This study explored scientists' epistemological views of science and compared views based upon scientists' discipline area and primary investigative approach. Participants were 24 scientists, averaging 25 years research experience, representing four discipline areas: life science (10), Earth and space science (5), physics (5), chemistry (5); and four investigative approaches: experimental (10), nonexperimental (5), combination (5), theoretical (4). Views of nature of science (NOS) and scientific inquiry (NOSI) were assessed through two open-ended questionnaires, the VNOS-Sci and the VOSI-Sci, and interviews.

The analysis revealed 16 categories of scientists' NOS/NOSI views that are applicable across the science disciplines and contexts of this study. The results show that these participants' epistemological views of science are complex and sophisticated, "informed" in some areas, but not necessarily. On a level of broad generality, scientists'
views are as similar within as across groups, demonstrating overall consistency in how these scientists’ view the 16 categories of NOS/NOSI.

Views expressed are contextualized within science practices. However, there are variations in finer details of description and applicability. Some variation is related to contextual issues of discipline and/or research approach, yet no overarching pattern emerges to explain all the tendencies. With a few exceptions, variances are idiosyncratic, emerging at levels of specificity tied to individual contexts and experiences. Such finer levels of specificity and sophistication are deemed impractical for the K-12 science classroom. Exceptions include views of justification and reproducibility.

Results suggest that explicit/reflective instruction should target general NOS/NOSI instruction emphasizing connections among aspects and inquiry contexts. Variety in inquiry experiences is recommended. Teachers should raise awareness that some epistemological features of science demonstrate variability depending on the type of investigation and system under study. As such, learners need exposure and explicit/reflective instruction that promotes inclusive views of authentic science practices. Secondly, results demonstrate a variety of authentic science contexts are appropriate for addressing core features of and interdependencies among NOS/NOSI. Thirdly, the results suggest consensus on categories of NOS and scientific inquiry. Finally, this study enhances understanding of the scientific community and authentic practices of science; elements that enable teachers to connect real-world science to classroom science.
Epistemological Views in Authentic Science Practice: A Cross-Discipline Comparison of Scientists' Views of Nature of Science and Scientific Inquiry

By

Renée S. Schwartz

A THESIS

submitted to

Oregon State University

in partial fulfillment of

the requirements for the

degree of

Doctor of Philosophy

Presented June 7, 2004

Commencement June, 2005
I would like to extend my sincere gratitude and appreciation to my major professor and mentor, Norman G. Lederman, for his unyielding support and guidance. Thanks for giving me the opportunity to learn from the very best and for pushing me to be my best. I would like to express appreciation to my committee members, Kate Lajtha, Larry Enochs, Thomas Chastain, and a special thank you to Larry Flick, my Co-Chair. It has been a long haul, and I sincerely appreciate your support and advice that helped shape this work. My heartfelt appreciation goes to all the scientists who gave their valuable time and insights as participants in this study. They made their research come alive through their stories.

For accepting me into the research group, teaching me the ropes, and making me laugh, I thank Fouad Abd-El-Khalick and Randy Bell. I also thank Sandi Abell, for getting me started down this path so many years ago and for her friendship. Thank you to Judy Lederman, for her amazing energy, friendship, and inspiration. A special thank you goes to my friend and colleague, Mary Gfeller. Academy girls...We get stuff done! I am grateful for the opportunity to learn from you all.

Finally, I owe this degree to my family and friends. You gave me hope throughout it all. To my friends at Aurora Martial Arts, especially Lou and Wendy Donadio and Amy Strader, thanks for all your support, for your friendship, for the bruises, and for helping me know I can make anything happen. For all their love, I thank my parents, who sacrificed time, energy, and frequent flier miles to always be there for me. Finally, me deepest gratitude goes to my husband Christopher, and my two sons Christopher and Cameron, for their continuous love and understanding. Even in the toughest of times, their smiles and hugs remind me that everyday is a gift. Thank you.
# TABLE OF CONTENTS

**CHAPTER I: THE PROBLEM**

- Introduction .......................................................... 1
- Nature of Science ..................................................... 4
- Scientific Inquiry ..................................................... 8
- Statement of the Problem ........................................... 10
  - Inquiry as a Context to Teach Nature of Science ............. 14
  - Issues of Authenticity: Scientists’ Views of NOS and Scientific Inquiry ........................................ 17
- Research Questions ................................................ 20
- Significance of the Study .......................................... 21

**CHAPTER II: REVIEW OF THE LITERATURE** ........................................ 23

- The Assumption ...................................................... 23
- Nature of Science and Scientific Inquiry ....................... 24
- The Purpose .......................................................... 26
- A Critical Review of the Literature .............................. 26
  - The Influence of Inquiry-based Curricula on Students’ Conceptions of Nature of Science .................. 27
    - Summary ......................................................... 64
  - The Influence of Different Inquiry Teaching Approaches on Students’ Conceptions of Nature of Science .................. 67
    - Summary ......................................................... 103
  - The Influence of Inquiry-Based Experiences on Teachers’ and Preservice Teachers’ Views Nature of Science .................. 105
    - Summary ......................................................... 148
TABLE OF CONTENTS (continued)

Table: Scientists Views of Nature of Science

<table>
<thead>
<tr>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>150</td>
</tr>
</tbody>
</table>

Table: Summary

<table>
<thead>
<tr>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>180</td>
</tr>
</tbody>
</table>

Table: Discussion and Conclusions

<table>
<thead>
<tr>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>183</td>
</tr>
</tbody>
</table>

Table: Explicit Teaching of NOS through Inquiry: Necessary...but content-dependent?

<table>
<thead>
<tr>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>196</td>
</tr>
</tbody>
</table>

Table: Recommendations

<table>
<thead>
<tr>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>201</td>
</tr>
</tbody>
</table>

Table: CHAPTER III: DESIGN AND METHOD

<table>
<thead>
<tr>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>203</td>
</tr>
</tbody>
</table>

Table: Method

<table>
<thead>
<tr>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>203</td>
</tr>
</tbody>
</table>

Table: Participants

<table>
<thead>
<tr>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>204</td>
</tr>
</tbody>
</table>

Table: Data Sources and Analysis

<table>
<thead>
<tr>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>212</td>
</tr>
</tbody>
</table>

Table: Questionnaires

<table>
<thead>
<tr>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>212</td>
</tr>
</tbody>
</table>

Table: Views of Nature of Science Questionnaire

<table>
<thead>
<tr>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>212</td>
</tr>
</tbody>
</table>

Table: Views of Scientific Inquiry Questionnaire

<table>
<thead>
<tr>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>214</td>
</tr>
</tbody>
</table>

Table: Administration of the Questionnaires

<table>
<thead>
<tr>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>215</td>
</tr>
</tbody>
</table>

Table: Semi-structured Interview

<table>
<thead>
<tr>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>216</td>
</tr>
</tbody>
</table>

Table: Analysis of VNOS-Sci and VOSI-Sci Questionnaires and Interview

<table>
<thead>
<tr>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>219</td>
</tr>
</tbody>
</table>

Table: Views of NOS and Scientific Inquiry

<table>
<thead>
<tr>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>219</td>
</tr>
</tbody>
</table>

Table: Researcher Journal Analysis

<table>
<thead>
<tr>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>222</td>
</tr>
</tbody>
</table>

Table: Final Categorization of Participants

<table>
<thead>
<tr>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>223</td>
</tr>
</tbody>
</table>

Table: Answering of Research Questions

<table>
<thead>
<tr>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>223</td>
</tr>
</tbody>
</table>

Table: Research Question #1

<table>
<thead>
<tr>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>223</td>
</tr>
</tbody>
</table>

Table: Research Question #2

<table>
<thead>
<tr>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>224</td>
</tr>
</tbody>
</table>

Table: The Researcher

<table>
<thead>
<tr>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>225</td>
</tr>
<tr>
<td>Section</td>
</tr>
<tr>
<td>------------------------------------------------------------------------</td>
</tr>
<tr>
<td><strong>CHAPTER IV: RESULTS</strong></td>
</tr>
<tr>
<td>Research Approaches</td>
</tr>
<tr>
<td>Main Categories and Emergent Subcodes</td>
</tr>
<tr>
<td>Views of NOS and NOSI</td>
</tr>
<tr>
<td>Tentativeness</td>
</tr>
<tr>
<td>Scientists’ Views</td>
</tr>
<tr>
<td>Inherent tentativeness</td>
</tr>
<tr>
<td>Approaches certainty or attains certainty</td>
</tr>
<tr>
<td>Levels of certainty</td>
</tr>
<tr>
<td>Discipline-based comparisons</td>
</tr>
<tr>
<td>Approach-based comparisons</td>
</tr>
<tr>
<td>Empirical basis</td>
</tr>
<tr>
<td>Scientists’ Views</td>
</tr>
<tr>
<td>Empirically grounded</td>
</tr>
<tr>
<td>Varies: theoretical/mathematical valid basis</td>
</tr>
<tr>
<td>Discipline-based comparisons</td>
</tr>
<tr>
<td>Approach-based comparisons</td>
</tr>
<tr>
<td>Subjectivity</td>
</tr>
<tr>
<td>Scientists’ Views</td>
</tr>
<tr>
<td>Theory-laden</td>
</tr>
<tr>
<td>Qualitative methods are more subjective than</td>
</tr>
<tr>
<td>quantitative methods</td>
</tr>
<tr>
<td>Science is not subjective</td>
</tr>
<tr>
<td>Discipline-based comparisons</td>
</tr>
<tr>
<td>Approach-based comparisons</td>
</tr>
<tr>
<td>Creativity</td>
</tr>
<tr>
<td>Scientists’ Views</td>
</tr>
<tr>
<td>Affirming creativity</td>
</tr>
<tr>
<td>Finding patterns and connections</td>
</tr>
<tr>
<td>Progress/success</td>
</tr>
<tr>
<td>No role in making meaning</td>
</tr>
<tr>
<td>Discipline-based comparisons</td>
</tr>
<tr>
<td>Approach-based comparisons</td>
</tr>
<tr>
<td>Socio/cultural influences</td>
</tr>
<tr>
<td>Scientists’ Views</td>
</tr>
</tbody>
</table>
TABLE OF CONTENTS (continued)

| Political, economic, societal pressures on what questions are probed | 265 |
| Socio-cultural influences on the processes of science (how science is conducted and reasoning) | 268 |
| No effect of society or culture on what science is done | 271 |
| Effect of an international community on minimizing socio-cultural influences | 271 |
| Discipline-based comparisons | 272 |
| Approach-based comparisons | 272 |
| | |
| Theory/Law | 272 |
| Scientists' Views | 273 |
| Hierarchical relationship: Theories turn into laws with sufficient time and/or repeated testing | 273 |
| Different types of knowledge, not hierarchical | 278 |
| Differs by discipline | 279 |
| No laws in field of research | 280 |
| Discipline-based comparisons | 280 |
| Approach-based comparisons | 281 |
| | |
| Observation and Inference | 281 |
| Scientists' Views | 281 |
| Discipline-based and Approach-based comparisons | 283 |
| | |
| Models | 283 |
| Scientists' Views | 286 |
| Explain or organize observations/used for prediction/testing | 286 |
| Complex made simple/abstract made visual | 289 |
| Mathematics | 291 |
| Directing framework | 291 |
| Discipline-based comparisons | 292 |
| Approach-based comparisons | 292 |
| | |
| Experiment | 292 |
| Scientists' Views | 293 |
| Traditional | 293 |
| Hypothesis-driven | 293 |
| Experiments are not required for the development of scientific knowledge | 297 |
TABLE OF CONTENTS (continued)

Experiments are required for the development of scientific knowledge ........................................ 298
Experiments require replicas ........................................ 298
Experiments are variable by context ........................................ 299
Discipline-based comparisons ........................................ 300
Approach-based comparisons ........................................ 301

Purpose ........................................................................... 302
Scientists’ Views ........................................ 302
Discipline-based and Approach-based comparisons ........................................ 303

Methods ........................................................................... 303
Scientists’ Views ........................................ 303
  Single scientific method ........................................ 303
  Multiple scientific methods ........................................ 309
  Role of hypothesis: All hypothesis-driven ........................................ 312
  Role of hypothesis: Not all hypothesis-driven ........................................ 313
  Hierarchy of methods ........................................ 315
  Discipline differences: complexity of systems, ability to repeat ........................................ 318
Discipline-based comparisons ........................................ 319
Approach-based comparisons ........................................ 320

Anomaly ........................................................................... 320
Scientists’ Views ........................................ 321
  Initial identification of anomaly ........................................ 321
  Repeatable: True anomaly ........................................ 325
  Excitement/Progress ........................................ 326
  Expand existing model/framework ........................................ 327
  Develop a new model ........................................ 329
  May set aside, but do not disregard ........................................ 331
Discipline-based comparisons ........................................ 332
Approach-based comparisons ........................................ 333

Justification ........................................................................... 333
Scientists’ Views ........................................ 336
  Reproducibility ........................................ 336
  Experimentation over description ........................................ 337
  Prediction/test ........................................ 339
  Differs with discipline/context ........................................ 340
  Peer review ........................................ 342
  Address alternatives ........................................ 343
Discipline-based comparisons ........................................ 343
TABLE OF CONTENTS (continued)

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Approach-based comparisons</td>
<td>344</td>
</tr>
<tr>
<td>Data/Evidence</td>
<td>347</td>
</tr>
<tr>
<td>Scientists’ Views</td>
<td>347</td>
</tr>
<tr>
<td>Data as observations/ Evidence as interpretation</td>
<td>347</td>
</tr>
<tr>
<td>No difference between data and evidence</td>
<td>348</td>
</tr>
<tr>
<td>No use of evidence in field of research</td>
<td>348</td>
</tr>
<tr>
<td>Discipline-based and Approach-based comparisons</td>
<td>349</td>
</tr>
<tr>
<td>Reproducibility</td>
<td>349</td>
</tr>
<tr>
<td>Scientists’ Views</td>
<td>349</td>
</tr>
<tr>
<td>Statistics</td>
<td>349</td>
</tr>
<tr>
<td>Requirement to be scientific</td>
<td>351</td>
</tr>
<tr>
<td>Depends on method/discipline</td>
<td>351</td>
</tr>
<tr>
<td>Discipline-based comparisons</td>
<td>353</td>
</tr>
<tr>
<td>Approach-based comparisons</td>
<td>353</td>
</tr>
<tr>
<td>Prediction</td>
<td>354</td>
</tr>
<tr>
<td>Scientists’ Views</td>
<td>354</td>
</tr>
<tr>
<td>Affirmed important</td>
<td>354</td>
</tr>
<tr>
<td>Requirement to be scientific</td>
<td>357</td>
</tr>
<tr>
<td>Prediction through model use</td>
<td>359</td>
</tr>
<tr>
<td>Differs with discipline/context</td>
<td>360</td>
</tr>
<tr>
<td>Discipline-based comparisons</td>
<td>361</td>
</tr>
<tr>
<td>Approach-based comparisons</td>
<td>361</td>
</tr>
<tr>
<td>A final note about prediction</td>
<td>361</td>
</tr>
</tbody>
</table>

CHAPTER V: DISCUSSION AND IMPLICATIONS ........................................... 363

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Scientists’ epistemological views of science</td>
<td>363</td>
</tr>
<tr>
<td>Tentativeness</td>
<td>364</td>
</tr>
<tr>
<td>Empirical basis</td>
<td>366</td>
</tr>
<tr>
<td>Subjectivity</td>
<td>367</td>
</tr>
<tr>
<td>Creativity and Observation/Inference</td>
<td>369</td>
</tr>
<tr>
<td>Sociocultural Influence</td>
<td>370</td>
</tr>
<tr>
<td>Scientific Theories and Laws</td>
<td>373</td>
</tr>
<tr>
<td>Models</td>
<td>374</td>
</tr>
<tr>
<td>Experiments</td>
<td>376</td>
</tr>
<tr>
<td>Purpose</td>
<td>377</td>
</tr>
<tr>
<td>Methods</td>
<td>377</td>
</tr>
<tr>
<td>Anomaly</td>
<td>379</td>
</tr>
<tr>
<td>Justification</td>
<td>381</td>
</tr>
<tr>
<td>Distinguishing Data and Evidence</td>
<td>382</td>
</tr>
</tbody>
</table>
# TABLE OF CONTENTS (continued)

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Reproducibility</td>
<td>383</td>
</tr>
<tr>
<td>Prediction</td>
<td>384</td>
</tr>
<tr>
<td>Views of a Common Subcode: Variability by Context</td>
<td>386</td>
</tr>
<tr>
<td>Views of a Common Subcode: Statistics</td>
<td>387</td>
</tr>
<tr>
<td>Views of a Common Subcode: Hypothesis Driven</td>
<td>387</td>
</tr>
<tr>
<td>Consistency and Interdependence among aspects of NOS and NOSI</td>
<td>388</td>
</tr>
<tr>
<td>Are Variations in Scientists' Epistemological Views of Science Related to Authentic Science Context?</td>
<td>389</td>
</tr>
<tr>
<td>Scientists’ Epistemological Views of Science: Informed or Naïve?</td>
<td>391</td>
</tr>
<tr>
<td>Implications of the Study</td>
<td>394</td>
</tr>
<tr>
<td>The Adequacy of the Generalized Treatment of NOS and NOSI</td>
<td>394</td>
</tr>
<tr>
<td>Redefining “Generalized”</td>
<td>395</td>
</tr>
<tr>
<td>Informing Explicit Instruction: Pedagogical Content Knowledge beyond Generalizations to Promote Inclusive Image of Authentic Science Practice</td>
<td>396</td>
</tr>
<tr>
<td>Authentic Scientific Inquiry as a Context for Developing Conceptions of NOS and NOSI</td>
<td>399</td>
</tr>
<tr>
<td>Consensus on NOS and NOSI</td>
<td>400</td>
</tr>
<tr>
<td>Contemporary Science Practice: Multiple Contexts, Multiple Perspectives</td>
<td>400</td>
</tr>
<tr>
<td>Implications of Methodology: The Importance of Facilitating Reflection</td>
<td>401</td>
</tr>
<tr>
<td>Limitations of the Study</td>
<td>402</td>
</tr>
<tr>
<td>Recommendations for Future Research</td>
<td>404</td>
</tr>
<tr>
<td>REFERENCES</td>
<td>410</td>
</tr>
<tr>
<td>APPENDICES</td>
<td>422</td>
</tr>
<tr>
<td>Appendix A: Initial Query</td>
<td>423</td>
</tr>
</tbody>
</table>
TABLE OF CONTENTS (continued)

Appendix B: Informed Consent Form .................................................. 424
Appendix C: VNOS-Sci ................................................................. 425
Appendix D: VOSI-Sci ................................................................. 429
# LIST OF TABLES

<table>
<thead>
<tr>
<th>Table</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Description of scientist participants</td>
<td>209</td>
</tr>
<tr>
<td>2. Summary of scientists grouped by discipline</td>
<td>211</td>
</tr>
<tr>
<td>3. Scientists' views of NOS and NOSI</td>
<td>234</td>
</tr>
<tr>
<td>4. Scientists' views of tentativeness: Grouped by discipline</td>
<td>239</td>
</tr>
<tr>
<td>5. Scientists' views of tentativeness: Grouped by approach</td>
<td>240</td>
</tr>
<tr>
<td>6. Scientists' views of empirical basis: Grouped by discipline</td>
<td>249</td>
</tr>
<tr>
<td>7. Scientists' views of empirical basis: Grouped by approach</td>
<td>249</td>
</tr>
<tr>
<td>8. Scientists' views of subjectivity: Grouped by discipline</td>
<td>253</td>
</tr>
<tr>
<td>9. Scientists' views of subjectivity: Grouped by approach</td>
<td>254</td>
</tr>
<tr>
<td>10. Scientists' views of creativity: Grouped by discipline</td>
<td>260</td>
</tr>
<tr>
<td>11. Scientists' views of creativity: Grouped by approach</td>
<td>261</td>
</tr>
<tr>
<td>12. Scientists' views of socio/cultural embeddedness:</td>
<td></td>
</tr>
<tr>
<td>Grouped by discipline</td>
<td>266</td>
</tr>
<tr>
<td>13. Scientists' views of socio/cultural embeddedness:</td>
<td></td>
</tr>
<tr>
<td>Grouped by approach</td>
<td>267</td>
</tr>
<tr>
<td>14. Scientists' views of theory and law: Grouped by discipline</td>
<td>274</td>
</tr>
<tr>
<td>15. Scientists' views of theory and law: Grouped by approach</td>
<td>275</td>
</tr>
<tr>
<td>16. Scientists' views of observation and inference:</td>
<td></td>
</tr>
<tr>
<td>Grouped by discipline</td>
<td>282</td>
</tr>
<tr>
<td>17. Scientists' views of observation and inference:</td>
<td></td>
</tr>
<tr>
<td>Grouped by approach</td>
<td>282</td>
</tr>
<tr>
<td>18. Scientists' views of models: Grouped by discipline</td>
<td>284</td>
</tr>
<tr>
<td>19. Scientists' views of models: Grouped by approach</td>
<td>285</td>
</tr>
<tr>
<td>Table</td>
<td>Title</td>
</tr>
<tr>
<td>-------</td>
<td>----------------------------------------------------</td>
</tr>
<tr>
<td>20</td>
<td>Scientists' views of experiment: Grouped by discipline</td>
</tr>
<tr>
<td>21</td>
<td>Scientists' views of experiment: Grouped by approach</td>
</tr>
<tr>
<td>22</td>
<td>Scientists' views of purpose: Grouped by discipline</td>
</tr>
<tr>
<td>23</td>
<td>Scientists' views of purpose: Grouped by approach</td>
</tr>
<tr>
<td>24</td>
<td>Scientists' views of methods: Grouped by discipline</td>
</tr>
<tr>
<td>25</td>
<td>Scientists' views of methods: Grouped by approach</td>
</tr>
<tr>
<td>26</td>
<td>Scientists' views of anomaly: Grouped by discipline</td>
</tr>
<tr>
<td>27</td>
<td>Scientists' views of anomaly: Grouped by approach</td>
</tr>
<tr>
<td>28</td>
<td>Scientists' views of justification: Grouped by discipline</td>
</tr>
<tr>
<td>29</td>
<td>Scientists' views of justification: Grouped by approach</td>
</tr>
<tr>
<td>30</td>
<td>Scientists' views of data and evidence: Grouped by discipline</td>
</tr>
<tr>
<td>31</td>
<td>Scientists' views of data and evidence: Grouped by approach</td>
</tr>
<tr>
<td>32</td>
<td>Scientists' views of reproducibility: Grouped by discipline</td>
</tr>
<tr>
<td>33</td>
<td>Scientists' views of reproducibility: Grouped by approach</td>
</tr>
<tr>
<td>34</td>
<td>Scientists' views of prediction: Grouped by discipline</td>
</tr>
<tr>
<td>35</td>
<td>Scientists' views of prediction: Grouped by approach</td>
</tr>
<tr>
<td>Appendix</td>
<td>Page</td>
</tr>
<tr>
<td>----------------------------------------------------</td>
<td>------</td>
</tr>
<tr>
<td>A. Initial Query</td>
<td>423</td>
</tr>
<tr>
<td>B. Informed Consent Form</td>
<td>424</td>
</tr>
<tr>
<td>C. Views of Nature of Science Questionnaire: VNOS-Sci</td>
<td>425</td>
</tr>
<tr>
<td>D. Views of Scientific Inquiry Questionnaire: VOSI-Sci</td>
<td>429</td>
</tr>
</tbody>
</table>
This work is dedicated to my family,

*Christopher, Christopher, and Cameron.*

Your unconditional love, encouragement, and sacrifice made this possible.

