-scientists to complete in an hour. Participants were presented with a set of data in 200 rows and 5 columns and were stepped through analyzing these data, first with more open-ended prompts to see what the participants were able to determine unaided, and then with increasing levels of instruction to approximate the steps taken by expert astronomers to analyze these data.

For the quantitative aspect of the study the authors developed a pre-/post-test with eight multiple-choice questions. These questions were designed to span Bloom’s taxonomy, with Questions 1-3 being at the level of knowledge or comprehension, Questions 4 and 5 requiring application or analysis, and Questions 6-7 needing synthesis or evaluation. Means and standard deviations were calculated for the pre-test and post-test, and for men and women separately. The authors examined the gains of the subjects on the quantitative questions, and using ANOVA determined whether any subsets of the subjects are distinct from the majority.

Qualitative data included open-ended questions from the pre-/post-tests (three questions from the same parts of Bloom’s taxonomy), artifacts from the activity itself, field notes taken during administration of the activity (e.g., what patterns the participants saw in the data sets), and interviews. The qualitative data was coded for themes, specifically searching for evidence of novice/expert characteristics.

**Findings**

**Quantitative**

The posttest means were higher than that of the pretests, with men specifically showing a statistically significant increase, as shown in Figure 1. The mean on the pretest

Figure 1: Pre-/Post-test Scores.

was 73.4% with a standard deviation of 18.8%. The posttest mean was 85.9% with a standard deviation of 12.8%. For 16 degrees of freedom, this results in a P-value of less than 0.050. Men had a pretest mean of 72.5% with a standard deviation of 18.5% and a posttest mean of 87.5% with a standard deviation of 11.8%. With 10 degrees of freedom, this is a P-value of less than 0.050. For women, the pre-/post-test means of 75% and 83.3% (standard deviations of 20.9% and 10.5%) were not significantly different. Men and women did not have significantly different pretest scores, nor significantly different posttest scores. Normalized matched gains (as per Hake, 1998) are shown in Figure 2. These were 0.32 overall, 0.38 for men, and 0.21 for women; these were not statistically significant from each other.

Figure 2: Normalized Matched Gains, calculated as per Hake (1998).

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  |  | Q2 |  | Q3 |  | Q8 |  |
|  | N | Pre | Post | Pre | Post | Pre | Post |
| Women | 6 | 2 | 5 | 5 | 6 | 3 | 5 |
| Men | 10 | 4 | 6 | 6 | 10 | 7 | 10 |
| All | 16 | 6 | 11 | 11 | 16 | 10 | 15 |

There were three questions where the participants’ pre- and post-test answers showed significant improvement, as shown in Table 1. Two of these were ranked

Table 1: Number of correct answers on questions 2, 3, and 8, by gender.

low on Bloom’s taxonomy, as knowledge or comprehension, while the third was at the higher level of synthesis or evaluation. Question 2 asked, “This type of plot or chart lets you examine one column of data and find its distribution.” Choices were “column plot,” “histogram,” “polygraph,” and “scatter,” with the correct answer being “histogram.” On the pretest 6 out of 16 got the question right, while on the posttest 11 people got it right, corresponding to a P-value of less than 0.050. Neither men nor women had a statistically significant change on this question.

Question 3 was “Quasars are…” with choices of “star clusters,” “pulsating stars,” “cores of active galaxies,” and “nearby extremely bright stars,” with the third choice being correct. On the pretest 11 people got this right while all 16 got it right on the posttest (P-value less than 0.050); five women originally got this question right and all six did on the posttest (not significant); six men originally got this question right on the pretest and all 10 did on the posttest (P-value less than 0.050).

The final question with statistically significant improvement was Question 8: “If you wanted to study quasars with similar types of jets, which of the following would you want to have similar values?” The choices were “Declination,” “radio magnitude” (correct answer), “redshift,” and “Right Ascension.” Ten people had this right on the pretest and 15 on the posttest, a P-value of less than 0.050. Seven men had this right on the pretest and all 10 on the posttest, a P-value of less than 0.050. The change in number of women having this correct (three to five) was not significant.