To my parents,

*Richard and Karen Butt*

My biggest cheerleaders. You have never doubted me.

*And to all survivors.* Never give up.
“Science is a way of knowing that is characterized by empirical criteria, logical argument, and skeptical review. Students should develop an understanding of what science is, what science is not, what science can and cannot do, and how science contributes to culture.” (National Research Council [NRC], 1996, p. 21)

This statement encapsulates the most recent reforms in science education (American Association for the Advancement of Science [AAAS], 1993; NRC, 1996). These efforts reclaim the importance of introducing students to the culture of science wherein they can develop conceptual understanding of traditional science subject matter, nature of science (NOS), and scientific inquiry. The intercept of these three knowledge domains, along with an understanding of the utility of that knowledge to the individual and society, represents the conceptual foundation for a scientifically literate individual. In broad terms, this vision for scientific literacy involves:

“being familiar with the natural world and respecting its unity; being aware of some of the important ways in which mathematics, technology, and the sciences depend upon one another; understanding some of the key concepts and principles of science; having a capacity for scientific ways of thinking;
knowing that science, mathematics, and technology are human enterprises, and being able to use scientific knowledge and ways of thinking for personal and social purposes. (AAAS, 1989, p. 20)

Thus, the scientifically literate individual not only has conceptual knowledge of science subject matter, but also holds sound epistemological views about science that are consistent with current perspectives. Through developing scientific literacy, learners gain abilities to understand, communicate, and make informed decisions regarding scientific and technology-based issues prevalent in today’s society (e.g., AAAS, 1989, 1993; Driver, Leach, Millar, & Scott, 1996; McComas, 1998; NRC, 1996; Schwab, 1962).

An epistemological view refers to one’s understanding of what knowledge is and where knowledge comes from (von Glasersfeld, 1993). Epistemological views of science involve one’s view of scientific knowledge as a way of knowing and explaining the natural world (NOS) and one’s view of the nature and rationale of the processes through which that knowledge is constructed and justified within the scientific community (nature of scientific inquiry, NOSI). NOS and NOSI encompass basic philosophies, characteristics, and underlying assumptions that are intrinsic to scientific knowledge and the processes of its development, including the influences and limitations that result from science as a human endeavor. If one were to describe an individual’s “nature,” that description would include qualities that make that individual uniquely that individual. For science, the qualities that make scientific knowledge uniquely “scientific” and acceptable within the scientific community are what is referred to here as “NOS” and “NOSI.” Although separately
stated, NOS and NOSI are not mutually exclusive entities, for the process and the product are influentially connected.

Four decades ago Joseph Schwab (1962) stressed the importance of public understanding of the source and justification of scientific knowledge. He stated, “The knowledge won through enquiry is not knowledge merely of the facts but of the facts interpreted. And this interpretation, too, depends on the conceptual principle of the enquiry” (p. 14) [emphasis added] (It has been suggested that Schwab chose this particular spelling of “enquiry” to draw attention to its importance). Without understanding the values and assumptions that are inherent to scientific knowledge (NOS) and the processes by which the knowledge was created and accepted (scientific inquiry), the learner can do little more than construct an image of science consisting of isolated “facts” void of context that make the knowledge relevant, applicable, and meaningful. As such, science education reform efforts, past and present, encourage the teaching of scientific inquiry and NOS within the various disciplines of science (i.e. life science, physical science, chemical science, Earth science). The aim is to enable the development of disciplinary-based subject matter knowledge while promoting understandings of NOS and scientific inquiry that are representative of authentic scientific practices that create the subject matter.

To promote this vision of scientific literacy, science educators and national education policy and decision-makers have shifted curricular and instructional foci from the teaching of science as a final body of knowledge to teaching science as a human endeavor that produces a durable (empirically-based and internally consistent), yet provisional (tentative) understanding of the natural world (see, for
example, Duschl, 1990; Hodson, 1988). The shift has involved the learner as the constructor of his/her own knowledge through inquiry-based practices such as observation, data collection and analysis, model construction, and argumentation. The teacher acts as the guide rather than the final authority (NRC, 1996). Through engaging in content-embedded inquiry activities, the learner is to develop conceptual understanding of subject matter, scientific inquiry, and NOS (NRC, 1996). This claim is intuitively appealing from the perspective of situated cognition (Brown, Collins, & Duguid, 1989) wherein learning is viewed as an integral component of the activity of knowledge development itself. Cognition is contextually situated. To promote the development of scientific literacy through enhanced understanding of scientific inquiry and NOS, reform advocates recommend scientific inquiry experiences as a context for learning (AAAS, 1993; NRC 1996; 2000). “Inquiry is a critical component of a science program at all grade levels and in every domain of science, and designers of curricula and programs must be sure that the approach to content, as well as the teaching and assessment strategies, reflect the acquisition of scientific understanding through inquiry. Students then will learn science in a way that reflects how science actually works” (NRC, 1996, p. 214). To what extent, however, might variance within the scientific context influence resultant epistemological views? The present study begins to address this question in terms of variance in scientific discipline and methods of inquiry.

**Nature of Science**

Lederman (1992) has described NOS as the values and assumptions that are inherent in the development and application of scientific knowledge. In describing
NOS from epistemological and associated sociological perspectives, Ryan and Aikenhead (1992) included such topics such as: the meaning of science, assumptions, values, conceptual inventions, scientific method, consensus making, and the characteristics of the knowledge produced in science.

A concise meaning or description of NOS is often debated among philosophers of science, historians of science, and science educators (Loving, 1997; Matthews, 1994). Various representations illustrated among these scholars are as dynamic as the knowledge and enterprise of science itself. In delving into the philosophical depths of “the” NOS, one would likely find the phrase “natures of science” to be more relevant. Loving (1997) described past and current perspectives that have been voiced in social science research, exploring a “positivist-to-postmodernist shift” (p. 422) in perspectives. The current NOS perspective is in contrast with the traditional positivist view that describes science as authoritative, objective, and free of cultural influences (Casti, 1989; Chalmers, 1982). The current, postmodern, perspective is based on the works of Kuhn (1970), Hanson (1958), and other philosophers who purport more relativist views of science (see, for example, Casti, 1989; and Chalmers, 1982 for reviews). The postmodern view acknowledges science as a human endeavor, directed by theory and culture, reliant on empirical observation, and subject to change. As previously stated, there are still areas of disagreement among the scholars, such as the ontological basis of scientific knowledge. However, there is agreement on certain aspects of NOS that have been deemed relevant and important to include in k-16 education (Matthews, 1994).
Institutions of reform such as AAAS (1990, 1993), the NRC (1996), as well as science education efforts outside the US (see, for example, Hodson, 1998; Matthews, 1998; Millar, 1989; Millar & Osborne, 1998; Ryan & Aikenhead, 1992) have presented descriptions of NOS that include common generalities. For example, as described in Millar (1989), the National Science Curriculum in the United Kingdom states that

Pupils should develop an understanding that science is a human activity, that scientific ideas change through time, and that the nature of scientific ideas and the uses to which they are put are affected by the social and cultural contexts in which they are developed. (p. 2)

The National Science Education Standards (NRC, 1996) includes “Nature of Scientific Knowledge” in the grade 9-12 content standards “History and Nature of Science.” The recommended aspects of NOS relevant for grade 9-12 students include:

- Science distinguishes itself from other ways of knowing through the use of empirical standards, logical arguments, and skepticism, as scientists strive for the best possible explanations about the natural world.

- Scientific explanations must meet certain criteria. First and foremost, they must be consistent with experimental and observational evidence about nature, and must make accurate predictions, when appropriate, about systems being studied. They should also be logical, respect the rules of evidence, be open to criticism, report methods and procedures, and make knowledge public. Explanations on how the natural world changes based on myths,
personal beliefs, religious values, mystical inspiration, superstition, or authority may be personally useful and socially relevant, but they are not scientific.

- Because all scientific ideas depend on experimental and observational confirmation, all scientific knowledge is, in principle, subject to change as new evidence becomes available. (p. 201)

Additionally, numerous science educators have reviewed reform documents and literature from philosophy of science to identify aspects of NOS that most commonly appear and pose little disagreement according to current philosophical perspectives (Driver et al., 1996; Hodson, 1998; Lederman & Abd-El-Khalick, 2000; Millar, 1989; Millar & Osborne, 1998; Ryan & Aikenhead, 1992; Smith, Lederman, Bell, McComas, & Clough, 1997; Smith & Scharmann, 1999; among others). Chief among these is that scientific knowledge is *tentative* or subject to change. Reasons for the inherent tentativeness of scientific knowledge stems from several other aspects including: (a) scientific knowledge has basis in *empirical evidence*, (b) collection and interpretation of empirical evidence is influenced by current scientific perspectives (*theory-laden* observations and interpretations) as well as *personal subjectivity* due to scientists’ values, knowledge, and prior experiences, (c) scientific knowledge is the product of human *imagination and creativity*, and (d) the direction and products of scientific investigations are influenced by the society and culture in which the science is conducted (*sociocultural embeddedness*). Additional important considerations to NOS include (e) the differences between *observation and inference*
in the development of scientific knowledge, and (f) the differences between and functional roles of scientific *theories and laws*.

**Scientific Inquiry**

"Scientific inquiry" refers to the characteristics of the processes through which scientific knowledge is developed, including the conventions involved in the development, acceptance, and utility of scientific knowledge. The NSES states:

[Inquiry] involves making observations; posing questions; examining books and other sources of information to see what is already known; planning investigations; reviewing what is already known in light of experimental evidence; using tools to gather, analyze, and interpret data; proposing answers, explanations, and predictions; and communicating the results. Inquiry requires identification of assumptions, use of critical and logical thinking, and consideration of alternative explanations. (NRC, 1996, p. 23)

Like NOS, the meaning of "science as inquiry" has been debated for decades, and precise descriptions of what inquiry means for science education seem to vary as much as the methods of inquiry (Bybee, 2000). The National Academy of Sciences (2002) has identified "guiding principles for scientific inquiry" that serve a common basis across disciplines of research including political science, geophysics, and education. These principles are similar to the items on the list of generalized inquiry skills identified in the NSES (NRC, 2000) as part of the content standards for "science as inquiry." The NSES content standards for inquiry include generalized inquiry skills (e.g. identify a question that can be answered through scientific investigations; design and conduct a scientific investigation; use appropriate tools
and techniques to gather, analyze, and interpret data" (p. 19) and a list of statements termed "fundamental understandings about scientific inquiry (e.g. "different kinds of questions suggest different kinds of scientific investigations; current scientific knowledge and understanding guide scientific investigations" (p. 20). To highlight the importance of understanding the nature of scientific inquiry, The Benchmarks for Science Literacy (AAAS, 1993) emphasizes what students should know in relation to inquiry, rather than skills of inquiry.

The NRC (2000), AAAS (1993), the National Academy of Sciences (2002), science educators (Chinn & Malhotra 2002; Reiff, Harwood, & Phillipson, 2002; Hodson, 1998; Minstrell & van Zee, 2000), and researchers who have explored scientists in practice (e.g. Dunbar, 2001; Knorr-Cetina, 1999; Latour & Woolgar, 1979) offer descriptions of scientific inquiry, beyond basic investigative skills, that share commonalties. These commonalties are utilized here as a basis for agreed upon aspects of the nature of scientific inquiry (NOSI) deemed relevant and important for science education. Furthermore, the identified common aspects are consistent with those set forth in general terms by Schwab 40 years ago. The general aspects of nature of scientific inquiry include: a) multiple methods of scientific investigations, b) multiple purposes of scientific investigations, c) the form and role of argumentation in the development and acceptance of new knowledge, d) recognition and handling of anomalous data, e) sources, roles of, and distinctions between data and evidence, and f) community of practice. The relevance of "community of practice" to science education is within processes of peer review and communication that are associated with negotiating meaning within a social setting.
Statement of the Problem

Within the framework of the reform’s vision, there is an implied connection between NOS, scientific inquiry, and science subject matter. The recommendation to teach these three components of science concurrently encourages the presentation of science in an authentic light, rather than through artificial separation that serves to decontextualize science into mere “rhetoric of conclusions” (Schwab, 1962).

Connecting NOS, skills of scientific inquiry, knowledge about scientific inquiry, and traditional science subject matter is suggested to enhance an overall understanding of science, and thus, scientific literacy (Driver et al., 1996; Lederman, 1998; Schwab, 1962).

This consideration of context is well recognized within cognitive science research as an integral part of learning (e.g. Brown et al., 1989; Lave & Wenger, 1991). According to Brown et al. (1989),

Activity, concept, and culture are interdependent. No one can be totally understood without the other two. Learning must involve all three. Teaching methods often try to impart abstracted concepts as fixed, well-defined, independent entities that can be explored in prototypical examples and textbook exercises. But such exemplification cannot provide the important insights into either the culture or the authentic activities of members of that culture that learners need. (p. 33)

Within this perspective, knowledge is linked to activity and the situation under which the knowledge is acquired. Does this perspective apply to knowledge of NOS and scientific inquiry?
Throughout reform documents and science textbooks, scientific disciplines are typically distinguished based on topics of study such as physics, life science, Earth and space science, and chemistry. However, scientific disciplines differ by more than classification of subject matter. For example, Bechtel (1986) describes three dimensions that define scientific disciplines: (1) the objects studied, (2) the cognitive activities involved, and (3) the social and institutional organization. Additionally, studies of scientific activity and reasoning within research laboratory settings have provided descriptions of cognitive activities, inquiry methods, and social and institutional organizations of scientific communities that support the notion of diversity among disciplines that extend beyond subject matter (e.g. Dunbar, 2001; Knorr-Cetina, 1999; Latour & Woolgar, 1979).

Science education literature and reform documents recognize variation in methods and purpose of investigations within science. The NSES states, “Scientific inquiry refers to the diverse ways in which scientists study the natural world and propose explanations based on the evidence derived from their work” (NRC, 1996, p. 23). The education community also recognizes variances according to scientific discipline. In addressing the question of “What makes good science good?” Ault (1998) described variance in the “criteria of excellence” that serve as the basis for judging the merits of scientific findings among fields of study. He stated,

What makes particular fields of inquiry distinct resides in the uniqueness of their criteria of excellence or in their related strengths of attachment to criteria of excellence they hold in common with other fields. In progressing through time, fields of inquiry produce their own methods and strategies for
solving problems, and these methods and strategies are characteristic of the phenomena under scrutiny. Geology is not physics. Ecology is in very important ways not reducible to chemistry. Neither geology nor ecology is evolving to become better adapted to the criteria of excellence governing the search for greater unity in physics theory. Fields produce and follow their own criteria, appropriate to the class of problems defined in large measure by how the conceptual apparatus of the field imposes boundaries – inventions of the mind – upon experiences. (p. 190-191)

Furthermore, philosophers of science and scientists recognize distinctions in the nature of scientific disciplines (e.g. Mayr, 1988; 1997; Ruse, 1998; Spieker, 1972; Ziman, 2000). These discussions have included topics such as degrees of tentativeness based on differences in inference and abstractness, such as that between geology and chemistry or physics (Van Bemmelen, 1961), with the former being regarded as more inferential due to the nature of the evidence and limitations of experimentation. Spieker (1972) presented a comparison of specialized areas under the domain of natural science based on number of variables. He arranged fields of natural science in the order: physics-chemistry-geology-botany-zoology, to indicate increasing number of variables, decreasing power of mathematics, and a transition from concentrative fields (physics and chemistry) to more distributive fields. The concentrative fields of physics and chemistry, Spieker (1972) argues, have been able to “derive fundamental and universal laws from relatively small and concentrated bodies of data, or so to support generalizations reached through deductive reasoning, this is rarely possible in geology and biology” (p. 75). He states that geology and
biology typically require large and diverse amounts of data in order to make
generalizations. Mayr (1988; 1997) also presents a distinction in laws in the physical
sciences versus biological science. "The so-called laws of biology are not the
universal laws of classical physics but are simply high-level generalizations" (p. 19).

Mayr (1988) described biology as consisting of two types of study,
distinguished by the type of causation addressed. One line is functional biology and
addresses questions of proximate causation. The second is evolutionary biology and
addresses questions of ultimate causation. "The functional biologist is vitally
concerned with the operation and interaction of structural elements, from molecules
up to organs and whole individuals. His ever-repeated question is 'How?' ... The
evolutionary biologist differs in his method and in the problems in which he is
interested. His basic question is 'Why?'" (p. 25). Moreover, Mayr sees the questions
and methods of the evolutionary biologists as distinct from those encountered in
other disciplines of science. "There is nothing in the physical sciences that
corresponds to the biology of ultimate causations" (p. 17).

It is important to note that the purpose here is not to delve into a
philosophical debate about THE nature of science or of scientific disciplines that
extend beyond what is advocated for K-16 education. The point is to acknowledge
the presence of diversity in content, methods, and institutional structures of the
scientific endeavors that lead to the development of the subject matter included in K-
16 science education.

Despite recognition of variance among and within the sciences,
recommendations for what a scientifically literate individual should understand about
NOS and be able to do and understand about scientific inquiry are typically
generalized in a discipline-independent manner (AAAS, 1993; Abd-El-Khalick,
Lederman, & Bell, 1998; Chinn & Malhotra, 2002; Elby & Hammer, 2001; NRC,
1996, 2000). That is, the science education literature describes aspects of NOS and
scientific inquiry (NOSI) in a manner specific to the domain of science, but not
specific for particular disciplines within science. Recognizing possible conflicts for
science education in keeping to strict generalized applications of positivist and/or
postmodern descriptions of NOS and scientific inquiry, Loving (1997) poses several
important questions for science educators to examine in order to present a balanced
view of science in the science classroom. These questions included, “How are
theories generated? How does evidence bear on theory and theory on evidence?
What is the relationship between methods, aims, and theories?” (p. 435) and most
relevant to the present study, “How are physics, biology, or geology the same or
different in the way questions get answered, and what kind of evidence is acceptable
to justify theories?” (p. 435). These questions and the role of NOS and scientific
inquiry in the push for scientific literacy leads to the additional question of “What
impact might different disciplinary and investigative contexts have on one’s
epistemological views of science?”

Inquiry as a Context to Teach Nature of Science

The literature demonstrates students and teachers persistently maintain naïve
views of NOS (Benson, 1989; Carey & Stauss, 1968; Driver et al., 1996; Duschl,
1990; Duschl & Wright, 1989; Gallagher, 1991; Kimball, 1967-68; Lederman, 1992;
Meichtry, 1992; Pomeroy, 1993; Ryder, Leach, & Driver, 1999; Ryan & Aikenhead,
Schwab (1962) attributed public (mis)understanding of science to the traditional science teaching methods that emphasize “unmitigated rhetoric of conclusions in which current and temporary constructions of scientific knowledge are conveyed as empirical, literal, and irrevocable truths” (p. 24). Lack of connection to the scientific community and experiences in conducting scientific investigations may contribute to teachers’ and students’ naïve NOS views (i.e. views inconsistent with currently accepted perspectives) (Gallagher, 1991; Harms & Yager, 1981; Shapiro, 1996; Schwab, 1962; Welch, Klopfer, Aikenhead, & Robinson, 1981). Recommendations for improving learners’ conceptions of NOS and scientific inquiry strongly emphasize the use of inquiry as an instructional strategy (AAAS, 1990, 1993; NRC, 1996; Schwab, 1962; Welch et al., 1981). The argument for the use of inquiry asserts that the engagement in scientific inquiry activities similar to those of scientists, or along with scientists, will help the learner develop an understanding of the methods and activities through which science progresses, and, in turn, develop an understanding of NOS and NOSI. Thus, inquiry activities provide a viable context for developing desired conceptions associated with scientific literacy.

Education in science is more than the transmission of factual information: it must provide students with a knowledge base that enables them to educate themselves about the scientific and technological issues of their times; it must provide students with an understanding of the nature of science and its place in society; and it must provide them with an understanding of the methods and processes of scientific inquiry. To achieve these goals, science should be taught as science is practiced at its best (AAAS, 1990, p. xii).
Engaging learners in scientific inquiry for the purpose of teaching them about NOS and NOSI has intuitive appeal, and is consistent with a situated cognition perspective (Brown et al., 1989; Lave & Wenger, 1991). The assumption would seem that the more authentic the inquiry experience, the more representative the resultant view of NOS and inquiry would be of authentic science. Yet, the empirical research does not support this claim. Regarding NOS learning outcomes, the research suggests that regardless of the type of inquiry (guided or open) or extent of immersion within the scientific community (participant observation of scientific community or apprenticeship model), students do not necessarily develop informed conceptions of NOS (Crawford, Bell, Blair, & Lederman, 1999; Crumb, 1965; Haukoos & Penick, 1983, 1985; Khishfe & Abd-El-Khalick, 2002; Spears & Zollman, 1977; Trent, 1965; Tamir, 1972; Yager, Engen, & Snider, 1969; Yager & Wick, 1966). The research makes similar claims relative to developments in teachers' views of NOS through inquiry experiences (Barafaldi, Bethel, & Lamb, 1977; Bianchini & Colburn, 2000; Billeh & Hasan, 1975; Haukoos & Penick, 1983, 1985; Meichtry, 1992; Riley, 1979; Schwartz, Lederman, & Crawford, 2000, in press; Westerlund, Schwartz, Lederman, & Koke, 2001). This line of research demonstrates insufficient NOS learning outcomes when NOS is viewed as tacit knowledge, gained solely from engagement in inquiry activities. This work suggests that experiences with scientific investigations must be accompanied by explicit instruction and reflection opportunities that encourage the learner to formalize his/her views about NOS in relation to those experiences. It should be noted that while these results are consistent across levels of inquiry (e.g. guided versus open),
none of this work purposely sought discipline-based comparisons. Questions about contextually-based variance in views of NOS have been raised by the science education community, but not purposely explored or extended to address views about scientific inquiry (Brickhouse, Dagher, Shipman, & Letts, 2002; Loving, 1997; Ryder et al., 1999).

The Issue of Authenticity: Scientists’ Views of NOS and Scientific Inquiry

The emphasis on inquiry as a pedagogical approach for teaching about NOS and the results of research on the effectiveness of inquiry experiences on NOS learning outcomes leads to several questions about the meaning of scientific inquiry and the generalized treatment of NOS and inquiry in science education. To what extent have the subjects in these studies engaged in authentic scientific inquiry as experienced by practicing scientists within the scientific community? Given the notion that authentic scientific inquiry is that which occurs within the scientific community by practicing scientists (Roth, 1995; 1997), school-based scientific inquiry cannot be considered “authentic” in the strictest sense. Even though the literature suggests that learners are engaging in activities similar to those of scientists, there remain fundamental differences between authentic scientific inquiry and scientific inquiry as experienced by students and teachers within a classroom community (Brown, et al., 1989; Driver et al., 1996; Roth & Roychoudhury, 1993; Roth, 1991; Ryder, Leach, & Driver, 1999; Samarapungavan, 1992). Although both contexts provide opportunities for social construction of understanding, the school community rarely promotes the complexity of reasoning and negotiation of meaning as it is expressed within the scientific community (Chinn & Malhotra, 2002; Lave &
Wenger 1991; Wenger, 1998). Furthermore, students and teachers are not experts in subject matter or skills of inquiry to the extent that practicing scientists are.

A logical next step, therefore, is an examination of scientists’ epistemological views of science to offer some insight into conceptions that may correlate with engagement in truly authentic scientific inquiry within the community of science. However, the few existing reports lack descriptive detail. In short, the reports of scientists’ NOS views are similar to reports of teachers’ and students’ views. They suggest that scientists do not necessarily hold epistemological views that align with currently accepted views advocated for K-16 science education (Bell, 2000; Glasson & Bentley, 2000; Kimball, 1967-68; Pomeroy, 1993). Glasson and Bentley (2000) offer a rationale for scientists’ views:

The overriding view among practicing scientists is that science is essentially experimental and empirical; however, the important role of theory, the multiplicity and complexity of science methods, and the value-ladenness of science require that scientists examine the assumptions underlying their own research and what goes into the decision-making that affects research design, funding, and public acceptance of results. (Glasson & Bentley, 2000, p. 483)

Regarding scientists’ views of scientific inquiry, one recent study suggests scientists in various disciplines share a common view of their approach to inquiry (Reiff, Harwood, & Phillipson, 2002). However, the proposed commonly viewed aspects of inquiry were very general such as “ask questions” and “making observations.” These are general skills scientists perform. This study did not include more detailed descriptions of scientists’ views concerning valid evidence, how
scientists’ deal with anomalous data, and how findings within their research area gain acceptance. These latter features are more aligned with epistemological views relative to the nature of scientific inquiry, and are not necessarily represented or understood in equal ways across research settings (Chinn & Malhotra, 2001; Dunbar, 2001; Knorr-Cetina, 1999; Ziman, 2000). Furthermore, “science is a complex social activity’ (AAAS, 1989) that progresses within a community that has established social processes and institutional norms” is a commonly stated feature of science epistemology relative to scientific inquiry (e.g. AAAS, 1989; Chinn & Malhotra, 2001; 2002; Knorr-Cetina, 1999; NRC, 2000; Ziman, 2000). Knorr-Cetina (1999) described how the social and institutional structure of different research groups was reflected in how scientists viewed their community and individual roles within that community.

Few scientists are actively reflective in their daily research, and there seems to be little relation between such reflection and successful scientific practice (Elby & Hammer, 2001; Glasson & Bentley, 2000). Yet given that scientists engage in specialized areas of scientific inquiry and specific subject matter, it would seem a valuable venue to examine scientists’ views of NOS and scientific inquiry and provide in-depth details of practicing scientists’ views. This study addresses the issue of views of NOS and scientific inquiry that exist through active participation in authentic scientific practices. Such descriptions are lacking in the existing literature base.

Given the emphasis on relating NOS and NOSI through inquiry-based learning experiences throughout science disciplines, and given the recognition of the
situated nature of learning, there is a need to explore potential contextual connections of epistemological views of science within and across scientific disciplines. The present study is a first step in this effort. Potential contextual connections were explored by examining the views of scientists from different scientific disciplines (e.g. life science, physical science, Earth science). The focus on the scientists as the participants enables a descriptive study relating authentic contexts of inquiry and corresponding views of NOS and NOSI, without limitations of subject matter knowledge or scientific inquiry skills relevant to the field of specialization.

Research Questions

This study is exploratory in nature, and presents details of scientists' epistemological views concerning NOS and scientific inquiry. Through intra- and cross-discipline comparisons of scientists' views, this study explores associations between epistemological views of science, scientific disciplines, and methods of scientific inquiry. There are two primary research questions:

1. What are practicing scientists' views of nature of science and scientific inquiry?
2. What are the relationships, if any, among practicing scientists' views of nature of science, views of scientific inquiry, approach to scientific inquiry, and the scientific discipline in which the scientist participates?
Significance of the Study

The need for this study stems from the aforementioned generalized treatment of NOS and scientific inquiry within science education. There is a void in the education literature of contextually-based explorations of epistemological views of science. Teachers are to engage students in inquiry activities and associate the activities with authentic scientific practices and relevant NOS and inquiry issues. Yet, current expectations regarding teaching of NOS and nature of scientific inquiry do not consider variations in epistemological views of science based on engaging in different types of investigative approaches or different scientific disciplines.

"In vivo" studies of scientists' practice (Dunbar, 2001) provide authentic context for describing communities of practice (Wenger, 1998) and the epistemological views that correspond to engaging in such activity. Yet, all authentic activity in science is not the same. Knorr-Cetina (1999) referred to the epistemic cultures of science, where "epistemic cultures" are "cultures that create and warrant knowledge" (p. 1). Through her two-year ethnographic investigation, she distinguished the epistemic cultures of a molecular biology laboratory and a high-energy physics laboratory through depictions of distinct processes and structures of research practices, collaborative structures, and identities (Knorr-Cetina, 1999). As mentioned, different scientific fields have their own "criteria of excellence" (Ault, 1998) and conventions determined by their community (Knorr-Cetina, 1999; Ziman, 2000). As such, perceptions of valid "evidence" and of the role of inference and creativity in scientific endeavors may also vary. Variance in action, purpose, and acceptance standards may translate into variance in how scientists' perceive NOS
and the nature of scientific inquiry. The present study is a beginning exploration of context-based epistemological views of science that will help clarify the appropriateness of the generalized treatment of NOS and scientific inquiry in science education.

The empirical research on the effectiveness of engaging in inquiry activities on NOS and inquiry views, and the few reported investigations of scientists' epistemological views, clearly indicate a need for explicit/reflective instruction for NOS and inquiry in K-16 education. The present study helps to identify possible variations based upon the science discipline and investigative methods that may facilitate context-specific explicit teaching of NOS and scientific inquiry.

Finally, this study leads to better overall understanding of the scientific community and authentic practices of science; elements that enable teachers to better connect real-world science to classroom science. Glasson and Bentley (2000) suggest dialogue between scientists and educators to enhance our understanding of the scientific community and scientists, and thereby enabling educators to increase the relevance of scientific research to the classroom. The proposed study incorporates dialogue between educator and scientists. Increased understanding of authentic scientific practice will inform curricular revision to more closely simulate authentic science and help teachers draw parallels between authentic science and classroom-based science (Chinn & Malhotra, 2002; Schwartz & Crawford, 2004).
CHAPTER II
REVIEW OF THE LITERATURE

The primary purpose of the present literature review is to investigate the NOS understandings of individuals who have engaged in scientific inquiry activities. A question of particular interest is whether or not people who engage in different types of inquiry possess different views of NOS. Because very few studies have explored how individuals understand the nature of scientific inquiry, this literature review mainly focuses on views of NOS that correlate with inquiry contexts and pedagogical approaches.

The Assumption

Reform efforts have proposed that by engaging in inquiry-oriented activities, learners will develop conceptions of NOS that are in agreement with accepted contemporary perspectives. Inherent to the abundance of subsequently developed curricula has been the assumption that inquiry-based instruction alone will lead to an intuitive understanding of NOS. Thus, the question, “Do individuals necessarily develop conceptions of NOS consistent with the currently accepted, or contemporary, views by engaging in scientific inquiry activities?” is affirmed by advocates of an implicit approach to the use of inquiry in science instruction as a means to teach NOS. Implicit refers to the absence of direct attention to NOS in relation to the activities. Within any science instruction, implicit messages about NOS may be communicated. However, the proper communication of these implicit
messages during inquiry-based laboratory activities is viewed, by some science educators, as a natural consequence of participation in the activities. The assumption is that by “doing science,” learners will acquire an image and understanding of NOS (and inquiry) that is in accordance with the learning objectives posed in the reform documents. Such an implicit approach was advocated by many of the inquiry-based curricula developed in response to the reform efforts since the 1960s (Hodson, 1988). The first portion of this literature review examines the validity of the assumption.

Nature of science and Scientific Inquiry

The concepts and contemporary perspectives of NOS and scientific inquiry have been described in Chapter I of this thesis. Because these concepts are dynamic, the studies reviewed here do not necessarily hold the same perspectives or present aspects of NOS or inquiry with clarity. The literature represents diverse perspectives and functions of inquiry and NOS. These studies examine developing “views about science” (combining NOS and scientific inquiry), science processes, or NOS when scientific inquiry is the pedagogical approach.

Teaching “science as inquiry” has been described as comprising three main elements: (1) teaching skills of scientific inquiry (what students should be able to do), (2) teaching about scientific inquiry (what students should understand about the nature of scientific inquiry), and (3) teaching science content through inquiry (Bybee, 2000; NRC, 2000). The NSES presents content standards for both skills of inquiry and knowledge about inquiry. In addition, in the Science Teaching
Standards, the NSES discusses the third meaning of "science as inquiry." Through inquiry teaching strategies, as outlined in the NSES, students learn science content. When one considers that NOS, and inquiry for that matter, are included as "science content," it becomes apparent that teaching NOS through inquiry makes sense. The perspective assumed for the present review of the literature is that NOS is considered the content and inquiry is the teaching strategy employed to teach that content.

Although there are distinctions between scientific inquiry and NOS, the two are not mutually exclusive. The conjunction between understanding the nature of scientific inquiry and understanding NOS is the perception that scientific inquiry is a creative process, driven by currently accepted theories and laws of the scientific community. The explanations of empirical evidence resulting from such inquiries are tentative and only as accurate, in an ontological sense, as the theories and laws upon which the inquiries were based. Hence, the interdependence of the nature of scientific inquiry and NOS may contribute to the intuitive appeal of using inquiry as an instructional approach to teach NOS. Yet, the distinction must be made between the process and nature aspects of the enterprise wherein NOS more specifically refers to the theory-laden assumptions and constraints of perceptual apparatus (Abd-El-Khalick, Bell, & Lederman, 1998) that are involved in the processes by which scientific knowledge is developed. This distinction has often been muddled and overlooked in the development and implementation of curricula and instructional materials intended to engender an understanding of NOS in learners.
The Purpose

The purpose of this critical review is to provide an overlook of published research on NOS views or "views about science" that have been examined in the context of scientific inquiry. This review examines the relative effectiveness of various pedagogical approaches and contexts. The criteria for inclusion in this review were: (1) subjects of the studies engaged in or intended to engage in some type of scientific inquiry-based activity, (2) researchers aimed to assess or describe subjects' views of NOS after engagement in the inquiry-based activity, and (3) the research report is published in a peer reviewed journal.

The literature on the use of inquiry to teach NOS contains various meanings of the phrase "NOS" and related phrases (e.g. "understanding of science"). The purpose here is not to evaluate the authors' interpretations of "NOS" relative to current perspectives. The purpose here is to critically review studies that have intended to look at learners' views of NOS, in whatever way the authors describe, in the context of inquiry-based activity, or what was intended to be inquiry-based activity. Likewise, these reports vary in their use of the term "inquiry" and its application in the classroom. As is possible, authors' descriptions of inquiry and the activities of the subjects are provided.

A Critical Review of the Literature

This critical review examines reports of the influence of scientific inquiry activities and programs involving scientific or science-related inquiry experiences on individuals' conceptions of NOS. The articles included in this review are divided
into four sections based on their primary focus: (1) The influence of inquiry-based curricula on students’ conceptions of NOS: This section examines the effectiveness of early curricula developed with the stated objective of enhancing students’ views of NOS through their engagement in inquiry activities. These reports compare the outcomes of the inquiry-based curricula with outcomes of traditional science instruction. (2) The influence of different inquiry teaching approaches on students’ conceptions of NOS: This section examines reports that compare NOS learning outcomes within various types of inquiry-based teaching strategies, including open vs. guided inquiry and active vs. inactive participation in inquiry. (3) The influence of inquiry-based experiences on teachers’ and preservice teachers’ views of NOS: This section examines reports of teachers’ NOS learning outcomes that developed within the context of inquiry-based experiences. (4) Scientists’ views of NOS: To address concerns of constraints due to the context of teacher preparation or classroom-based inquiry, this section examines NOS views held by scientists involved in authentic scientific inquiry. The Discussion and Conclusions section then relates these four sections in terms of the differential effectiveness of inquiry experiences in developing views of NOS.

The Influence of Inquiry-based Curricula on Students’ Conceptions of NOS

Efforts emphasizing an inquiry approach to science instruction are not new. According to Yager (1997), early efforts in science education reform stressing the process skills of scientists were seen in the 1930s. In the wake of the Sputnik success of 1957, the push for improved science achievement and advancement in the United States sparked the development of a multitude of science curricula that focused on
scientific inquiry. The main emphasis was toward active learning, stressing hands-on laboratory experiences for the students. The shift was away from textbook-centered toward student-centered classrooms, with the students engaging in inquiry-based laboratory activities rather than rote “book” learning (Sund & Trowbridge, 1973). The curricula intended to be inquiry approaches to science, encouraging students to make observations, define problems, formulate hypotheses, test hypotheses, interpret data, formulate inferences and make generalizations (Shulman & Tamir, 1973; Sund & Trowbridge, 1973). Included in the stated laboratory objectives of the 1960s secondary science curricula, such as BSCS (Biological Sciences Curriculum Study) biology and PSSC (Physical Science Study Committee) physics, was the objective of enhancing students’ understanding of NOS. The specific aspects of NOS indicated as important for the students to comprehend pertained to “the scientific enterprise, the scientists and how they work, the existence of multiplicity of scientific methods, the interrelationship between science and technology and among the various disciplines of science” (Shulman & Tamir, 1973, p. 1119). It is apparent from this list that NOS and nature of inquiry are not distinguished, but are equated within the objectives.

These early curricula that aimed to engage students in inquiry-oriented science laboratory activities in an effort to increase students’ understanding of the nature and processes of science were designed and implemented under the aforementioned assumption that by “doing science” participants would come to understand NOS. It should be noted that this was not the only objective of these programs. Other objectives included better understanding of science content and processes, greater interest and attitude toward science, and greater ability to evaluate
and make decisions scientifically (Sund & Trowbridge, 1973). However, evidence to support their effectiveness in any of these areas was lacking. Many research studies following implementation of these curricula compared the effectiveness of the new inquiry-oriented approaches to the previous, more traditional teaching methods in achieving the various objectives. Included in this review is an examination of those studies that evaluated program effectiveness in fostering better understandings of NOS. As noted, these curricula included understanding of science processes (inquiry) in their concept of NOS.

Trent (1965) compared the relative effectiveness of a traditional high school physics curriculum and the inquiry-based PSSC physics curriculum in developing a “greater understanding of the nature of science” (p. 224) in secondary students. In the PSSC physics course, the textbook and associated materials were prepared by the Physical Science Study Committee. The PSSC curriculum is a laboratory-centered, experimental approach with objectives toward developing appreciation of science and scientists and advancing students’ understanding of the structure and NOS. The traditional, textbook-centered, high school physics curriculum used in this study met four criteria: (1) physics course taught in the junior or senior year of secondary school, (2) a physics course that is a college prep science course, (3) a physics course that included units on heat and sound, and (4) a physics course that did not use the PSSC textbook or materials. However, both courses “claim that a major outcome of physics courses is to ‘develop an understanding of science’” (p. 224), where “understanding of science” referred to the “development of science and the scientific enterprise, the structure and methods of science, and science as a produce of human
intelligence” (p. 224). No further description or criteria of the PSSC or traditional approaches were provided.

Fifty-two California secondary schools were chosen for this study, 26 schools that used the PSSC curriculum, and 26 schools that offered the traditional physics course. Questionnaires were sent to 80, out of a possible 136, randomly selected California secondary schools that were teaching the PSSC course during the 1962-63 school year. Questionnaires were also sent to 80 randomly selected California secondary schools that were teaching the traditional physics course. Based on an undetermined rate of questionnaire return, 26 schools from each group were used in the study. The author stated that the control group was the randomly selected 26 groups where the “traditional” high school physics was taught. The experimental group comprised the 26 randomly selected high schools where the PSSC physics course was taught. However, because of the dependence of the return of questionnaires, the sample was one of volunteers, rather than a random selection. There was no discussion of the nature of the resultant sample, or relation of the sample’s “representativeness” to the general population of high school students. As such, generalization of results should not extend beyond the sample of California schools involved in this study. The precise nature of the questionnaires was not explicated in this report. However, Trent (1965) reported that 22 of the 26 teachers from each group indicated that “understanding science” was an important objective in the classroom. The focus of this investigation was on the success in attaining this objective. No explanation or discussion of the 22 teachers’ views of what it means to “understand science” was provided.
The experimental groups were the classrooms that employed the PSSC physics curriculum. The control groups were those who used the "traditional" approach. The present report indicated the investigator relied strictly on the questionnaires to verify the treatments. No classroom observations or analyses were done to determine the consistency of the instruction within treatment groups, or the extent to which they matched the criteria for the "traditional" and PSSC approaches to teaching high school physics.

The measurement instruments in this study were the Test on Understanding Science (TOUS) and the Otis Quick Scoring Mental Ability Test. The TOUS was intended to measure students' "understanding of science" as described by Trent (1965). No source, validity or reliability information on the TOUS was reported in this study. The author offered no description, source, validity, or reliability information regarding the Otis Quick Scoring Mental Ability Test.

Thirteen of the 26 schools in both the PSSC and traditional groups were given the TOUS as a pretest during the first two weeks of the school year. All students were given the TOUS as a posttest during the last three weeks of the school year. Only half the students were administered the pretest so that any testing effect due to taking the pretest could be identified. Students in both groups also took the Otis Quick Scoring Mental Ability Test during the first two weeks of the school year. Some students failed to take all the tests for various reasons, including dropping out of school or being absent on the test dates. Scores for only those students who took all required tests were used for analysis. No report of the extent of sample mortality was provided. This was a possible source of an internal validity problem given that
no information on the original and resulting sample was provided. The author addressed this possibility in the statistical analysis. No information was provided concerning the administration of the tests.

For all statistical measures, the school mean, not individual student data, was appropriately used as the unit of analysis. Four statistical analyses were performed on the collected data. First, an analysis of variance was used to determine the effects of experimental treatment (PSSC course), the effects of pretesting, and the interaction effects between testing and experimental treatment. Second, to assess the effect of sample deterioration, an analysis of variance was conducted using all student data, including those who had not taken all required tests. Third, analysis of covariance was conducted to assess the effect of scholastic aptitude and prior achievement in science on students’ understanding of science as measured by the TOUS, using scores on the *Otis Quick Scoring Mental Ability* and TOUS pretests as covariates in this analysis. Finally, a 2 x 2 factorial analysis of covariance was conducted using school size and type of physics course (PSSC or traditional) as variables, and mental ability and prior science understanding as covariates. No criteria were provided pertaining to school size categories. These statistical measures were appropriately conducted to assess differences between the experimental and control groups.

No differences were found in mean posttest TOUS scores between the pretested and unpretested groups ($\alpha = 0.20$). Likewise, interaction between testing and treatment showed no significant difference. The PSSC schools did, however, demonstrate significantly higher achievement on the TOUS than the traditional
schools (F = 9.78; p = .004). Consistent results were obtained when scores from all
students were included, indicating experimental subject mortality did not influence
the results of this investigation.

When effects due to prior science understanding and mental ability were
appropriately controlled (ANCOVA), no significant difference in the mean TOUS
scores between the PSSC schools and traditional schools was indicated (p > 0.20). In
addition, no differences were indicated due to type of course, school size, or
interaction between type of course and school size (p > 0.20).

The only inference, the author argued, permitted from these results was that
when prior science understanding and mental ability were statistically controlled,
there was no significant difference in students’ understanding of science between
those students in the PSSC course and those in the traditional physics course. This
study implied that “the PSSC and traditional physics courses were equally effective
in attaining student science understanding as measured by the Test on Understanding
Science” (Trent, 1965, p. 229). There was no discussion of whether the achieved
scores on the TOUS represented an adequate understanding of science. Hence, the
attainment of the “understanding of science” objective was not directly addressed in
this report. According to the presented mean TOUS scores, the PSSC and traditional
groups achieved approximately 38 out of the possible 60 points on the TOUS. The
adequacy of this achievement level on understanding of science must be based on a
preset criterion. Trent did not provide any such criterion. Nevertheless, the results of
the analysis indicate the inquiry-based PSSC approach did not enhance students’
understandings of science compared to students in the “traditional” classroom.
Trent (1965) noted differences among the individual schools' mean TOUS scores, suggesting other factors may influence students' understanding of science. Among such factors was perhaps the teachers' behavior and emphases in the classroom. Trent recognized the need for further studies to assess the relationship between teaching style and student achievement in physics. Teachers in the present study indicated an "understanding of science" was an important objective of science teaching. However, no studies had yet examined the relationship between teacher objectives and classroom practice. Furthermore, Trent recommended investigations into the relationships between teacher understanding of science, teacher classroom behavior, and student understanding of science. Along the same line, Trent suggested an examination of teachers' understanding of science, classroom objectives, and success of attaining said objectives is worth pursuing.