Comparing the women’s answers to the men’s answers on the pretest, and also on the posttest, showed no significant differences between their responses by gender in either case.

**Qualitative**

In coding for themes related to novice and expert characteristics, the most frequently coded theme was that of recognizing patterns. For example, the data in the activity was such that in one of the columns 90% of the entries were 0 (indicating that they had no detectable radio emission) while the remaining 10% of entries had values in the teens (indicating values for the radio light measured). All participants recognized this pattern, though they assigned it different meanings. For example, during the second workshop, the new participants debated whether this meant that no attempt had been made to observe these sources, or if it meant that the attempt had been made but the source had no radio emission – this latter interpretation was the correct one, as had been previously explained to the participants in the introduction to the activity.

The theme of understanding the big picture also emerged as important. This theme was used when participants’ responses indicated that they were looking at the purposes or goals of their actions, or that they were taking actions designed to answer larger questions about the underlying science.

From the first workshop, out of the nine participants, one showed evidence of looking at the big picture in his pre-test, saying, “Start with a question to answer,” a specific phrasing which he also used in the post-test. On his pretest, he added more detail: “Evaluate to see patterns or if your question is answered or not.” A second individual from this workshop said on the posttest, “I would look for trends, and then graph the data based on what I want to analyze.”

The second workshop consisted of eight individuals, of whom three had participated in the first workshop and study. On the pretest a total of three individuals (two repeaters, one new person) had free response answers indicative of looking at the big picture. The posttest also had three individuals referring to the big picture, the same two repeaters, and a different new person.

Transcriptions of recordings of the participants in the second workshop performing the activity in groups were also coded for themes. Two of the three repeaters referred to big picture ideas, for example, “the field of view grows as you look further away, causing the telescope to see more quasars further away.” This idea is a relatively advanced one; she had originally touched upon it in her first time participating in the study at the first workshop, and had developed it further in the one-on-one interview. Three of the five new participants also showed evidence of looking at the big picture, for example one of the men proposed looking for correlations in variables not haphazardly, but by examining the meaning of the variables: “About the only thing you could maybe find here is those that do have radio magnitudes, how do they compare in distance?”

**Discussion**

**Conclusions**

The participants in this study improved their scores from their pretests to their posttests from 73.4% to 85.9% with a P-value of less than 0.050. This indicates that they experienced a statistically significant gain in their exam scores. While the K-12 teacher participants started off with a larger basis of knowledge than do most college students in introductory astronomy classes (53% being more typical of that population, as per an accompanying paper by Schwortz, Burrows, and Myers, 2015a).

The participants’ matched normalized gains were 0.32 overall. Traditional (lecture-like) classrooms usually result in gains under 0.30, while active learning (or interactive engagement) classrooms’ gains are over 0.40 (Hake, 1998). These values however have been more commonly measured before and after an entire course, rather than before and after a single activity. With the small sample size here, the gains of 0.32 cannot be conclusively stated to be closer to those of traditional or active classrooms.

Men’s and women’s pretest scores were not significantly different. Women did not have a statistically significant increase from their pre- to post-tests, though men did. This difference between men’s and women’s scores is sustained in an accompanying study of college students in introductory astronomy completing the same activity with a larger sample size of 95 participants (Schwortz, et al., 2015a).

Out of the eight multiple choice questions on the pre-/post-test, there three questions with statistically significant improvement. Two of these questions dealt with vocabulary, one of the terms associated with spreadsheets and the other with the astronomical term quasar. The improvement on these questions can be easily explained by the participants’ increased familiarity with these terms through completing the activity which repeatedly used these terms in appropriate context.

**Implications**

There were multiple instances where observers during the activities noted that men participants had a tendency to talk over the women participants, for example interrupting the women as they talked to a greater extent than the women interrupted the men, or otherwise limiting the women’s participation in the group. In the second workshop, the group of five new participants sat so that the two men and one of the women were close enough to interact with the computer, while the other two women sat farther from the computer. These two tended to passively observe what the three louder participants in front were doing and quietly murmur to each other rather than interacting with the others.