The results of Trent's study indicated that, in such research endeavors, prior science understanding is an important factor that should be statistically controlled. Furthermore, this study exemplified the need for classroom observations to assess variations in instruction. The claims and limitations stated in this present study, namely those students in the PSSC group and traditional group scored equally on the TOUS, were valid and consistent with the research results. The importance of classroom observations, assessment of prior science understanding, and evaluation of the teacher as a variable were identified by Trent and is also supported by contemporary views of science education research.

Trent's study spanned a short time period, suggesting perhaps a longer exposure to the PSSC approach may be necessary for enhanced understanding of
science. In response, Crumb (1965) examined developments in student understanding over the course of a school year, or two semesters. In this investigation Crumb (1965) compared the PSSC curriculum to traditional high school physics with respect to students' understandings of science as measured by the Test on Understanding Science (TOUS). The purpose of the study was to determine if there was a significant difference in understanding of science between students who had studied PSSC physics and those who had studied traditional high school physics, with some attention given to gains over time of study. The report included no description of specific objectives of the traditional curriculum.

The sample comprised 1,275 physics students nonrandomly selected from 29 rural and urban high schools in four Midwest states. The sample was stratified according to type of physics course being taught, PSSC or traditional, and whether or not the teacher of the class had participated in a special program on the teaching of PSSC physics. Thus, there were four groups for this study, 360 PP students (PSSC course with PSSC-trained teachers), 78 PT students (PSSC course with non-PSSC-trained teachers), 336 TP students (traditional physics with PSSC-trained teachers) and 501 TT students (traditional physics with non-PSSC-trained teachers). The author stated that classroom observations and teacher interviews were conducted to verify classification. No further elaboration of these verification procedures or course instruction was reported.

The criterion instrument utilized in this study was the TOUS. No source, validity, or reliability details are provided in this report. The TOUS was administered by the investigator at the beginning of the school year and at the end of each of two
semesters. General ability was measured by the *Otis Quick Scoring Tests of Mental Ability*, Gamma, Form Em, and the participants' backgrounds in science were measured by the *Iowa Tests of Educational Development* (ITED), Test Two, Background in Natural Science. Again, the author provided no information in regard to the validity or reliability of these instruments. Furthermore, the time of the single administration of the Otis Test and the ITED was not indicated.

The specific null hypotheses tested were:

1. There is no significant gain in understanding science, as measured by the TOUS, among those students studying high school physics.

2. There is no significant difference in understanding of science, as measured by the TOUS, between those students studying one semester of PSSC physics and those students studying one semester of traditional physics.

3. There is no significant difference in understanding of science, as measured by the TOUS, between those students studying two semesters of PSSC physics and those students studying two semesters of traditional physics.

In the statistical analysis of the scores, means for the groups (PP, PT, TP, TT) were used as the unit of analysis, with group means determined based on individual students' scores.

In the test of the first null hypothesis, t-tests were performed for subgroups of pre-posttest scores on the TOUS for each group and pre-mid for each group. The total means scores for the pre and post-test were also compared as well as the total
means for the pre and mid-test scores. In this analysis, a total of 10 t-tests were performed. To reduce the occurrence of a Type I error due to multiple comparisons, analysis of covariance should have been performed on the mean TOUS scores for the groups and total, with the pre-test scores as the covariate. Differences in the mean pre-test and post-test TOUS scores for each of the four groups as well as the total sample were significant (p<0.01). Thus, it was concluded that students in all classes made significant gains in understanding science during the two semesters of study. Absent from this discussion was the practical significance of the gain in scores. The range in gain in mean scores was 2.2 to 2.8 points on the TOUS with averages ranging from 35.6 to 40.6 points. Thus an increase in 2 or 3 points (about 6% of the average score) has limited practical significance.

The t-tests to detect differences between the pre-test and the mid-test indicated the mean scores were significantly different for most groups and the total. Although there was an increase, the PT group was the only group not showing significant gains in mean TOUS score between the pre-test and the mid-test. Thus, it was concluded that significant gains in understanding of science could be achieved during one semester of study in most cases. Again, the practical significance of the gains, which were about half of the pre to post gains (approximately 3%), was not explored.

An analysis of the means of the ITED, the Otis Test, and the TOUS pre-test indicated no differences in the PP and PT groups. Likewise, there were no differences in these means between the TP and TT groups. No details of this analysis were provided. The PP and PT groups were subsequently treated as one
subpopulation, and the TP and TT groups were also combined and treated as one
subpopulation. Lacking from this analysis was the consideration of the variable
effects of individual teachers within the groups. Other studies indicated the
importance of the teacher as a variable in measuring the effects of curricula (Trent,
1965).

Analysis of covariance was appropriately used to compare mean TOUS
scores for the PSSC group and the traditional group. Dependent variables in the
analysis were students' scholastic aptitude, as measured by the Otis Test, and
students' prior science backgrounds, as measured by the ITED, and prior
understanding of science, as measured by the TOUS pre-test. A significant difference
in mean mid-test TOUS scores was determined between the PSSC and traditional
groups when the above factors were controlled (p<0.01). Moreover, a significant
difference was determined between the two groups with respect to mean post-test
TOUS scores (p<0.01). Thus, the author rejected the second and third null
hypotheses. With such a large sample size, 1,275 students, small differences may be
statistically significant, but the practical significance needs to be addressed. Again,
the practical significance of the apparent score differences in this analysis is minimal
(approximately 3 points) and ignored by the author.

Crumb (1965) concluded that this study provided evidence for differential
gains in understanding of NOS and its process and methods as measured by the
TOUS, with the PSSC physics students having greater gains than those students in
the traditional high school physics courses. In addition, Crumb indicated the greater
success of the PSSC program over a short period of one semester. Generalization of
the results to other similar student populations is suggested due to the sampling of schools by size, type, and location.

Crumb (1965) noted that the TOUS pre-test scores for students in the PSSC classes were higher than those scores for students in the traditional classes. Crumb suggested the school, the staff, and their attitudes toward change may be important considerations for further study.

Like Trent, Crumb failed to consider the overall effectiveness of either curriculum in advancing students' conceptions of NOS. Adjusted post-test TOUS scores for the PSSC groups and the traditional groups was 38.31 and 37.62, respectively. With 60 points possible, these scores are approximately 63%. Crumb does not address the adequacy of such scores for understanding of NOS. The quantitative results do not provide ample description of the respondents' conceptions of NOS. It would appear that neither curriculum is particularly effective in promoting high levels of understanding, but the reader is uninformed of what conceptions of NOS or inquiry the respondents purported.

The Biological Sciences Curriculum Study, BCSC, was another curriculum designed with the vision of incorporating more inquiry experiences in school science. Jungwirth (1970) attempted to evaluate the effectiveness of the BSCS-Biology Yellow Version on Israeli secondary students' understanding of NOS. The BSCS-Biology program in Israel spans three years, from grade 9 through grade 11. Jungwirth, like Trent (1965), argued that the new curricula had been implemented without testing whether the objectives of the curricula were actually achievable, and to what extent various student populations were obtaining them. The purpose of the
investigation by Jungwirth was to assess the effectiveness of the BSCS-Biology curriculum towards its intended objectives for Israeli high school biology students.

Using the work of Schwab as a framework, Jungwirth outlined the objectives for the BSCS-Biology Yellow Version as defined by a previously published Teachers' Handbook. According to Jungwirth (1970), the Handbook describes the main objective of the BSCS curriculum to be for students to gain “a genuine understanding of the nature of enquiry” (p. 141). The skills and abilities necessary for such achievement include:

1. Ability to discern problems.
2. Ability to formulate and screen useful hypotheses.
3. Ability to infer what data to seek (if...then logic).
4. Ability to plan experiments appropriate to a problem.
5. Ability to interpret data (draw appropriate conclusions).
6. Ability to identify assumptions. (p. 141)

The Handbook subdivides these objectives into an analytical mode and a constructive mode. The analytical mode refers to the abilities to understand and make judgments on scientific information. The constructive mode refers to the skills required to properly engage in scientific inquiry.

These educational objectives stated for the BSCS-Biology Yellow Version were used by Jungwirth (1970) in the development of an instrument to assess students' achievement in both the analytical and constructive modes. In the development of his instrument, Jungwirth utilized the TOUS for validation purposes.
Based on this approach, Jungwirth considered "the nature of inquiry" and "NOS" synonymous concepts.

Part I of the test aimed to evaluate the analytical aspects of the nature of scientific inquiry. The six sub-areas included for testing were (1) Reading comprehension of data, (2) Manipulation of data, (3) Experimental procedures, including judgment of adequacy of samples and controls, (4) Relevancy of assumptions underlying an hypothesis and/or experiment, (5) Relevancy of hypothesis to a given experiment, and (6) Relevancy of conclusions. For Part I of the instrument, five multiple-choice items represented each sub-area. Information regarding the precise nature of the multiple-choice items was not provided.

Part II of the instrument aimed to evaluate the constructive aspect of the nature of scientific inquiry, namely the skills or abilities inherent to the processes. This part of the test was developed by the author to be a presentation of a problem that the test respondent was to identify, evaluate, and suggest a means to solving the problem. Specifically, the constructive mode of the instrument included

1. A one-page printed text on a biological topic not covered by the BSCS Yellow Version (absorption of minerals by plant-leaves).
2. A problem situation (given problem) based on the text and requiring the following pupil-responses:
   - Formulation of (a) relevant assumption (s).
   - Formulation of a relevant (working)-hypothesis.
Suggestions for the testing of the hypothesis, to include specifications as to sample, proper and adequate controls, experimental conditions and replications. (Jungwirth, 1970, p. 142)

The author-developed instrument was criterion-referenced. Jungwirth (1970) reported that face validity was established by a group of BSCS Adaptation-team leaders. The author reported results of point-biserial correlations as measures of validity. When point-biserial correlations between items and mean sub-test totals ($r = .53$) and between items and mean grand total ($r = .36$) were considered along with the determination that all items had a higher correlation with sub-test total, Jungwirth concluded that the items were properly classified. This exercise was an effort to establish validity of the sub-test items by showing the items measure the same constructs. This is also a measure of internal consistency. The correlation values do not indicate particularly high internal consistency of the instrument.

The author reported an adequate KR$_{20}$ reliability measure of 0.86 for Part I of the instrument, and a rather low Cronbach's alpha coefficient of 0.69 for Part II. A product-moment correlation (Pearson's $r$) was determined using both the Test on Understanding Science and the Milta group-intelligence test against both parts of the developed instrument. The reported correlations ranged from 0.33 to 0.58. Jungwirth (1970) concluded that the developed instrument was appropriate for its intended measurement of students' understanding of the nature of scientific inquiry. It is assumed that the correlation of the developed instrument with the TOUS was an attempt to establish concurrent validity, although this was not explicitly stated in the report. There was no description of the sample used to make external correlations.
between the TOUS scores and the author-developed instrument. In any case, correlation values obtained were not sufficient to conclude the instruments measure the same construct. There was no explanation of the correlation of I.Q. measures to the developed instrument.

The sample for this investigation comprised 693 Israeli tenth-grade students from 32 schools. Of the 25 schools in the study who adopted the BSCS- Biology curriculum, 10 were selective urban schools, nine were non-selective rural schools, and six were non-selective agricultural boarding schools. The control, or comparison, group consisted of four selective urban schools and three non-selective agricultural boarding schools. There were no non-selective rural schools in the control group. The author correctly stated that the groups were not random samples because the teachers of the experimental groups were members of the Israel BSCS Adaptation writing-team. No description of selection of comparison schools or the basis for any discrimination in classroom selection was provided. Jungwirth (1970) did report, however, there were differences in mean I.Q. scores between the city schools, rural schools, and agricultural schools, although no specific data were provided. No attempt was made to verify the consistency of curriculum implementation at the different schools.

The test was administered to the grade 10 students in all school groups at the end of the school term. No description of the manner in which the instruments were administered and scored was reported. Because of the subjective nature of Part II of the instrument, scoring criteria should have been clearly explicated in this report. Indeed, the lack of accurate validity measures and details of scoring of the
instrument are critical omissions that lead the reader to hold any results from this study suspect.

Analysis of test scores was based on comparisons among mean test scores for the students in the three school groups using the BSCS curriculum (city, n=323; rural, n=208; and agricultural, n=162) and the two comparison school groups (city, n=138; and agricultural, n=77). The intact classes for each school should have been used as the unit of analysis, rather than each individual student score. However, no statistical analyses were performed for these comparisons. The author only stated relative differences (higher or lower) or provided mean percentage data but did not confirm any differences by statistical measures. As such, and because the reader is not informed of the representativeness of the sample, any conclusions drawn from this study pertain only to the specific sample, and may not apply to the general population. Indeed, caution in accepting any of the reported conclusions is needed.

The results from this investigation were divided into four sections. The first considered student achievement in the analytical mode (Part I of the test). For each school group, the mean scores were determined in each of the six subsets of the analytical mode. The subset of lowest achievement for all groups was subset iii, experimental procedure. Jungwirth (1970) noted differences in achievement between city schools, rural schools, and agricultural schools, with the city schools tending to score higher. He suggested these differences could be explained by differences in average intelligence of the students at the different schools. He reported that the city schools were more homogeneous in I.Q. than the other schools. However, he provided no evidence to support this claim.
Subsets of Part I were compared relative to I.Q. level of the students in the BSCS groups. Level of I.Q. was divided into four categories: below 95, 96-104, 105-113, and 113-137. The author did not make clear how the students’ I.Q. levels were measured or what criteria were used to establish the cut-off values for I.Q. categories. The BSCS students in the lowest I.Q. level also achieved the lowest scores (range from 30% - 47%) on all six subsets of Part I. Students in the highest I.Q. level achieved the highest scores (range from 48% to 78%). Again, no statistical measures were provided to assess the extent of the correlation between measured I.Q. and Part I of the instrument.

Results from Part II, the constructive mode of the test, indicated that almost half of all subpopulations suggested irrelevant experiments (47% in the BSCS group and 48% in the non-BSCS group). The author suggested that students confused the given data with the problem. Similar confusion was attributed for low scores in hypothesis-formulation (correct responses for 54% of BSCS and 39% of non-BSCS) and identification of assumptions (correct responses for 26% of both groups). Again, no statistical analyses were employed to determine the extent to which the BSCS and non-BSCS scores were different. Given that no criteria were explicated for the analysis of Part II of the instrument, no valid implications can be drawn from these stated results.

In the category of making suggestions for experimental design, 60% of the BSCS population and 70% of the non-BSCS population suggested uncontrolled experiments. Of the BSCS students who suggested relevant experiments, 88% indicated a need for control, and 75% indicated a need for adequate samples. Of the
non-BSCS students who suggested relevant experiments, 70% indicated a need for control, and 50% indicated a need for adequate samples. Again, the lack of stated operational definitions for these categories necessarily negates any implications.

Taking I.Q. levels into consideration, the author concluded that students in all I.Q. levels “show unsatisfactory achievement in ‘formulation of assumption’” (Jungwirth, 1970, p. 147). Percent mastery in formulation of assumptions ranged from 17% for the lowest I.Q. group to 30% for the highest I.Q. group. He considered "unsatisfactory" to be less than 65% achievement.

Intercorrelations were determined between sub-areas of Parts I and II. These correlations were based on individual scores of the BSCS population, with 691 degrees of freedom. No further description of this calculation was provided. From the intercorrelations, it was determined that reading comprehension was a contributing factor for success when multiple-choice recognition items were used (total intercorrelation of 0.77) but only a minor factor when original formulations were necessary (total intercorrelation of 0.37). Also, low correlation values were determined between achievement on parallel tasks in Parts I and II. For example, when “identification of assumptions” from Part I is correlated with “formulation of assumptions” from Part II, the resulting values were 0.07, 0.09, and 0.34 for the BSCS city schools, rural schools and agricultural schools, respectively. The author suggested these results indicated the two modes of the test measure two different aspects of science inquiry. Finally, the intercorrelation data indicate, “the ability to formulate a correct working-hypothesis does not ensure correct suggestions of relevant experiments” (Jungwirth, 1970, p. 148).
Although no specific results were presented, the author reported that “no significant differences were found between males and females in any of the subtests of Parts I and II, based on analysis of covariance, with means adjusted for I.Q. This no-difference phenomenon was found in all sub-populations” (p. 148). This was the only mention of exploring gender effects on achievement. The reporting of this finding appeared almost as an afterthought.

Jungwirth (1970) erroneously concluded that attainment of an understanding of scientific processes, as measured by the analytical mode, was possible through use of the BSCS curriculum. He suggested that the low achievement in the areas experimental procedures and conclusions was disturbing. He determined that without considering these two areas for the BSCS population, overall achievement in the analytical mode approaches 65% for 2/3 of the BSCS population, his criterion for successful achievement. Thus, the author was suggesting an understanding of experimental procedures and the formation of conclusions were not necessary requirements for understanding the nature of inquiry. Simply disregarding students’ scores on these two measures elevated their status from “low achievement” to “successful achievement.” Yet these two aspects comprised two of the six established criteria for understanding of inquiry, upon which this author’s instrument was developed. Either the BSCS curriculum did not enhance students’ understanding of science processes, or the analytical mode of the instrument did not accurately assess students’ understandings. There is always the possibility of both.

The author concluded that, “in general, concept-mastery in the constructive mode approached a satisfactory level only with the upper half (by I.Q.) of the BSCS
population as contrasted with 75% in the analytical mode” (Jungwirth, 1970, p. 150).

From this study, he concluded that use of the BSCS-Biology Yellow Version was only mildly more successful in developing an understanding of science processes as compared to the traditional approach. Jungwirth suggested that redirected teacher emphasis could enhance student achievement of an understanding of science processes.

The present report contained numerous errors in design, measurement and analysis. Because of these problems, the results of this study are meaningless in terms of evaluation of the BSCS curriculum on the attainment of an understanding of the nature of scientific inquiry, as suggested by Jungwirth, or NOS as implied by his comparison with the TOUS.

Another study set in Israel, examined the relative effects of three inquiry-based curricula in biology, chemistry and physics on Israeli secondary students’ understandings of the processes of science (Tamir, 1972). The three science curricula in this study were adapted from American curricula for use in Israeli high schools in 1965. In Israel, the three subjects were studied concurrently through grades 9, 10 and 11, whereas in American schools, students tended to take one subject at a time. Tamir suggested that by studying students’ understanding of NOS and inquiry processes when the curricula are taken concurrently he would be able to better determine the relative effectiveness of each curriculum. The biology curriculum was the BSCS Yellow Version; the chemistry curriculum was an adaptation of CHEM-study; and the physics curriculum was based on PSSC. A major difference between the new curricula compared to the old was the increased laboratory focus with the
objective of gaining an understanding of the processes of science through inquiry-oriented activities. Therefore, there was the assumption that “assessment of the understanding of the processes of science will serve as an appropriate index to the degree of change and to the extent to which each of the new curricula has been successful in attaining this important objective” (Tamir, 1972, p. 240). The terminology and approach reported suggests the author equated understanding of “science process” with understanding of NOS.

The sample for this study consisted of students in science classes of 44 Israeli high schools. The author reported the sampling was stratified, but only to the extent that each type of Israeli high school was represented. Whether the sample of high schools was adequately stratified is questionable. The sample included 19 city academic schools, 17 agricultural cooperative settlement schools, 4 agricultural boarding schools, and 4 occupational technical schools. The schools within each of these categories were reported to be randomly selected. From each school, one class from each of grades 9, 10, 11, and 12 was used in the study. The total number of classes was 220, with 3500 students involved. For the analyses, the class was used as the appropriate unit.

The *Welch Science Process Inventory* (SPI) (Welch & Pella, 1967-68) was sent to each of the sample schools at the end of the academic year. Tamir reported choosing the SPI over the TOUS for several reasons, including potential validity problems with the TOUS, some students already taken the TOUS, and format preference. The SPI contains 135 agree/disagree statements regarding the assumptions, activities, products and ethics of science. It had previously been
reported to have a K-R$_{20}$ reliability of 0.86 on a sample of 171 high school physics students. Cronbach's alpha coefficient was determined here to be 0.74. Evidence for face validity of the scoring key was based on 19 practicing research scientists and had been reported elsewhere. Tamir (1972) stated that the SPI was an appropriate instrument for assessing all areas of high school science. The instrument was translated from English into Hebrew. Validity of the Hebrew version of the SPI should have been established with a sample of high school Israeli students, representative of the sample in the present study. No such validity measures are reported.

Students in grades 10, 11, and 12 took the SPI in May following the year during which they studied the BSCS Yellow Version, the CHEM-study chemistry, and the PSSC physics, concurrently. The grade 9 students were given the SPI as a pretest during the first week of classes, and again as a posttest. Tamir (1972) reported that the schools were sent the instrument. No information concerning the administration of the SPI was provided. Inconsistencies in test administration would add unknown and unaccountable variation to the study. Furthermore, no classroom observations were conducted to verify the integrity of the curricular implementation.

An item analysis was reported, using scores from a random sample of 140 students. The percentage of accepted answers for each of the 135 SPI statements was reported. On the average 50% of the answers were correct for 87% of the items, and 70% were correct for 63% of the items. The author identified five areas of consistent misconceptions. These were (1) interrelationships among hypotheses, assumptions and theories; (2) the role of the scientists in creating models and classification
schemes; (3) the plurality of scientific methods and approaches; (4) the nature of experiment and control; and (5) the relationship between scientific research and practical application.

The mean scores for the different grade levels were determined. These scores demonstrated a gradual increase in SPI scores with grade level. Norms for Israeli 9th and 12th grade students were established based on these mean scores. However, it was aptly noted that these data result from this cross-sectional study, rather than a longitudinal study, and were, therefore, only representative of the situation in Israeli high schools in 1969.

An objective of this study was to determine the relative effects of the curricula on students' understandings of the processes of science. The mean yearly gain was determined by the difference in mean SPI scores for the students beginning grade 9 and the mean SPI scores for the students completing grade 12, and dividing by 4. Although not mentioned, weighted means should have been used in the calculations to account for differences in group sizes. The mean yearly gain in SPI from grade 9 to grade 12 was used as the dependent variable for assessing the differential effects of the three new curricula because this gain in score was determined not to be correlated with type of school ($r = -0.08$). Means for school type were not reported. Because this was a cross-sectional study, the students in each grade were not the same. As later discussed, the mean gains varied with grade level. Nonetheless, Tamir (1972) used the mean yearly gain score in the analysis of the curricula. Achievement scores and school type did correlate ($r = 0.37$) with city academic school showing higher achievement. This $r^2$ value indicates only 13.7% of
the scores can be attributed to school type. These results more appropriately indicate no differences in achievement between students in the different types of schools.

The relative effects of the three curriculum projects were measured in several ways. First, an adoption index was determined for each science area based on the number of grades studying each of the curricula (BSCS, CHEM-study, and PSSC). For example, a school in which PSSC physics was studied in grades 9 and 10 has an adoption index of 2. The index for each curriculum was determined for each type of school (city academic, cooperative settlement, agricultural, and occupational). For the city academic schools, the cooperative settlement, and the agricultural schools, the BSCS curriculum had the highest adoption indices (68.2, 68.3, and 60.0, respectively). The occupational schools had not adopted any of the new curricula (adoption index of 0). It is assumed that the occupational schools employed the more "traditional" instruction of the subjects. The cooperative settlement school had equal use of BSCS and PSSC-based programs. The city academic school had an adoption index of only 4.5 for the PSSC-based curriculum, whereas the agricultural schools had not adopted the PSSC-based curriculum. Next, the relationship between the adoption index and mean yearly gain in SPI was determined using a Pearson r correlation. As expected, the BSCS had the highest positive correlation with mean yearly gain in SPI scores (r = 0.25). However, the author noted, this correlation was quite low and limited the implications of the study. Indeed, only 6.25% of the mean yearly gain in SPI score could potentially be attributed to the adoption of the BSCS curriculum. Use of the relative adoption indices of the various curricula to explain gains in SPI scores implied that adoption alone assured consistent and proper
implementation of the curricula. No assessment of instructional approaches was included in this study to verify this assumption.

Tamir (1972) further probed the relative impact of the three curricula on the mean yearly gain by employing a multiple regression analysis. The independent variables in this analysis were BSCS biology, CHEM-study chemistry, PSSC physics, and the four types of schools. The regression equation was $G = 2.65 + 0.94B - 0.85s + e$, where $G =$ mean yearly gain, $B =$ BSCS biology, $s =$ occupational school, and $e =$ "error". The BSCS variable and the occupational school variable were found to be the only significant factors (t values = 1.81 and 1.54, respective; $p = 0.05$). By eliminating the impact of the BSCS curriculum on the effects seen for the other two curricula, the mean yearly gain dropped from 14.36 points to 10.60 points. Thus, the author stated that the BSCS courses contribute 31% increase in SPI scores yearly. This was determined to be significant ($p < 0.05$).

Tamir (1972) suggested that the BSCS biology curriculum had not only the highest effect, but the only significant effect, on students' understanding of the processes of science. From the presented analysis, however, the practical significance of the impact of the BSCS context on students' conceptions is questionable, especially to the American educational system. He related the success of the BSCS curriculum to the unique characteristics of biological sciences, implying the contexts of physics and chemistry were less likely to enhance learners' views of science. Moreover, the BSCS curriculum in Israeli schools was the most familiar of the three due to the more rapid adoption of the BSCS project compared to the chemistry or physics curricula. Finally, in-service teacher training and teaching
materials had been developed by the BSCS project. Because no observations were made to verify the equal implementation of BSCS, CHEM-study and PSSC, the results may be biased in favor of the more familiar BSCS curriculum. The purported differential effects of the biology context versus physics and chemistry contexts is interesting, but in need of verification of equal exposure and engagement in inquiry-based activity within the different disciplinary contexts. Furthermore, the reader has no indication of how the different disciplinary contexts may have impacted conceptions. The SPI scores do not offer a description of conceptions that enable qualitative comparisons.

The cooperative settlement schools had the highest use of all three curricula, and the occupational schools did not adopt any of the new curricula. Because the BSCS had the greatest effect on performance on the SPI, Tamir (1972) concluded that students in occupational schools were at a disadvantage compared with students in the other types of schools regarding their understanding of the processes of science. The fairly low correlation of achievement with school type argues against this conclusion. A significant difference between achievement scores on the SPI for the various school types was not indicated.

Consistent with past efforts of the 1960s, the more recent reform efforts in science education continue to stress the need for improved instruction toward meeting the goal of improving conceptions of NOS. Additional curricula have amassed in response to these recommendations. The purpose of a study conducted by Meichtry (1992) was to examine the influence of one of these more recently developed curricula, the BSCS middle school science program Science and
Technology: Investigating Human Dimensions. Meichtry detailed the arguments that led to the development of science instructional materials intended to provide teachers with activities and approaches that enhance learning about NOS. The recommended strategies emphasize inquiry-oriented science activities that involve students in all stages of scientific investigations such as questioning, designing, conducting, and analyzing evidence. Meichtry recognized the lack of empirical evidence to support the use of the most recently developed materials to develop more adequate understandings of NOS. Such was the aim of this present study.

Two of the assumptions upon which the Science and Technology: Investigating Human Dimensions program was based included the importance of scientific literacy and the need for understanding NOS. This program is a 3-year integrated science curriculum that emphasized questions of interest to middle school students, inquiry activities to promote higher order thinking skills such as reasoning and decision-making, and the use of cooperative learning. Activities were modeled after a 5-stage learning model wherein students are first engaged, then actively explore their ideas, explain and elaborate on their experiences and understandings, and finally evaluate their progress.

Specifically, the purpose of this study was to determine the extent to which the BSCS middle school science program Science and Technology: Investigating Human Dimension influenced student understanding of NOS. Meichtry (1992) looked for differences in students’ understandings of NOS before and after experiencing the BSCS program. She looked for differences in students’ understandings of NOS before and after experiencing the usual science program
(non-BSCS) employed by the school. Finally she looked for differences in understandings of NOS between students in the BSCS program and students in the non-BSCS program.

The aspects of NOS used as the dependent variables in this study were: (1) science is a creative endeavor; (2) science is developmental, or tentative; (3) science is testable, or scientific knowledge is capable of empirical test; and (4) science is a unified set of properties, or an interrelated network of theories, laws, and concepts. The inclusion of these four dimensions of NOS, Meichtry (1992) states, is based on what is reported in the literature as representative of NOS.

The student participants in this study were 6th, 7th, and 8th grade students who attended schools in the same midwestern suburban school district. Eleven science teachers in one of the three middle schools in the district taught the BSCS curriculum to the 1004 student participants in that school. The control group comprised 603 students from one of the remaining middle schools. The number of teachers from the control group was not reported. The two middle schools from this district were chosen based on their similarities in socioeconomic status and ethnic origin of the student body. No information regarding the teachers was provided.

The control group experienced science as it had been taught for 10 years prior to the study. Students in the different grade levels used different text programs, and high and lower-ability students in the 7th and 8th grades also experienced different programs. None of the science programs experienced by the control group participants were BSCS-based. The author did not explicate goals or specific emphases of the non-BSCS curricula. However, classroom observations, discussed
below, enabled comparisons of the programs in terms of attention to NOS and student activities.

A modified version of the Nature of Scientific Knowledge Scale (MNSKS) was used to measure student understandings of the four targeted aspects of NOS. The instrument was modified to include 32 items, 5-point Likert scale response, specific for measuring the creative, developmental, testable, and unified dimensions of NOS, as well as a measure of an overall understanding of NOS. There were eight statements for each of the four subscales, allowing subscale scores to range from 8 to 40 points. Thus, the range for the overall test was 32 to 160 points. A neutral score on the test was 96 points overall, or 24 points for a subscale. Scores higher than neutral were considered to indicate views toward the accepted understandings of NOS. Lower scores indicated views inconsistent with currently accepted understandings.

Four science educators established content validity of the MNSKS. Construct validity was determined statistically by factor analysis of the pretest results. The appropriateness of the reading level of the instrument was determined by interviews with middle school students. Reliability measures for each of the subscales were determined from the pretest scores and ranged from $\alpha=0.45$ to $\alpha=0.60$. The overall test had a higher reliability coefficient of 0.77. Meichtry (1992) recognized the relatively low reliability measure of the MNSKS compared to other reported measures. She also reported that this finding is consistent with the literature that shows the decrease in reliability of the Nature of Scientific Knowledge Scale, the
instrument that was modified for this study, as the age of the test population decreases. This is an appropriately recognized limitation of Meichtry’s study.

In this nonequivalent control group design, both the experimental and control groups were administered the MNSKS as a pretest and posttest to measure their understanding of NOS. The pretest was administered to both groups at the beginning of the school year. The same test was again administered after 26 weeks of instruction. No information regarding test administration was provided. Other data sources included classroom observations, formal interviews, and examination of curricular materials.

The purpose of the classroom observations was to collect information on the science content, the instructional approaches, and details of attention to NOS during instruction. There was no further information reported about the number of observations conducted or if all teacher participants were observed. Moreover, the author did not describe the types of events that constituted “attention to NOS.” Formal interviews were conducted with each of the observed teachers, the school principals, the district science supervisor, and the research assistant for BSCS. The purpose was to collect information about the science content, teaching techniques, and the extent to which the teachers in the experimental group were implementing the BSCS program. The curricular materials of both groups were examined for program goals, disciplinary orientation, science content, instructional activities and methodology, grouping strategies, and emphasis on NOS as defined by this study.

Multivariate analysis of variance and analysis of covariance were used in the analysis of the pretest and posttest scores. To aid in the interpretation of the
statistical findings, qualitative analysis techniques were employed for the classroom observations, interviews, and curricular materials. Scores for students in each group (experimental N = 809, control N= 491) were averaged and the means, adjusted means and standard deviations reported. The reason for the decrease in number of student participants for each group is not explained in the report. The reported findings for both the experimental and control group were a decrease in mean overall score between the pretest and posttest. In addition, the pretest and posttest scores revealed neither group held conceptions of NOS that would be considered “adequate” as measured by the MNSKS (posttest mean overall score for the experimental group = 80.983; posttest mean overall score for the control group = 82.389 out of a possible 160). Furthermore, neither group demonstrated adequate understandings for any of the subscales for either administration of the test.

For both groups, a repeated-measures multivariate analysis of variance was performed to determine any differences in understanding of NOS between the pretest and posttest measurements for each of the subscales and the overall test. The individual student was inappropriately used as the unit of analysis. The more appropriate approach would have used the class as the unit of analysis. The results indicated a significant difference in the overall test scores for both the experimental and control groups (p = 0.006 and 0.001, respectively). For the experimental group, paired t-tests (α = 0.01) determined students’ understanding of the developmental and testable nature of science decreased significantly (two-tailed, p = 0.001 and 0.000, respectively). For the control group, scores on the subscale for the creative nature of science decreased significantly (two-tailed, p=0.003). It should be noted
that the largest difference in mean scores for these subscales was 0.7 points. Although significantly different, a drop of 0.7 points out of 40 possible points does not hold any practical importance toward students’ understanding of NOS.

Analysis of covariance (α = 0.01) was employed to determine differences between the experimental and control groups’ mean posttest scores for each of the four subscales. The mean pretest scores served as the covariates. The results determined students in the control group to have a significantly higher mean score on the subscale for the testable nature of science than students in the experimental group (F=12.84, p = 0.000). The actual difference in adjusted mean posttest scores on this subscale was 0.5 points. Again, this minimal difference, although statistically significant, imparts little practical importance.

Analyses of classroom observations, interviews, and curricular materials were based on the recommended instructional practices to enhance student learning of NOS. The BSCS program was found to be highly activity-based, 85% compared to 40% for the control group program; more student centered; 85% cooperative learning environment, compared to no cooperative learning in the control; and more consistent with the 5-step learning model. In addition, the BSCS program was found to be highly inquiry-based compared to the emphasis on knowledge verification found for the control program. Regarding attention to NOS, the BSCS instruction offered explicit instruction toward the creative, developmental, and unified nature of science, and implicit reference to the testable nature of science. The control group instruction made implicit reference to the developmental nature of science, and no attention to the creative, testable, or unified nature of science. The investigator
determined the teachers of the BSCS program were fairly consistent in implementing the curriculum as designed. The components least represented were the use of cooperative learning and the evaluation of student progress.

Lacking from this description of student activities and instructional practices are the criteria used to categorize the activities. For example, Meichtry (1992) did not provide the criteria used to categorize an activity as "highly inquiry-based." There was no information concerning the depth of inquiry experienced by the students in these programs. Regarding the meaning of "explicit" and "implicit" attention to NOS, she provided sample passages from the BSCS textbook and a non-BSCS textbook in order to compare explicit and non-explicit references to the testable NOS. Although informative to the reader, the passages did not necessarily reflect the observed implicit or explicit classroom instruction. Such details of the analyses would be beneficial in the interpretation of the reported results.

The results of the analyses indicated that the BSCS program, although more representative of the type of instruction believed to be conducive to teaching NOS, was not more successful in enhancing middle school students' conceptions than the control-group program. In fact, for the BSCS group, student understanding decreased on two of the four subscales and was significantly less than the students in the control group on one of the subscales. Meichtry (1992) suggested several possible reasons for these results and limitations of the study. First, the lack of increase in student conceptions of NOS may have resulted from the lack of consistency to which NOS was represented by the BSCS curriculum. She reported that the creative, developmental, testable, and unified nature of science were not always directly
presented to the students by either the curricular materials or the teachers. She reported that even though students were engaged in inquiry-based investigations requiring creativity, the connection to science as a creative process in the development of scientific knowledge was never directly made. She argued for the need for *explicit* attention to these concepts for students to gain understanding. Just doing the activities or learning the biology content was not sufficient for students to make connections to the four aspects of NOS measured in this study.

Meichtry (1992) stated several other possible reasons for the outcomes of the present study. The age of the student participants could influence the validity of the measurement instrument, as previously described. In addition, the cognitive developmental characteristics of the early adolescents could contribute to varying views of NOS. She recommends conducting student interviews in combination with the paper-pencil test to help determine the actual views of the student participants. Regardless of the developmental stage of the subjects, interviews have been suggested as a means to establish validity of the paper-pencil instrument and expand on the subjects' responses to gain a more detailed description of their views (Lederman & O'Malley, 1990).

A further limitation influencing this study, Meichtry (1992) argued, includes the process of change itself. She cited other reports that state for effective change to be realized in any program, a period of two to three years is required. During the transition from the traditional curriculum to the BSCS curriculum, teachers must use new materials, attempt different teaching approaches, and must restructure their beliefs about their role in teaching. The teachers implementing the new program
have a difficult and complex task. As such, “it may not be realistic to expect a change in student understandings after a period of only 26 weeks” (p. 403). Finally, this study was conducted during the first of a two-year field test of the BSCS curriculum. The program could be considered in the developmental stages and subject to revision during the field-testing. Thus, Meichtry appropriately stated the results of her study could not be generalized beyond the program, as it existed at the time of the study. She recommended further investigations to probe the effects of the stated limitations.

Even with such broad limitations, Meichtry (1992) reported several implications for the teaching of NOS resulting from her study. Meichtry concluded that students’ views of NOS were not enhanced by the BSCS program due to the inadequate representation of the BSCS program toward a constructivist view of learning. Specifically, the program inadequately considered students’ prior knowledge as important and necessary for building new knowledge or changing existing knowledge. Also, she suggested the low pretest scores for the students in this study may have made their conceptual change toward more adequate understanding of NOS attainable through the BSCS program. The suggestion was made that the BSCS program might be more effective if greater emphasis were placed on the identification of students’ misconceptions. Her argument detailed the suggested benefits to student learning that such considerations add, yet is inappropriate as an implication of the present study. The classroom observations and interviews did not focus on the use of students’ prior knowledge during teaching. Further study of the issue is warranted.
The results of the present study, similar to the others reviewed here, illustrate that just because the curriculum states a particular goal and is designed accordingly does not guarantee the intended outcomes. To aid in increasing students' understanding of NOS through the BSCS approach, Meichtry (1992) stated three formal recommendations consistent with the results from this study. First, there must be *explicit* attention to all aspects of NOS by both the curricular materials and the teacher. Second, the content and processes students are learning must be directly related to the various aspects of NOS. Third, instructional methods to facilitate conceptual change must be used.

**Summary.** The five articles included in this section compare the influence on students' views of NOS and inquiry of curricula designed to provide opportunities for students to engage in inquiry-based science laboratory activities to more traditional curricula that relied heavily on textbooks and "cook-book" laboratory demonstrations. This collection of reports indicates the inquiry-oriented curricula alone, as implemented in these studies, do not lead to improvements in student understanding of NOS or inquiry, as measured in these studies.

None of the studies revealed *practical* differences among students studying the BSCS, PSSC, or CHEM-study and those studying the "traditional," noninquiry-oriented, curricula. These results do not support the assumption that the increased emphasis on inquiry and inquiry-based activities would effectively enhancing students' conceptions of NOS or inquiry. Recall that the curricula studied in these reports were designed with the specific intention of improving students' understanding of science, among other aims. One possible explanation for the
ineffectiveness of these inquiry-based curricula relates to the discrepancies between the intended inquiry levels and actual inquiry levels of the designed activities. The intentions of the curricula were to present opportunities for students to practice scientific inquiry, with the assumption that the experiences were similar to the actions of scientists. Only one study reviewed here (Meichtry, 1992) conducted classroom observations in an effort to describe the inquiry nature of the activities. Even though she described them as “highly inquiry-based,” the activities did not lead to better understandings of NOS for those students. However, in that report Meichtry failed to clearly define what she considered “highly inquiry-based.” In fact, the majority of the activities in the BSCS and PSSC curricula have been described as fully guided inquiries complete with stated problems, procedures and expected outcomes (Herron, 1971). As such, the experiences of the students in these studies were most likely far from scientific inquiries as practiced by scientists. The lack of considerable opportunities for students to engage in high-level scientific inquiry (Herron, 1971; Schwab, 1962) may have contributed to the failure of the programs to achieve their stated objective of enhancing students’ understandings of NOS. The assumption that one will develop an understanding of NOS through the act of “doing science,” in the sense of actual participation in scientific inquiry as performed by scientists, remains untested.

An additional assumption common to the reviewed investigations was that the curricula, regardless of the actual nature of the activities, were implemented in a manner consistent with their design. As previously mentioned, only one of the studies (Meichtry, 1992) employed classroom observations to verify the instruction.
All the investigators, except Meichtry, assumed the curricula were so-called "teacher-proof," whereby all teachers would obtain the same results simply by using the developed materials. Yet, no empirical evidence had been collected to verify such an assumption. Indeed, the results of these studies strongly suggest different teachers have different impacts on the effectiveness of a set curriculum. As suggested by Trent (1965), because of the variability in achievement among students at different schools adopting the same curriculum, other factors must influence students' conceptions of NOS. This failure to produce consistent results, positive, negative, or neutral, suggests the teachers play an integral role in students' development of conceptions of NOS. Consequently, the research focus turned toward the teacher and the manners in which the teacher could be most effective in using inquiry-based activities to teach aspects of NOS.

Finally, none of the studies reported here provide sufficient description of respondents' views of NOS or inquiry to allow qualitative comparisons among groups of learners. The test scores are comparable through statistical measures, but they fail to describe how learners in different contexts understand NOS or inquiry. Tamir (1972) suggests the biology context lead to better learning about science and science processes, but the methods of the study did not provide details of differences in views. Furthermore, the students experiencing inquiry within different subject areas may not have had similar levels of experience. The lack of classroom observations to ensure similar time, type, and level of inquiry exposure within the different subject areas is a limitation of Tamir's study.
The Influence of Different Inquiry Teaching Approaches on Students' Conceptions of NOS

This section presents research studies that examined the effectiveness of various approaches to implementing a given curriculum. Two of the reports examine eighth grade students' understanding of science in classroom laboratory situations that are structured differently (Yager, Engen, & Snider, 1969; Yager & Wick, 1966). Both of these reports used the BSCS Blue Version biology curriculum. One report examines the effects of two inquiry approaches on grade 6 students views of NOS (Khishfe & Abd-El-Khalick, 2002). The three remaining reports reviewed in this section deal with developing understandings of NOS aspects through various approaches to laboratory instruction in college level science courses (Haukoos & Penick, 1983, 985; Spears & Zollman, 1977). Attempts were made in these studies to maintain an inquiry focus as well as minimize the “teacher variable.” Variations in treatments included level of engagement of the students in the activities, type of supporting materials and discussion focus, and amount of guidance or structure provided by the teacher.

Yager, Engen, and Snider (1969) studied the relative effects of three approaches to the same biology curriculum on students’ abilities for critical thinking, understanding of NOS, understanding of scientists, attitude toward science, knowledge of general science, knowledge of biology, and ability to use biological tools. The three approaches compared were a laboratory approach, demonstration approach, and an approach that incorporated neither laboratory nor demonstration (discussion only). The BSCS Blue Version Course Biological Science: Molecules to
Man was used as the curriculum program for the three 8th grade classes in this study. This study arose from conflicting results in previous reports, such as those included in the present review, of the effects of inquiry-based laboratory curricula on student achievement of various goals of science instruction, including an understanding of NOS.

A pretest-posttest, 3-treatment design was utilized in this study. There were five general areas of student achievement that were the focus of the comparisons. First, the investigators aimed to identify any differences among the three approaches in students' development of critical thinking skills. For this measurement, the Watson-Glaser Critical Thinking Appraisal was used. Second, students' understandings of NOS and of scientists were measured by the Test On Understanding Science (TOUS). Third, student attitudes toward science, biology in particular, were measured by the Silace Attitude Scale and the Prouse Subject Preference Scale. Fourth, student achievement in science was measured using three different tests. The Read General Science Test was used to measure student knowledge of science. The Nelson Biology Test was used to measure student mastery of general biology. The Biological Science Curriculum Study Comprehensive Final was used to measure general achievement in biology. Finally, to identify differences in students' abilities to use laboratory tools in biology, a laboratory practical examination was developed for this study. All other measurement instruments, except for the laboratory practical exam, had been previously published and described. Also, except for the laboratory practical exam, the seven instruments were
administered as both a pretest and posttest. The laboratory practical was administered as a posttest only.