While this group dynamic was not universal (one woman in this group for example did tend to ask leading questions that helped bring the two men to a deeper understanding), it is possible that it contributed to the fact that the men’s overall pre-/post-test scores increased more than did the women’s. On the third question with statistically significant improvement, which dealt with the deductive process that radio emission from a quasar indicates the possession of jets, it is possible that the statistically significant increase in the number of men answering this question from seven to 10 (P<0.050) is partially influenced by the fact that the men were overriding the women’s participation in the activity.

In this iteration of the study, participants were allowed to choose their own groups, and for the most part groups ended up being of mixed gender, sometimes consisting of participants who knew each other from their home schools or from previous workshops. A future iteration could investigate whether assigning same sex groups benefits the women participants.

Three of the six expert traits commonly identified (e.g., Bransford et al. 2000) have experts looking at the big picture of what they are studying. Of the 14 participants, nine exhibited an understanding that a big picture existed, an understanding which ideally all teachers should exhibit: in order for students to progress along the path of novice to expert themselves, they need models of expertise. Even should K-12 teachers be unaware of what the big picture is, they should still express to students that such a big picture exists and encourage students to seek it out themselves.

Those teachers who participated in the study a second time did show a qualitative progression in their thinking, indicating that teacher professional development in the sciences does produce observable positive results. While the sample size to this is small with only three individuals repeating the study, the implications are promising for teacher improvement in novice/expert characteristics as a result of professional development activities.

**Limitations**

Cross-evaluator differences between the four individuals taking field notes during the activity were observed. The subjects were already familiar with both the instructor administering the task and two out of the four field note takers. Due to constraints of the activity, it was not possible to determine inter-rater reliability.

This group may not represent teachers in general, since the sample participating in this study was self-selected as K-12 teachers who wished to attend a professional development opportunity focusing on science content. This could potentially lead to two disparate sub-populations within the subjects: teachers who are already science content experts and enjoy learning more, and teachers who are weak in the content area of science and either choose to or are required by administrators to participate in an attempt to shore up their content knowledge. There was definitely evidence that participants who repeated the activity had a higher baseline than did those who were participating for the first time, for example the repeat participants had a better grasp of the overarching ideas related to the data they were working with.

References

Abell, Sandra K., and Norman G. Lederman, ed. 2013. Handbook of Research on Science Education: ProQuest Tech Books. Focal Press.

Bransford, John D., Ann L. Brown, and Rodney R. Cocking. 2000. How People Learn: Mind, Brain, Experience, School. National Academies Press.

Hake, R. R. (1998). Interactive-engagement versus traditional methods: A six-thousand-student survey of mechanics test data for introductory physics courses. American Journal of Physics, 66(1), 64. doi:10.1119/1.18809

Larkin, Jill, John McDermott, Dorothea P. Simon, and Herbert A. Simon. 1980. “Expert and Novice Performance in Solving Physics Problem.” Science.

Lawrenz, F., Huffman, D., & Appeldoorn, K. (2005). Enhancing the instructional environment: optimal learning in introductory science classes. Journal of College Science Teaching, July/Augus, 40–44.

Schwortz, Andria C., Burrows, Andrea C., Myers, A. 2015. “Learning to Work with Databases in Astronomy: Quantitative Analysis of Science Educators' and Students' Pre-/Post-Tests.” American Astronomical Society Meeting 250. *Accepted*.

Schwortz, Andria C.; Eftekharzadeh, Sarah; Myers, Adam D.; Shen, Yue. 2015. “Quasar Clustering from SDSS DR7: Dependencies on FIRST Radio Magnitudes.” American Astronomical Society Meeting 250. Accepted.

Vygotsky, L. (1978). Social Constructivism. Mind in Society.

Westerman, D. A. (1991). Expert and Novice Teacher Decision Making. Journal of Teacher Education, 42(4), 292–305. doi:10.1177/002248719104200407