Sixty 8th grade students from the Iowa Laboratory School comprised the sample for this study. They were randomly assigned to one of three biology sections. Because of scheduling conflicts, the class sizes varied. Nonetheless, the authors reported the treatment groups were equivalent with respect to abilities, skills, interests, and attitudes. The only descriptor provided relative to the students was the mean IQ scores of 118.0, 114.4, and 119.5 for the three groups. Pretest scores on the instruments supported their claim of equivalent groups. There was no report of how representative this sample was to all grade 8 students, however. No generalizations could be made beyond this group of grade 8 students.

There were three teachers involved in this study. They were similar in terms of teaching experience, education, and familiarity with the BSCS curriculum. Previous classroom observations determined the teachers to “practice the use of ‘science as inquiry’ in the classroom” (Yager et al., 1969, p. 79). The authors did not further explicate the meaning of “science as inquiry.” As such, it can only be concluded that the three teachers in this study demonstrated similar practices, but the actual nature of those practices toward “science as inquiry” remains in question.

The three treatment groups experienced the same course in terms of textbook, class period time and length, examinations, and teachers. The three teachers established common lesson plans and instructional materials to ensure each section experienced the same course sequence and material. The design attempted to control the “teacher variable” by having the teachers rotate among the three treatment groups
every 4 weeks. The authors reported attempting to establish all laboratories, regardless of treatment group, as “situations for inquiry” (Yager, et al., 1969, p. 78). The difference among the treatment groups was in the type of exposure to laboratory investigations presented in the BSCS Blue Version text. The laboratory group was the only group to actively engage in laboratory activities (50 of the 57) either individually or in groups of two or three students. However, no details or examples of investigations or student responsibilities were included in the report. So, the level of inquiry and type of investigative approaches cannot be determined. The demonstration group experienced the same laboratory investigations in the form of a teacher or student demonstration. In this group, most students were observers of the investigations. The teacher provided additional data sets to the students for their analysis. In the discussion group, the laboratory investigations were discussed, but never experienced or demonstrated. The discussions in this group focused on interpreting the teacher-provided results, relating those conclusions to other situations, describing the experimental design, and suggesting alternative experiments. There were no reports of classroom observations to verify the instructional approaches or elaborate on the nature of the activities or discussions. As such, the extent to which the laboratory situations were similar and inquiry-based cannot be determined.

Analysis of covariance was appropriately utilized for all measures where both a pretest and posttest were used. For these analyses, the mean pretest scores for each group served as the covariates. The F test was used to measure differences between adjusted posttest mean scores. Analysis of variance was employed to determine
differences in the laboratory practical examination mean scores. The level of
significance for these analyses was 0.05.

In the analyses of covariance calculations, no significant differences between
the corresponding mean scores of the three treatment groups were determined
(p>0.05 for all analyses). Thus, there were no differences determined for measures of
critical thinking, understanding of NOS, attitudes toward science, mastery of content
or biology achievement.

For the achievement of laboratory practical skills, the laboratory group scored
significantly higher than the other two groups in each of the four skill tests. Twelve
t-tests were performed to compare the individual groups in each of the four skill
tests. The laboratory group scored significantly higher (p<0.05) for each comparison
with the demonstration group and three of the four comparisons with the discussion
group. There were 12 t-tests performed in this analysis, yet no adjustment made to
the cumulative effect of the multiple tests. Without such adjustments in alpha level
when performing 12 comparisons, each with a 0.05 level of significance, there is a
46% chance of committing a Type I error. Nonetheless, the trend in scores revealed
the laboratory group tended to score higher than the other groups on the skill
assessments.

The authors claimed the results of this study do not support the assumptions
concerning the values of the laboratory. They suggested these results show that
teachers with sufficient preparation can achieve the same results with or without
laboratory experiences in the classroom. The authors emphasized that this study
differs from other studies that compare a laboratory approach to a lecture approach.
The group that did not experience the laboratories or demonstrations was not a lecture-oriented approach, but these students were taught to be critical of research design, interpret results, and state conclusions similar to those students in the other groups. The authors claimed that they experienced science as inquiry even though they did not experience inquiry through laboratory activities. They asserted that, except for the development of specific laboratory skills, many desirable outcomes could be achieved with or without a hands-on laboratory component. However, for some teachers, the use of a laboratory approach may be easier to achieve the desired outcomes of inquiry instruction. Regardless of the approach, the authors stressed the need for emphasis on the nature of the scientific enterprise in an inquiry-oriented curriculum.

Although the results of the present study do not demonstrate significant differences in understanding of NOS, as measured by the TOUS for students in the three treatment groups, the authors failed to elaborate on an additional finding from this study, mainly the relatively low achievement on the TOUS for all the groups. Regardless of treatment, the students maintain poor understandings of NOS. The adjusted posttest scores for the laboratory, demonstration, and discussion groups on the TOUS were reported as 34.67, 32.26, and 33.84 points, respectively, out of 60 possible points. The aforementioned suggestion by the authors of the need to further emphasize the nature of the scientific enterprise clearly applies to the BSCS curriculum utilized in this study. These students learn little about NOS, regardless of the style of inquiry.
This report also exemplifies the need for classroom observations to clarify the instructional approaches, activities, and discussions. Without such details, the investigator cannot make appropriate claims as to the extent to which NOS and scientific inquiry were actually incorporated into the instruction. The statement that all of the approaches presented “science as inquiry,” as claimed in the present report by Yager et al. (1969), is too vague to be blindly accepted. More information regarding investigative approaches and respective science content is needed. Furthermore, the manner in which NOS and scientific inquiry were presented and explained, with implicit assumptions or explicit detail, is important for identifying factors that influence students’ conceptions of NOS in an inquiry-oriented classroom.

In a study by Yager and Wick (1966), the effects of different teaching emphases on students’ understanding of science as well as ability to do critical thinking were examined. Specifically, they compared three laboratory teaching styles that differed in use of supplementary materials and student ideas. The purpose “was to determine if it is possible to affect a student’s understanding of science and his ability to do critical thinking by altering the emphases of the teacher in the classroom” (p. 16).

The study was conducted at the University of Iowa Laboratory School with 70 students in three grade 8 general biology sections. The students were told they were part of a research study, but were not aware of the research purpose, methods or results. The “Hawthorne effect” was minimized due to frequent participation in such studies by the students in the laboratory school. No further information regarding the sample was provided. Students were randomly assigned to one of three treatment
groups. No information concerning group compositions was provided. It was assumed that the students were typical of the grade 8 students at this school. The degree of representativeness of this sample to other grade 8 students was not described. As such, no generalizations could be made to grade 8 students outside of this particular population.

The general biology course emphasized the molecular level of biology with frequent laboratory experiences. Three teachers were used for each of three sections. The teachers had similar training and philosophy. They were reported as being better than average teachers, but no criteria for this assessment were provided. Each of the teachers averaged five years teaching experience and 50 semester hours of graduate level biology. Furthermore, they all had some experience in the history and philosophy of science. The teachers were aware of the intentions of the study and were involved in its development. It was reported, however, that none of the teachers had any preconceived notions about the outcome of the investigation. This is a questionable assumption considering the teachers' backgrounds and that they were aware of the purpose of the study.

In an effort to reduce teacher variability, the teachers collaborated and worked as a team, periodically teaching units in all three sections. All sections spent the same amount of time on units, performed the same laboratory activities, and took the same exams. The sections differed in the materials used and emphasis placed on opinion, interpretation, and ideas concerning the development of scientific knowledge. The first section, the textbook-laboratory (TL) group, used only the BSCS Blue Version textbook and accompanying laboratory. There were no
supplementary materials used or discussed in this class. The TL group avoided discussions that included opinion, interpretations and reports of new findings. The main emphasis was defined by the authors of the textbook as mastery of basic concepts through an inquiry approach in the laboratory.

The second group, the multi-reference-laboratory (MRL) group, used the same materials as the TL group with additional materials from reference books, textbooks, and articles of original work. Materials were paraphrased and prepared as handouts for the students. There was an avoidance of identifying information with any specific textbook. All laboratory procedures were prepared as handouts. Occasionally, student suggestions were incorporated into the laboratory investigations, making the students more involved in the planning of activities than the students in the TL group. The teachers in the MRL group avoided discussion of any controversy that occurs among scientists, and gave attention only to varying interpretations given by modern writers. There was no discussion of history or development of ideas and theories in science.

The third group, the multi-reference-laboratory and idea (MRLI) group, had the same focus as the MRL group with additional emphasis on how and by whom major scientific knowledge was developed. The teachers made constant reference to major contributors in science as well as the culture at the time of contribution. The same laboratories were performed with emphasis on how the investigation would be viewed at various times throughout history. Additional emphasis was placed on how the major ideas have evolved through history, and how present ideas are likely to change with time. "The teachers in this group utilized the spirit of inquiry in the
laboratory to emphasize it as the technique employed by the major contributors of
the big ideas in biology” (Yager & Wick, 1966, p. 17).

Yager and Wick (1966) do not provide information as to training of teachers
in the three approaches, length of treatment or rotational schedule of the three
teachers. Moreover, there was no indication that observational evidence was
collected to demonstrate the consistency of the teaching approaches with the stated
operational definitions. The three teachers rotated through all three groups. Without
observational analysis of teaching behaviors, student behaviors, and classroom
discourse, consistency of the three teachers to the three approaches cannot be
verified.

Measurements of student outcomes were determined in a pretest/posttest
design. The Test on Understanding Science (TOUS) was administered to measure
students’ understanding of science. The Watson-Glaser Critical Thinking Appraisal
was administered to measure students’ skills in critical thinking. The Nelson Biology
Test was administered to determine mastery of content. Yager and Wick (1966) refer
the reader to other sources of documentation on these instruments. No specific
validity or reliability measures were reported.

Yager and Wick (1966) described six preconditions that need to be satisfied
in order to draw meaningful conclusions from the data in this study. Their intent was
to infer any differential causal relationship between teacher emphasis and
achievement on the instruments. Thus the context variability was “teacher
emphasis.” The first precondition states that there must be random selection of
treatment groups. This precondition was assumed to be met by the random
assignment of students to the three sections. Any variations in motivation or other potentially influential aspects should have been evenly distributed among the three groups. The second precondition states that the pretest scores must be unaffected by the measures. This precondition was met by administering the pretest during the first week of classes. The third precondition states there must be homogeneity of regression of the pretest and posttest scores for all treatment groups. A test for homogeneity of regression ($\alpha = 0.01$) for the three instruments was performed and found to be insignificant ($F_{2,64} > 4.98$) for all three groups. The fourth precondition is the necessity for linear regression in each case. This requirement was satisfactorily determined by an F test ratio of mean square for departure from linearity over mean square within (no values reported). The fifth precondition states that the adjusted scores for each treatment must be normally distributed. The investigators made this assumption. Finally, there must be an equal distribution of variance in each case. A Bartlett’s test for homogeneity of variance showed no significant differences (no values reported). With all the preconditions satisfied, an analysis of covariance was justly performed on the collected data, with the class as the unit of analysis and pretest scores as the covariate, to determine if any differences existed between any of the mean scores on each of the three instruments ($\alpha = 0.05$).

The analysis of the TOUS scores demonstrated a significant difference between treatment groups ($F_{2,67} = 29.3823; p < 0.001$). Multiple $t$ tests comparing the three possible pairings of the TL, the MRL, and the MRLI groups were performed and determined that the students in the MRLI group scored significantly higher ($\alpha = 0.01$) on the TOUS than the MRL group. As a post-hoc procedure for
determining differences between groups, the Scheffé method would have been preferred over the multiple individual $t$ tests. The students in the MRL group scored significantly higher than students in the TL group. A similar result was found concerning the ability to do critical thinking ($F_{2,66} = 23.1793; p<0.001$). The $t$ test comparison of group pairings indicated the MRLI and MRL groups both scored significantly higher than the TL groups on the Watson-Glaser Test of Critical Thinking. No difference was found in mastery of content for the three groups ($F_{2,66} = 0.0015; p>0.05$).

The results of this present study suggested that different teacher emphases influenced these students' understandings of science and development of critical thinking skills in biology. Yager and Wick (1966) concluded that the use of a "multireference approach in the biology classroom caused students to develop more skill in critical thinking than when a single textbook is used with the same laboratory investigation" (p. 20). Secondly, they suggested that the "multireference approach was equally superior to a single textbook-laboratory approach in causing students to understand science to a higher degree" (p. 20). However, the multireference-laboratory approach that explicitly emphasized the history of science and the processes of scientific discovery had the greatest effect on increasing students' understanding of science as measured by the TOUS. The results indicate that the "MRLI method is significantly superior to the other two, and that the MRL is significantly better than the TL method. This, of course, is based on the premise that the TOUS really does measure ability to understand science" (p. 18).
When comparing the actual mean score differences between the groups, the practical significance of these results comes into question. The mean difference in TOUS scores, adjusted for pretest values, was only 4.5 points between the MRLI and the MRL groups, and 7.8 points between the MRLI and TL groups. The TOUS is a 60-item instrument. The largest gain, 13% between the TL and MRLI groups, does indicate moderate practical significance. Even though the MRLI treatment group made the largest gain on the TOUS, the MRLI mean score of 37 (adjusted posttest mean) out of 60 points on the TOUS reveals that none of the teaching approaches was particularly effective in establishing an adequate understanding of science, as measured by the TOUS. In the analysis of the effect of teaching emphasis on critical thinking, the adjusted mean posttest score on the Watson-Glaser Test of Critical Thinking between the MRLI and TL groups and between the MRL and TL groups was 7.2 and 7.4 points, respectively. This result indicates limited practical significance to the statistically significant gain in critical thinking due to the MRL and MRLI approaches.

The design, analyses and comparisons reported in this study were appropriate to address the established research goals of assessing the effects of different teaching emphases on these students' understanding of science, ability to do critical thinking, and mastery of content. The investigators were able to infer a causal effect of the MRLI treatment on increased achievement on the TOUS and Watson Glaser Critical Thinking Appraisal due to the assumed random assignment of subjects into the three treatment groups, as well as meeting the preconditions necessary for such inferences. The lack of classroom observations to verify treatments limits the legitimacy of these
results, however. The practical value of the differential gain in understanding of science for the MRLI group and the gain in critical thinking for the MRLI and MRL groups, as explicated above, is minimal although a shift in the desired direction.

Khishfe and Abd-El-Khalick (2002) compared the influence of explicit and implicit inquiry-oriented instruction on sixth graders’ views of NOS. The study emphasized the tentative, empirical, inferential, and creative aspects of NOS. The authors support their focus on these aspects based on their inclusion in reform documents as relevant for sixth grade science instruction. The implicit approach to NOS instruction assumes that the learner will come to understand NOS as a by-product of engaging in inquiry-based activities. The explicit and reflective approach assumes understanding NOS is a cognitive learning outcome and must be planned for accordingly. The empirical literature cited in this study supports the effectiveness of the explicit and reflective approach over the implicit approach in advancing learners’ NOS conceptions. The purpose of the present study was to explore the relative effectiveness of these two approaches on sixth grade students’ NOS views.

Sixty-two sixth grade students enrolled in two class sections in a private school in Beirut, Lebanon, served as the sample for this study. The authors give no indication of how the students were divided into the sections. One section, 33 students, was the explicit group. The other section, 29 students, was the implicit group. The authors state equivalent achievement levels for the two groups. They administered a six-item open-ended NOS questionnaire in a pre/post format. Validation procedures for the questionnaire included expert reviews and pilot testing. Eight students from each group were purposefully selected for interviews after both
administrations, for a total of 32 student interviews. The interviews served to further validate the questionnaire by ensuring common interpretation of items and responses between respondents and researchers. Interviewees were chosen based on science achievement and gender, to allow for wide variance in types of responses. Different students were interviewed at the beginning and end of the intervention. The first author served as the teacher of both groups, administered the questionnaires, and conducted the interviews. Given that she was the teacher, the students may have viewed the questionnaire and interview as evaluative and in need of a “correct” answer. There is no indication by the authors of attempts to minimize students’ concerns.

The teacher taught six inquiry-based activities that included five activities in physical science (atomic structure, mixtures, phase changes, heat and heat transfer, and combustion) and one activity in Earth science (fossils). The teacher taught two, 50-minute science lessons per week to both groups, for the ten weeks of the intervention. Both groups experienced the same six activities using a guided inquiry model. Students worked in pairs or groups of three. The teacher introduced the activity by posing a question or problem. The students then clarify their questions, make predictions, design data collection, defend their procedures to the whole class, collect data, organize and analyze data, and pool results with the class. Students then engaged in whole-class discussions to derive generalizations from the investigations. The teacher guided discussions to address the science content and science process skills.
The explicit group then received reflective NOS instruction wherein the teacher addressed aspects of NOS depicted in the different activities. The authors report that by the fourth activity, the students in the explicit group were able to initiate NOS discussions, rather than wait for the teacher to ask the guiding questions. Targeted NOS aspects relative to each activity were not reported. There is no indication that different aspects or perspectives of aspects were addressed based on the content area or investigation studied. The implicit group did not address NOS. To equalize instructional time, the implicit group continued discussion about content or process skills.

Classroom observations were conducted through videotaped sessions. The two authors reviewed the tapes to ensure the validity of the intervention. Given that both authors knew the categories of each group, their observations were somewhat biased. A blind observation of classroom instruction to determine the extent of explicit and implicit instruction and similarity of guided instruction during the activities would be preferred for validating the integrity of the purported intervention.

Data were first analyzed by the first author (the teacher) and a blind analysis conducted by the second author. They discussed results until consensus was reached. Pre- and post instruction profiles were generated for each participant. Profiles described participants' views of the four targeted NOS aspects.

Results indicate preinstruction NOS views were similar for both groups, with 85% demonstrating naïve views of tentativeness, creativity, empirical, and inferential NOS. Khishfe and Abd-El-Khalick present quotes and descriptions of participants.
preinstruction views. The descriptions are categorized based on NOS aspect. The explicit group advanced in their conceptions of NOS more than the implicit group. Again, representative quotes and descriptions are presented. The implicit group showed gains in conceptions of the inferential NOS (7% initially to 18% post). The authors attribute this advance to the teacher’s need to clarify observation and inference during discussions of science process skills. However, the 31% of the explicit group gained more informed views of the inferential NOS. Overall gains for the explicit group included: 46% (tentative); 31% (observation vs. inference); 42% (empirical); and 31% (creative). The implicit group demonstrated a decrease in understanding of the creative NOS (7% initially to 4% post).

Khishfe and Abd-El-Khalick appropriately state that their findings indicate their explicit and reflective inquiry-oriented approach was more effective than the implicit approach. They also recognize the limited advancements in their participants. Less than 50% of the participants in the explicit group held informed views of the NOS aspects at the end of the intervention. There is much room for improvement. Understanding the creative NOS seemed difficult for this group of sixth graders. The authors suggest the difficulty in understanding NOS in general, as well as specific aspects of NOS, might be due to developmental constraints of 11-year olds. Furthermore, they suggest difficulties might arise due to general difficulties learners have with changing misconceptions in just 10 weeks of instruction. The researchers suggest a conceptual change approach with historical case studies might improve the effectiveness of the explicit instruction.
Finally, the authors suggest the context of the learning situation might impact learners' developing NOS conceptions. The science content that was more familiar with these participants was dinosaurs. The context of structure and matter was less familiar. Students had more difficulty relating NOS to the structure and matter context. It is not clear from the discussion whether the difficulty was encountered during the class discussions or whether the responses on the questionnaire were different for the dinosaur question and atomic structure question. If there was a difference in the classroom discussions, the difference might be due to the fact that the structure and matter lesson occurred early in the intervention and the fossil activity was the final activity. One would expect the students to be better able to relate NOS aspects to the fossil activity after having engaged in five previous similar focused discussions. Furthermore, it would be expected that the students' responses on the dinosaur and atomic structure questions could be similarly disparate. The authors do not raise this issue. So the question remains, is the difference in student ability to relate NOS within different contexts due to the subject matter itself, or due to the chronology of instructional experiences. They do recommend more investigation is needed into possible relationships between subject matter knowledge and conceptions of NOS.

The three remaining reports reviewed in this section deal with developing understandings of NOS and/or inquiry through various approaches to laboratory instruction in college level science courses. Spears and Zollman (1977) compared the influence of two teaching strategies in college physics laboratory courses, structured and unstructured, on students' understanding of the process of science. They stated
that college laboratory experiences are generally intended to provide the student with experiences that enable them to understand the process of science, not just science content. However, instructional activities in the science laboratory vary from quite structured activities with detailed procedures to exemplify the content students' learn in lecture, to more unstructured, open-ended inquiry-type activities that leave procedures and analysis to the discretion of the student. There was a suggestion that a student must engage in some type of unstructured inquiry-type activity to develop an understanding of the nature of scientific inquiry. The purpose of this study was to compare the effectiveness of structured and unstructured approaches to inquiry-oriented laboratory activities on college students' understanding of the process of science. There were two specific intentions of this study: (1) to examine whether a student who engages in some degree of scientific activity gains a better understanding the process of science as a result of that activity, and (2) to examined whether the structure of the laboratory activity influences the degree of understanding of the process of science.

The student sample for this study was from four lecture sections of Man's Physics World I during the spring semester of 1973 at Kansas State University. Ninety-six percent of the students were freshman and sophomores taking the course to fulfill a science requirement. Ninety percent had taken high school biology. Only 23% had taken high school physics. The majority of the students had taken college biology or geology. The authors defined this group of students as typical nonscience students. The authors reported that out of the 171 students enrolled in the four sections of the physics course, only 50% of the students were usable in the study...
because of absent pretest or posttest scores. This 50% mortality rate affected the internal validity of this study and could limit the implications of results. The loss of subjects suggested the sample was potentially no longer representative of typical nonscience students.

This study used a pretest-posttest, two-treatment design. The independent variable was the type of laboratory experience, structured or unstructured. Students from four different lecture sections were randomly assigned laboratory sections. There was no report of how many students were in each section, or how many sections of each type were assigned. Both laboratory approaches required students to investigate physical principles presented in lectures. Both posed a problem to be studied and informed students of available equipment. The structured laboratory then stated explicit instructions the students were to follow to address the given problem. Instructions included how to do procedures and analyze the data. Students were given specific questions to address regarding the investigations. The unstructured laboratory required students to design their own approach to the given problem. Students then collected, analyzed, and presented the data in ways they determined to be appropriate. Both approaches required students to draw conclusions or inferences from their investigations. The authors contended that these requirements were appropriate for students in the concrete operational stage of cognitive development. However, they did not provide any evidence in support of their claim with respect to these subjects. In addition, they provided no description of the laboratory instructor(s). There was no information concerning the instructor’s knowledge of or training in the teaching approaches. Moreover, there was no report of any
observations being made to verify the consistency of intended laboratory instruction. Such descriptions and observations are required to properly assess the treatment groups.

The dependent variable was represented by the student scores on the *Welch Science Process Inventory Form D* (SPI) (Welch & Pella, 1967-68), a test to measure students' understanding of the process of science. Predictive and construct validity had been established previously (Welch & Pella, 1967-68), but not for this particular study. A Kuder-Richardson reliability of 0.86 was established on a sample of 171. No information on reliability of the instrument for the reduced sample was provided.

The instrument addressed four elements of the scientific enterprise: Assumptions, Activities, Nature of outcomes, Ethics and goals. The test required an agree/disagree response for each of 135 statements concerning these aspects of the process of science.

Students were pre- and post-tested during a lecture session of the first and last week of classes, respectively. Data on past science experience, major, years in college, grades, reading ability and mental ability were collected and used as covariates in the analysis. These variables were considered to contribute to students' pretest scores. In addition, there were four different lecture sections, with four different instructors. Therefore, there was increased variance due to different instructional procedures and degree of discussion of scientific processes. The authors included this as an additional covariate in the study. However, no further information regarding the extent to which lecture variability was analyzed as an influence on students' SPI scores was reported. No observations of lecture sessions were reported.
Analysis of covariance was aptly performed on students' SPI scores. A separate analysis of each of the four components of the SPI was performed. The pretest score, laboratory grade, and lecture instructor were determined to be major contributors to the posttest scores, and so were included as covariates along with past science experience. The adjusted posttest scores were then compared to the type of laboratory instruction. In the analyses, individual student scores were inappropriately used as the unit of analysis, rather than mean scores for intact laboratory sections. The authors do not report a final sample size. Because of the 50% mortality of the sample, there is uncertainty regarding the extent to which those students who actually took both the pretest and posttest consistently went to the laboratory sessions. There was no indication that attendance in the laboratory portion, thus exposure to treatment, was considered in the analysis.

The adjusted posttest scores indicated no significant differences ($\alpha = 0.05$) in the two laboratory groups in the areas of Assumptions, Nature of Outcomes, or Ethics and Goals. However, the two groups were significantly different in the Activities component of the SPI ($F = 4.7, p < 0.05$). The adjusted mean scores on the Activities component for students in the structured laboratory was 46.3, and 45.0 for students in the unstructured laboratory. Thus, it was reported that students in the structured laboratory group gained a better understanding of the process of science than the students in the unstructured laboratory. However, what is the implication of the 1.3 points difference in mean scores between the two groups? Even though this difference was determined to be statistically significant, the practical importance of such an increase is minimal. Interestingly, the mean scores for the Activities
component of the SPI were much higher than mean scores for the other components. After the Activities component, the highest score was in the Nature of Outcomes component with the structured group mean of 27.6 and the unstructured group mean of 26.9. The other two components, Assumptions and Ethics and Goals, were 10 points lower for both groups. The mean scores for the structured group were consistently higher than the mean scores for the unstructured group. However, the percent difference in mean scores between the treatment groups was 2.5%, 2.8% and 2.6% for the Activities, Nature of Outcomes and Assumptions components, respectively. The fourth component, Ethics and Goals, had a mean score difference of 1.3%, and reported the lowest scores of all four components for both treatment groups. Thus, even though the Activities component was reported to have statistically significant differences for the two treatments, the magnitude of the difference was the same as two of the other components, only a 2.5% increase.

Spears and Zollman (1977) argued that the students in the structured laboratory gained a better understanding of the process of science because they were explicitly led through the process with examples of activities of scientists. Students in the unstructured laboratory were left to decide for themselves an approach to take on a problem. They suggested that because many of the students in this study “did not as yet apply formal-operational processes to physics, they seldom followed the steps of observation, model building, and testing of the model” (p. 37). The assumption made by the authors was that students must have formal-operational cognitive abilities to effectively engage in scientific inquiry. Thus, the authors concluded, these students were not intellectually able to hypothesize or predict. The
students tended to relate their observations to information presented in lecture or the textbook, which, the authors argue, was more consistent with Piaget’s concrete-operational processing. Again, however, no evidence to support this interpretation was provided.

Spears and Zollman (1977) proposed two research questions in this study. The first was whether a student who engages in some degree of scientific activity gains a better understanding of the process of science, as measured by the SPI, as a result of that activity. They did not directly address this question by their design because they did not have a comparison group who did not do any activities. They can only state that both groups in this study showed gains, but they cannot make any causal statements relative to the laboratory experiences. Their second question of whether the instructional structure of the science activity influences the degree of understanding of the process of science was addressed. They concluded that college students, with little or no experience in scientific investigation, would gain a better understanding of the process of science in a course providing structured activities. Nonetheless, the extent of differential gain in understanding of the process of science as measured by the SPI was quite minimal, and pertained only to the “Activities” component where both groups scored considerably higher than on the other components of the SPI. Moreover, neither group scored particularly high on any component of the SPI. They do acknowledge that the results were limited to this population of liberal arts students.

Perhaps the most relevant aspect of this paper in regard to this present literature review was that the unstructured laboratory did not enhance students’
understandings of science as measured by the SPI. The unstructured laboratory was designed as an open inquiry-based approach with students only being given the problem to address. Thus, when compared to the more traditional, guided laboratory, the open inquiry approach did not enhance students’ understandings of science as measured by the SPI. It should be noted, however, that the high mortality of the sample might have impacted the results of this study. Finally, due to the lack of description of activities and learners’ views about science, the reader doesn’t gain information on possible relationships between contexts of inquiry (e.g. methods, subject matter) and views about science.

Classroom-based inquiry and students’ conceptions of NOS may correlate with the classroom climate. The influence of two types of classroom climates on process and content achievement was the focus of two investigations by Haukoos and Penick (1983, 1985). They designated two types of biology classrooms, one as discovery classroom climate (DCC) and the other as nondiscovery classroom climate (NDCC). Here discovery refers to the degree of freedom given the students by the teacher. The degree of freedom is dependent on the amount of directness in the teaching. Direct teaching, or direct verbal behavior by the teacher, includes lecturing, giving directions, reciting facts, and criticizing or praising. Thus, in direct teaching, active student involvement is minimal. Indirect teaching, on the other hand, includes eliciting student statements and listening and accepting student ideas wherein the student has more freedom of interaction. Thus, indirect teaching environments are more student-centered than teacher-centered. Haukoos and Penick (1983) cite published research studies that support the use of indirect teaching environments to
improve student achievement in high school biology, development of inquiry skills, creativity, and more independent or exploratory learning. Furthermore, they cite studies that demonstrated the effectiveness of less restrictive classroom environments in student gains in abilities and understanding of scientific investigations. These studies, however, are restricted to primary and secondary classrooms. At the time of the initial study by Haukoos and Penick (1983), there was little information about the effects of classroom climate in the college science classroom.

The purpose of the two studies by Haukoos and Penick (1983, 1985) was to compare the influence of indirect teaching in a Discovery Classroom Climate (DCC) and the more direct teaching in a Nondiscovery Classroom Climate (NDCC) on college students’ understanding of the processes of science and content achievement. The initial study (Haukoos & Penick, 1983) involved 78 college students from a midwestern community college with an approximate enrollment of 26,000. Four sections of an introductory college biology course were selected for treatment groups. Two were randomly designated as the DCC sections, and two were randomly designated as the NDCC sections. Both NDCC (n=25 and 19) classes and one DCC class (n=23) lasted 10 weeks. The other DCC class (n=11) was designed as an accelerated five-week course, still covering the same material and activities. No further information is given regarding the students in these sections or how sections were selected. The investigators did not report collection of background information from the students, such as grade level, science background, major, reason for taking the class, and reason for choosing a particular section (if choice was an option). Without such information, there is an unjustified assumption that the four groups
were equal. In addition, there is no way to determine how representative the sample was to the student population. Any conclusions drawn from this study can only apply to this particular sample.

Both treatments utilized the same content, but differed in teacher behavior (direct or indirect) and the nature of the activities performed by the students. The lecture/discussion sessions followed the same text materials and incorporated 35-mm photographic slides into the presentations. The DCC lecture sessions included student-centered discussions of slides and text materials. Questions in the DCC sessions allowed students to explore their own ideas about the topics without judgment by the teacher as being right or wrong. Lecture sessions in the NDCC classes included presentation of text materials in a manner that restricted student discussion and ideas. The impression conveyed by the format of the NDCC classroom was intended to be that “science was a complete, unaltered body of knowledge” (Haukoos & Penick, 1983, p. 634). The authors did not cite the reference of the text used in the study.

In the laboratory, students in both groups utilized the same materials and performed the same experiments. In the two DCC sections, students were encouraged to explore, select, manipulate, and explain the materials and results in whatever way they saw appropriate given a variety of questions (i.e. more open inquiry). The teacher’s role in the DCC classrooms involved asking open questions, actively observing, acknowledging student behaviors and listening to students. In the DCC classroom, the teacher offered little direction or evaluation of students’ activities. In contrast, students in the NDCC were told what and how to do the
laboratory activities. The teacher's role in the NDCC classrooms was high-profile which involved giving praise or rejection of students' results, frequently asking recall questions, demonstrating procedures, and giving directions. The treatment was 10 or 5 weeks, depending on groups. However, no information was given regarding the amount of instruction time in lecture/discussion sessions or laboratory sessions. No information was give regarding specific subject matter addressed.

To reduce variability between treatments, one of the investigators was the teacher for both groups. Audio-tapes of daily student-teacher interactions were analyzed and placed into categories defined as the Science Laboratory Interaction Categories (Shymansky & Penick, 1979). No information regarding the validity or reliability of this instrument was provided in this report. These interactions were then compared with the behavior criteria stated for the DCC (asks extended thought questions, acts as active observer, acknowledges student behaviors, listens to students) and NDCC (asks factual recall questions, demonstrates procedures, gives directions, gives positive or negative responses to students) classroom climates. For both treatments, a Learning Condition Index (LCI) was determined by considering the total class time spent on each of the coded behaviors. The LCI values range from 0, indicating teacher behavior is nonrestrictive to students, to 1, indicating complete restriction. The average LCI values for the DCC and NDCC classes were 0.06 and 0.64, respectively. The authors included the LCI values for weekly random assessments of DCC and NDCC teaching patterns. Values were given for each of the four treatment groups for 9 weeks and then averaged to determine the overall LCI values. The table in the report was contradictory to the text discussion of the
determination of the LCI values. Haukook and Penick first stated that daily audiotapes were analyzed. The presented data, however, indicated analysis was done on a randomly chosen tape or tapes per week. What was actually done to verify the treatments is unclear. Furthermore, without explanation, they presented 9 weeks of data for the accelerated 5-week DCC treatment group. There was no explanation of how they obtained data for 4 extra weeks. They concluded the results of the LCI analysis assured teaching behavior congruent with the established criteria for the two treatments. However, apparently the lecture/discussion portion of the course was not recorded or observed. Similar analysis in the lecture/discussion portion would be necessary to fully validate the treatment. Furthermore, and perhaps most problematic, the analysis of the audio-tapes, assuming done by the authors, could have been biased because they knew which groups were the DCC and the NDCC. A blind analysis, including classroom observations, would eliminate such possible bias.

This present study utilized a pretest/posttest design. The Science Process Inventory (SPI), Form D (Welch & Pella, 1967-68) was used to measure the influence of classroom climate on students' achievement of science processes. The SPI consists of 135 agree/disagree statements. A Kuder-Richardson reliability of 0.86 was determined in this study. Predictive and construct validation involved "using the literature, devising a model, employing the judgment of 'experts,' getting feedback from preliminary studies, and testing the instrument for its ability to distinguish among different groups of examinees" (Haukoos & Penick, 1983, p. 632). The reported reliability values and previous validation of the instrument appear adequate to justify the use of the SPI in this study.
The *Biology Achievement Test* (BAT) was used to assess content achievement after treatment only. This test was developed by biology faculty at the same community college for use in the Office of Testing. The authors reported that neither formal validity nor reliability of the BAT was measured. Therefore, there was no justification for the use of the BAT in this study. Conclusions based on the BAT scores are precarious due to the lack of validation and reliability.

The authors provided no further information on when, where, or by whom the tests were administered. Additional variation was introduced if the instrument was not administered in a consistent manner to all groups.

Analysis of covariance, using SPI pretest scores as covariate, was appropriately used to determine if differences existed in SPI scores for the DCC and NDCC groups ($\alpha = 0.05$). Duncan’s Multiple Range Test for variability was used to identify the source or sources of any differences. The class was appropriately used as the unit of analysis for the ANCOVA. Analysis of variance was appropriately used to determine differences between groups in terms of content achievement as measured by the BAT. The class was appropriately used as the unit of analysis in the ANOVA.

ANCOVA indicated significant differences in SPI scores for the DCC and NDCC treatments (Covariate F value: 94.76; Treatment F value: 3.86; $p < 0.05$). Duncan’s Multiple Range Test for variability identified the 10-week DCC treatment as having a significantly higher mean (111.17) on the SPI than the NDCC (103.36 and 100.58) and the five-week DCC (102.09) classes ($\alpha = 0.05$). ANOVA indicated no differences in content achievement between the groups as measured by the BAT.
As the BAT had neither confirmed validity or reliability nor pretest measures, the results of this analysis are meaningless.

Haukoos and Penick (1983) drew three main conclusions from the information gathered on students' understanding of science processes: (1) Classroom climate does influence the learning of science processes; (2) Sufficient time would appear to be needed to learn science processes; and (3) Students may loose their knowledge of science processes if they are not continually emphasized through use during instruction.

Students in the 10-week DCC group gained a better understanding of the processes of science than students in the NDCC groups and the 5-week DCC group. The reported difference in mean SPI scores between the 10-week DCC class and the other treatment groups ranged from 8 to 11 points on a 135 item test. However, no statistical error analysis is provided. Thus, the practical significance of the reported difference is somewhat questionable. The authors concluded that classroom climate, specifically an indirect or discovery environment, improved students' understanding of science processes. Given the extent of improvement reported, the DCC approach is only minimally effective compared to the NDCC approach.

The significant difference in mean SPI scores between the 10-week and 5-week DCC classes, the authors suggested, is indicative of the need for extended time to develop the understanding of science processes in such an unrestricted classroom climate. These results support other cited research findings. Again, the practical significance of this difference is minimal considering the two DCC mean SPI scores differ by only nine points.
Upon comparison of mean pretest and posttest SPI scores, the authors identified both the 10-week and 5-week DCC classes increased student knowledge of science processes. In contrast, the posttest scores on the SPI were actually lower than the corresponding pretest scores for the NDCC classes. The authors concluded that if science processes are not emphasized and developed by students, the students might lose their knowledge of these processes. From a graphical display of the change in scores between the pretest and posttest for each treatment group, the drop in mean scores for the two NDCC groups appears to be approximately 3 points for both groups. The gain in mean scores for the DCC groups is also 3 points. The practical significance of such a drop or a gain is again questionable. Even if one accepts the drop in SPI scores for the NDCC group as significant, a likely explanation for apparent loss of understanding is not that the processes of science are not emphasized in the NDCC classes, but that the emphasis promoted misconceptions and inaccuracies of the nature of the processes of science. The classroom climate of the NDCC classes was such that students were given set procedures and told if their results were right or wrong. The intent of the NDCC, as explicated by the authors, was to present science as an unchanging, authoritative piece of knowledge. As such, the students acquired inaccurate conceptions of the nature and the processes of science.

Finally, Haukoos and Penick (1983) reported no statistical differences in content achievement among any of the treatment groups. The authors concluded that the DCC and NDCC treatments were equally effective in influencing the learning of course content. In addition, “concentrated course time was equally effective for
learning course content as it was for minimal learning of science processes” (p. 635). The authors suggested the equal content achievement results from students having equal motivation to pass the course, regardless of classroom climate. The aforementioned lack of validity and reliability for the BAT instrument makes any conclusions drawn from this test dubious. Worthy of notation is that the authors did not confine their conclusions to the sample in the present study. The description of the sample does not sufficiently identify the participants in this study as representative of any larger population. Hence, the generalizations made by the authors were inappropriate.

In a replication study, Haukoos and Penick (1985) addressed the same research question concerning the differential effects of classroom climates on students’ understanding of processes of science. Again, they compared the effects of a Discovery Classroom Climate (DCC) and a Nondiscovery Classroom Climate (NDCC) on college students’ understanding of science processes as well as content achievement. The investigators in this study were the same as in the initial study. They used the same materials, classroom, and teacher (the primary investigator) as in the initial study. Differences in conditions between the two investigations included dates of the treatment, individual students, and the time span of the two treatments. All sessions lasted 10 weeks.

The investigation was again a pretest/posttest, two-treatment design. The sample for this investigation consisted of 61 college students from the same community college as the initial study. Twenty-three students were assigned to one NDCC group, and 38 students were assigned to one DCC group. Fifty-two percent of
the NDCC and 61% of DCC students were female, with a mean average age of 24.6 and 22.0 years, respectively. No reasons are stated for the difference in group sizes. No further information was provided on the sample. As in the initial study, the results of this investigation should not be generalized past this particular sample.

Classroom climates for the discovery and nondiscovery groups were as previously described (Haukoos & Penick, 1983). The climate was determined by teacher behavior and the nature of the activities the students performed. In summary, the teacher in the DCC class asked extended thought questions, acted as an active observer, acknowledged student behaviors, and listened to students. In contrast, the teacher in the NDCC class asked factual recall questions, demonstrated procedures, gave directions, and gave positive or negative responses to the students.

Teacher behavior data was again obtained by use of audio-tapes of laboratory sessions and categorization of interactions according to the Science Laboratory Interaction Categories. In this study, a 10-minute segment was randomly selected for categorization and then compared to the aforementioned established behavior criteria for DCC and NDCC. A Learning Conditions Index (LCI) was determined as previously described. The mean LCI values for the NDCC and DCC treatments were 0.63 and 0.07, respectively. These results could be biased, as in the previous study, due to the investigators having knowledge of the type of classroom climate they were analyzing. This type of analysis would best be done blind, without prior knowledge of intended classroom climate. In addition, a similar analysis of the lecture/discussion portion of the class is necessary to more fully verify the consistency of the treatments.
Achievement in science processes was measured by the *Science Process Inventory* (SPI). Achievement in content was measured, posttest only, by the Biology Achievement Test (BAT). Validity and reliability measures were described in their initial report (Haukoos & Penick, 1983). Reliability for the SPI was not reported on this sample. It is assumed that neither reliability nor validity was yet established for the BAT. As previously, they appropriately used analysis of covariance to determine effects of different classroom climates on students’ SPI scores. They appropriately used analysis of variance to determine differences in student achievement based on BAT scores. The unit of analysis was again based on the class (n=2) and not individual student scores.

Students in the two treatment groups again achieved equally well on the BAT. The authors suggested that the consistency between the initial and replication study on student content achievement lends support to the often criticized view that a discovery-type environment was as effective for student learning as a direct, content-presentation, approach. All that could really be stated about the results on the BAT was that both samples achieved equally well on this instrument. Because the BAT had not been validated, yet alone validated for college level biology students, and no pretest measures were taken, there was no certainty that the test measures achievement of content gained in this college biology course.

In this replication study, Haukoos and Penick (1985) did not get results consistent with the initial study in regard to student achievement on the SPI. The same statistical analyses were employed ($\alpha = 0.05$). Analyses revealed no significant differences in mean SPI scores between the DCC and NDCC groups (Covariate F
value: 17.91, p < 0.05; Treatment F value: 0.55, p = 0.46). The authors suggested explanations for the inconsistency between the two studies by addressing possible problems with both studies. First, they considered the possibility of instructor bias in the initial study due to the novelty of the discovery and nondiscovery classroom climates (Hawthorne effect). They did not expand on the nature of this potential bias, though. As previously explicated in this review, the bias could have been due to the investigators’ having knowledge of the treatment groups during the categorization of teaching behaviors.

They suggested the instructor could have biased the second study by making the two treatments more similar than in the initial study. Such behaviors might not have been distinguished with the Science Laboratory Interaction Categories (SLIC) or LCI determinations. They argued that the SLIC records specific teacher behaviors, but does not consider specific patterns of teaching behavior. Thus, differences in teaching patterns between the two studies would not be detected. Such differences might include timing of questions and reactions to specific student responses. An outside observer would address these potential problems.

The authors appropriately recognized that there might have been significant differences in intelligence, motivation, or other aspects between the two student samples. There is no way to obtain identical participants in such replication studies as the one reported. Collection and analysis of demographic information from subjects would aid in determining equivalence of two separate samples.

Haukoos and Penick (1985) concluded this report with suggestions for improving such investigations. They suggested expanding the interaction analysis to
include nonverbal behaviors and student-student interactions. Finally, they suggested the need to examine specific teaching patterns in all treatment groups. These additional examinations of interactions would help in the analysis of actual treatments administered. Blind observations would be required to minimize the potential bias.

Missing from their interpretation of the two studies was the possibility that classroom climate had no effect on students’ understanding of science process as measured by the SPI. Rather, they chose to limit the causes of inconsistencies to methodology. Clearly from their unwillingness to recognize a plausible alternative, there is inherent bias to the study. The lack of consistency between the two studies and the authors’ incomplete speculations raise doubts concerning the legitimacy of either study. On that account, it becomes apparent that these studies do not support the notion that classroom climate, specifically direct or indirect teaching behaviors, necessarily affects the degree to which students develop an understanding of science or science processes during inquiry-oriented activities. Again missing from the results is any qualitative description of the learners’ views of science that might be associated with classroom climate.

Summary. The six reports reviewed here reveal a consistent pattern to the effectiveness of inquiry teaching. That is, different emphases or approaches to inquiry teaching have different effects on different students taught by different teachers. Even given the limitations of the studies, no single described inquiry approach works consistently better than any other at establishing adequate conceptions of NOS or the nature of scientific processes in students. With the
exception of the Khishfe and Abd-El-Khalick (2002) study, there are no descriptions of NOS or inquiry conceptions that might relate to subject matter or investigative methods. Even though the Khishfe and Abd-El-Khalick study provides qualitative descriptions of learners’ views, their method did not allow for distinguishing among effects of different subject matter or investigations on NOS views. Their method does suggest an overall differential effect on NOS views due to explicit or implicit instruction. They also suggest that different views may be expressed based on the context of subject matter under consideration. Whether differences are due to levels of familiarity with the subject matter or differences in the nature of the subject matter itself cannot be determined from this study. More exploration is needed.

Yager and Wick (1966) demonstrated that an inquiry-based curriculum was only effective, albeit minimally, in enhancing students’ conceptions of NOS or the nature of the processes of science if the teacher emphasized those aspects purposefully in the biology classroom. Attention to the historical development of science, the cultural influence on science, the tentative nature of science, and the connection between the laboratory activities performed with actual science processes resulted in students’ better achievement on the TOUS. In the study by Spears and Zollman (1977) the structured group achieved higher gains on the Activities component of the SPI. These students were given structured physics activities to follow in the laboratory. Students in the unstructured group were left to decide for themselves what activities to perform. As such, they did not demonstrate a gain in the Activities component of the SPI. However, problems with high sample mortality may have impacted these results. The two Haukoos and Penick studies (1983, 1985)
demonstrated that classroom climate does not have a predictable differential effect on students' learning. Yager et al. (1969) found that none of the approaches they studied in the biology classes was effective in fostering adequate views of NOS in the students. They do suggest the inclusion of reference to NOS may aid in effecting the desired changes in students' views, regardless of the inquiry approach. This suggestion was confirmed by Khishfe and Abd-El-Khalick (2002) with regard to the impact of explicit vs. implicit instruction.

The Influence of Inquiry-based Experiences on Teachers' and Preservice Teachers' Views of NOS

The impact of studies that investigated the effects of different teaching approaches to the same inquiry-based curriculum turned much of the research focus on the teachers themselves. None of the studies thus reviewed assessed the views of the teachers who implemented the inquiry-based curricula. Questions arise concerning teachers' understanding of NOS, including the development of teachers' views. The current section of this literature review examines published reports of teachers' NOS learning outcomes that developed within the context of inquiry-based experiences. Although the literature base contains reports on the effectiveness of entire programs or courses designed around a NOS theme (e.g. Akindehin, 1988; Lederman, Schwartz, Abd-El-Khalick, & Bell, 2001; Schwartz & Lederman, 2002), these reports are not included here because the effects of learning of NOS based on inquiry experiences or different inquiry contexts can not be distinguished from NOS learning resulting from a thematic focus on NOS within a variety of contexts. Thus,
associations between learning experiences (e.g. inquiry activities, direct instruction, readings, peer interactions) and NOS or inquiry conceptions are not discernable.

In as much as the teacher must have an adequate understanding of NOS to effectively teach NOS (Carey & Stauss, 1970; Hurd, 1991; Lederman, 1992; Ramsey & Howe, 1969; Robinson, 1969; Rutherford, 1964; Schwartz & Lederman, 2002), he/she must also have an understanding of the processes by which scientific knowledge is created to effectively incorporate inquiry-based activities or projects to teach NOS (Gallagher, 1991; Herron, 1969, 1971; Horner & Rubba, 1978; Ramsey & Howe, 1969; Riley, 1979; Robinson, 1969; Rutherford, 1964; Schwab, 1962; Schwartz & Lederman, 2002). Studies conducted to evaluate the adequacy of teachers’ views of NOS have revealed the poor status of teachers’ understandings (Behnke, 1961; Benson, 1989; Carey & Stauss, 1968, 1970; Duschl, 1990; Duschl & Wright, 1989; Gallagher, 1991; Kimball, 1967-68; Lederman, 1992; Miller, 1963; Pomeroy, 1993; Schmidt, 1967). Furthermore, teachers indicated a lack of experience and understanding in the processes of science (Barufaldi, Bethel & Lamb, 1977; Harms & Yager, 1981; Riley, 1979; Shapiro, 1996). It was suggested that teachers need experiences with scientific inquiry in order to develop knowledge of NOS. “Prospective teachers have limited knowledge of, and experience with, the processes by which scientific knowledge is generated. This puts serious limitations on their ability to plan and implement lessons that will help the students develop an image of science that goes beyond the familiar ‘body of knowledge’” (Gallagher, 1991). Furthermore, Rutherford (1964) recognized the need for teachers to have
adequate philosophical views concerning the nature of scientific inquiry in order to teach science as an inquiry process:

[Teachers] must come to understand just how inquiry is in fact conducted in the sciences. Until science teachers have acquired a rather thorough grounding in the history and philosophy of the sciences they teach, this kind of understanding will elude them, in which event not much progress toward the teaching of science as inquiry can be expected (p. 84).

Presented here is a review of five studies that examined courses or programs that incorporated some level of inquiry activities with the intent of enhancing teachers’ views of NOS.

Billeh and Hasan (1975) conducted a study to identify factors that enhance science teachers’ understandings of NOS. They developed a summer training course for secondary science teachers in the country of Jordan. The program consisted of four components: (1) lectures and demonstrations in methods of teaching science and basic science concepts, (2) laboratory investigations related to secondary science, with an emphasis on what they called guided discovery (authors offered no further description), (3) enrichment activities including readings and films to present science concepts and represent scientists and their work, and (4) twelve, 50-minute lectures on NOS, including science and common sense, science and technology, art of scientific investigation, nature of scientific knowledge (characteristics, classification, scientific theories, and models), growth and development of scientific knowledge, and sociological aspects of science. The two main purposes of the study were, (1) to identify if teachers’ understandings of NOS were significantly increased as a result of attending the summer course, and (2) to identify any significant relationship between teachers’ gains in understandings and four variables (number of years...
education in college science, science subjects taught, number of years science teaching experience, and kind of previous professional training).

The Ministry of Education requested all 186 secondary science teachers in Jordan attend the 4-week summer program. Ninety-two percent of these teachers participated, and thus comprised the sample for the study (N=171). The average biographical information on the sample was collected by means of a questionnaire. The data included the median age of the teachers as 26 with 2.8 years of teaching experience, and 12.8% who had attended other in-service science professional training. Most had taught more than one science subject, and 68% had graduated from universities rather than two-year teacher training colleges. The teachers were divided into four groups based on their primary teaching area: Biology (N=30), Chemistry (N=41), Physical Science (N=58), and Physics (N=42).

The Nature of Science Test (NOST), a 60-item, multiple-choice test, was used to measure participants' views of four components of NOS. These included the assumptions in science, products of science, processes of science, and ethics of science. The authors contend that the lectures on NOS were not focused on specific items on the NOST. A panel of 25 scientists, science educators, and science supervisors established content validity. Split-half reliabilities of the test administered to secondary students, preservice teachers, and undergraduate science majors ranged from 0.58 to 0.82. For the sample in this study, reliability values ranged from 0.58 to 0.78. The authors did not address the low reliability value of 0.58 calculated for the physics posttest as a potential problem.
The NOST was administered at both the beginning and end of the 4-week program to all the teachers. All teachers experienced the same course, except the teachers in the biology group. To establish a reference in an effort to test the influence of NOS lectures, the investigators did not have the biology teachers participate in this part of the program. There was no indication of instructional observations to verify the consistency of the course as experienced by all the teachers.

Analysis of covariance was appropriately used to determine if differences existed in NOST scores among the four groups ($\alpha=0.05$). The mean pretest scores were used as covariates in the analysis. The mean posttest scores for the four teacher groups served as the dependent variable. Duncan’s, New Multiple Range Test was used to identify where differences existed. One-way $t$-tests ($\alpha=0.05$) were performed to determine the significance of any increase in mean pretest and posttest scores for each group and the entire teacher sample. Analysis of covariance ($\alpha=0.05$) was also used to determine any significance in mean gain scores as a function of the teacher variables.

Analysis of the pretest scores revealed no significant differences between the four groups of teachers in their understandings of NOS as measured by the NOST. However this analysis wasn’t necessary given that an ANCOVA was used. ANCOVA revealed significant differences in the mean posttest scores ($F=5.78$, $p<0.01$). The physical science and chemistry groups scored significantly higher than the biology and the physics groups. The directional $t$-tests determined significant gains ($p<0.0001$) in mean NOST scores for the physical science (mean gain of 5.66),
physics (mean gain of 2.00), and chemistry (mean gain of 4.15) groups, but not for the biology group (mean gain of 1.67). For the entire teacher sample, there was significant gain in mean NOST score (mean gain of 3.71, p<0.0001). No relationship was found between mean NOST scores and the any of the background variables of the teacher participants.

Billeh and Hasan (1975) drew several conclusions from this study. First, because the biology group did not receive specific instruction on NOS, but participated in all other components of the program (instruction in methods of teaching and science concepts, laboratory investigations focusing on guided discovery, and enrichment activities), it was concluded that the gains in mean NOST scores for the other groups was due to the formal “nature of science” instruction they received. Also reported by Billeh and Hasan (1975) was the lack of a relationship between mean gain scores and educational background. They indicated that the teachers who had not attended 4-year universities had greater mean gain scores than those who did attend universities. The mean posttest scores for both groups, however, were the same. Similarly, the authors concluded that the mean gain scores were not related to subject taught, science teaching experience, or professional in-service training. Those teachers who initially scored lower on the NOST showed the highest gain scores, but the gains could not be attributed to any of the variables examined in this study.

The authors did not comment on the practical significance of the mean gain scores. It would seem that a gain of 6 points would be considered practically significant because that would account for a 10% gain on the NOST possible score.
The mean gain scores for these groups were less than 6 points, but the scores approach practical importance for the physical science and chemistry groups. Another way of measuring the level of gain in score is in the percent increase from the initial score. In the case of the physical science, chemistry, and physics groups, they gained 18.6%, 12.8%, and 6%, respectively. These results indicate the physical science and chemistry groups made substantial gains over their initial NOST scores. The minimal gain showed by the physics group, although statistically significant, have no practical significance. In terms of gains, it seems the subject area of the teacher may have a differential impact. The authors do not elaborate on how they determined subject matter did not relate to gains.

Moreover, Billeh and Hasan (1975) did not elaborate on their interpretation of the adequacy of the views held by the teacher participants. The mean posttest NOST score for the entire sample was 35.12 with the biology group averaging 33.17, chemistry with 36.51, physical science with 36.02, and physics with 33.64. With 60 points possible on the test, the teachers averaged 58.5%. This mean score would not likely be considered “adequate.” It should be noted that the purpose of the study was not to measure the adequacy of the views, but rather the significance of improvement in understandings as measure by the NOST. Nonetheless, the physics and biology teachers held similar posttest scores, as did the physical science and chemistry teachers. However, no description is provided regarding how similar their views really were, that is, if they scored similarly on the same instrument items.

An outcome of this study was that the necessary component to enhance teachers’ understandings of NOS was the explicit instruction on NOS. The groups
that received the 12 lectures showed significant gains in NOST scores. All groups participated in the laboratory component. Therefore, just participating in the guided discovery-type laboratories of this program was not sufficient to significantly improve the biology teachers' understandings of NOS. It must be stated, however, that there were not observations of the instruction the teachers received or the laboratory activities in which they engaged. Thus, a complete description of the program was not provided. The actual nature of the laboratory activities and the participants' engagement in inquiries is not known.

Furthermore, observations were not made for any of the components of the summer course. Because the investigators were also the instructors of the summer course, the potential for biased instruction and interpretation are increased. Classroom observations by an additional individual would help verify the equality of instruction for all groups. Even though the authors stated the test items could be divided into four categories of NOS, they do not divide the participants' scores by categories for comparison. Grouping the participants' scores by category in conjunction with making observations and descriptions of the formal lecture component on NOS would enable a comparison of the instruction to the aspects of NOS for which the teachers demonstrated better understandings. This method would also allow comparisons of scores within NOS categories with teaching area.

A study conducted by Barufaldi, Bethel, and Lamb (1977) was based on the premise that teachers' views of NOS as tentative or absolute influence the views developed by their students. As such, they argued, "teachers should view science as tentative if they are to adequately promote this objective of inquiry teaching" (p.
Science teacher education programs should, therefore, engender a view of the tentative nature of science in the preservice teachers. In their report, Barufaldi et al. explicitly stated that gaining an understanding of science as a tentative, rather than absolute, is an affective goal of science teacher education. Thus, they assumed, like many of the investigators in this review, that learning about NOS is achieved through activities and learning science content. This view is consistent with proponents of implicit teaching of NOS.

The purpose of their investigation was to study the effect of a science methods course on the philosophical view of the tentative nature of science among elementary education majors. The authors reported the experimental design was a nonrandomized, equivalent control group, pretest-posttest approach. Participants for this study were elementary science education majors in their junior and senior years at The University of Texas at Austin. Subjects were self-selected based on enrollment in the four elementary science methods sections offered. The experimental group consisted of 56 students in three elementary science methods sections, and one control group consisting of 32 students in an elective elementary mathematics course. Of the 56 students in the experimental group, 12 were senior elementary education majors enrolled in a student teaching block program that involved practice teaching, an elementary science methods course, and a course on the cultural foundations of education. Twenty-one were junior and senior special education majors, enrolled in an elementary science methods course. Twenty-three were junior elementary education majors enrolled in a field-based block program requiring observations in an elementary school setting, an elementary science
methods course, and an educational psychology course. The control group consisted of 32 junior elementary education majors enrolled in a field-based observation block program, involving observations in an elementary school setting, an elementary reading methods course, and an educational psychology course. No further demographic information was provided on the participants.

The science methods course taken by students in the experimental groups served as the treatment. The mathematics methods course served as the control group. The criterion measure was participants' responses to the Views of Science (VS) instrument, administered at the beginning and end of the courses. The VS was intended to measure students' views of the tentativeness of scientific knowledge. The instrument is composed of 40, 5-point Likert-type statements that are either consistent or inconsistent with the tentative nature of science. A low VS score reflects a view that scientific knowledge is absolute, whereas a high VS score reflects a more appropriate view that scientific knowledge is tentative in nature. No criteria concerning a "high" or "low" score were defined. Barufaldi et al. (1977) reported that face validity, predictive validity, and alpha reliabilities were established and reported in previously published studies. The referenced predictive validity measure were not specific toward elementary education majors, but rather a sample of college teaching assistants in physical science, college physical science students, secondary science teachers, and ninth-grade physical science students. Thus the instrument was not validated for the type of sample in this study. They did, however, establish reliability for the VS on the new population of participants. The alpha reliability for the combined group on the pretest was 0.78, and the alpha values for
the three treatment groups and the control group were 0.71, 0.84, 0.80, and 0.76, respectively. These values were similar to previously published measurements and adequately established the reliability of the VS for this sample. Nonetheless, a reliable instrument that is invalid is still invalid.

To support their assumption that the student participants were homogeneously distributed, the authors conducted several statistical analyses. A one-way ANOVA ($\alpha = 0.10$) indicated that the groups differed significantly in the number of college science semester hours taken by students in each group. However, $t$-tests determined that only the special education majors group had significantly fewer science hours than the other three groups. A one-way ANOVA also revealed the four groups did not differ significantly in mean pretest scores on the Views of Science (VS) instrument used in the study. A principal-components analysis and Varimax Rotation was conducted to assess any differences in mean group scores with respect to the factor structure of the VS instrument. No significant differences were determined. In addition, a multiple discriminate function analysis was calculated where the VS items were the independent variables and the group membership served as the dependent variable. Again, no differences were found. No correlation was detected between pretest scores on the VS and the semester hours of college science by calculation of a Pearson correlation coefficient. As such, the authors appropriately concluded the four groups were similar, and therefore comparable for the intended study.

The elementary science methods course taken by the experimental groups met for two and one-half hours a week for 14 weeks. The nature of the course was
such that the authors believed the students would gain the view that science is tentative. The authors stated “the uniqueness and variety of the learning experiences in the courses provided the students with many opportunities to understand the tentativeness of scientific findings” (Barufaldi et. al., 1977, p. 291). The students explored content, methodology, and processes of science through hands-on, activity-centered, inquiry-oriented science experiences.

Students in the elementary science methods course were presented with the following tasks:

1. Present a science experience to a child by trying out a science skill-oriented learning event.

2. Describe in writing the interaction that occurred while trying out the science skill-oriented learning event.

3. Construct and try out a mystery box. Audio-tape a small segment of the interaction that occurred while trying out the mystery box.

4. Complete a science skills test. Demonstrate a proficiency with the science skills: observing, measuring, inferring, classifying, predicting, verifying predictions, hypothesizing, isolating variables, interpreting data, and experimenting.

5. Present a science experience to a small group of children by teaching an activity from Science-A Process Approach, Elementary science Study, Science Curriculum Improvement Study, or BSCS-Elementary School Science Program. Audio-tape 10 minutes of the lesson. Analyze (verbal interaction tactics) the learning event.
6. Investigate some science problem. Set up, conduct, and report on an experiment involving a science question (Barufaldi et al., 1977, p. 291-292).

The authors provided no criteria for assessment of these tasks. These tasks were not required of the students in the elementary mathematics teaching methods course. The VS was administered as a posttest at the completion of the course.

The given description of the treatment is lacking in several important aspects. The authors provided no information regarding the contexts within which the above tasks were conducted. There was no indication that classroom observations were conducted to assess the extent to which the tasks were carried out as intended. As such, even though the authors describe the activities as "inquiry-oriented science experiences," the application of their intent is not known. Furthermore, the extent to which the tentative NOS was not a topic of discussion is not known. The authors make no mention of the instructor(s) for the methods courses. Therefore, it is not known if the instructors were aware of the purpose of the investigation and if the instruction reflected this purpose.

The authors used a 2-group covariance regression model to compare the mean VS scores between specific pairs of groups. They compared each treatment group with the control group and with each other. No differences were found between treatment groups (no significance level reported). It is not clear from the report what the authors used as the unit of analysis. It is assumed that the whole class (an individual methods course) was the unit of analysis.
Three-group comparisons were calculated, and significant differences were found between pairs of treatment groups and the control group. However, no differences were found among the three treatment groups. No significance levels were reported.

In all, 10 separate pairwise comparisons of means were calculated. There is a 40% chance of a Type I error by performing so many comparisons on the same data (p = 0.05). Six out of the 10 comparisons were found to be significantly different. An analysis of covariance, using the pretest VS scores as a covariate, would be a more appropriate method to determine if any differences exist between the groups. If a difference is indicated, post hoc tests such as Duncan’s multiple comparison or Scheffe’ method could determine the source or sources of the differences. Such an approach would minimize the occurrence of a Type I error.

Barufaldi et al. (1977) concluded from these data that the preservice elementary teachers in the experimental group did gain a better understanding of the tentative nature of science as compared to the control group. They suggested that an elementary science methods course that “stresses inquiry methods and procedures, emphasizing a hands-on approach integrated with individual problem solving, develops, alters, and enhances students’ philosophical view of science” (p. 293). Such a broad claim is an overstatement of the results and beyond the scope of the reported study. Given the lack of instructional observation, claims specific to the influence of the activities alone are unfounded. Other reports have described the actions and emphases of the teacher as influential as well.
Riley (1979) suggested that it is a lack of understanding of NOS that may make a major contribution to elementary school teachers’ uneasiness towards science. He argued that teachers are to teach a process-centered science curriculum rather than the previous content-centered curriculum. However, similar emphasis changes had not consistently occurred in elementary science teacher preparation programs. Riley cited empirical research suggesting the lack of understanding of NOS is the result of the lack of exposure to the processes of science in teacher education. He stressed the need for preservice teachers to be exposed to hands-on science experiences in order to become proficient in science process skills. Such proficiency, it was suggested, would provide the understanding and attitude necessary to increase and improve science instruction. Similar to other researchers, such as Barufaldi et al. (1977), Riley endorsed an implicit approach to teaching about NOS whereby students learn NOS simply by “doing science.” Riley examined two approaches to inquiry-based instruction utilized in an elementary science methods course.

The purpose of the investigation by Riley (1979) was to compare the effects of hands-on versus non-manipulative training in process skills on preservice teachers’ (1) knowledge of process skills, (2) understanding of science, and (3) attitude toward science, science teaching, and method of instruction. The relationship between science grade point average and the procurement of science process skills was also examined. Included as a covariate in the analyses was students’ science background as determined by credit hours of college science courses.
Participants in this study consisted of student teachers enrolled in an undergraduate methods program. Students were student teaching in grades 1, 2, 3, or 4. These students were divided into three groups based on undergraduate science grade point average (high, medium, and low), and then 30 students from each group were randomly assigned into one of three treatment levels. The final number of students involved in the study was 90. There was no report of the total number of students enrolled in the courses, or the relative proportion of students in the high, medium and low grade point average categories. Thus, it cannot be assumed the resulting sample was a stratified representation of the original student population in the course. Groups were assumed to be equivalent in terms of average group grade point average. Because of the random assignment of students into the final treatment groups, this was a valid assumption. No further information was provided on the student sample.

The investigation utilized a 3 x 3 factorial design. There were two independent variables: (1) science grade point average (high, medium, or low) and (2) treatment with three levels. The three levels of treatment were Active-Inquiry, Vicarious-Inquiry, and Control treatment. The active-inquiry treatment group was trained in use of “hands-on” and manipulative process skills. The vicarious-inquiry treatment group was trained in process skills involving the same content as the active-inquiry group, but without student manipulation of materials. In the vicarious-inquiry treatment, the teacher did the only manipulation through classroom demonstrations. The control treatment group spent the same amount of time viewing
films on geology, meteorology, and physical science. This was an effort to present a neutral treatment compared to the two process training approaches.

The AAAS Guide for In-service Instruction was used as the basis for treatment procedures in the active-inquiry and vicarious-inquiry treatments. The treatment was presented in four, one and a half hour sessions with the teaching strategy being the only difference in the treatments as described above. The first session focused on the skills of observing and classifying. The second session emphasized skills of inferring, communicating with graphs, and predicting. The third session focused on using numbers, measuring, and using the metric system. Finally, the fourth session emphasized using space/time relationships and using space/time relationships rate of change.

The Teaching Strategies Observation Differential (TSOD) was employed to ensure the treatments were delivered as defined. This classroom observation instrument utilized a matrix sheet to encode classroom activity at one-minute intervals. The TSOD potentially differentiates between verbal direct teaching and nonverbal activity-oriented class environment. No validity measures were reported. The author did not further explicate criteria for classification of the observed activities. Criteria for scoring of the instrument were not provided, but it was indicated that a higher score was consistent with students working in a laboratory setting (again no description of type of student involvement) and a lower score was consistent with students in a passive role. Two observers analyzed videotapes of all treatment presentations for the active-inquiry group and the vicarious-inquiry group, and inter-rater agreement was reached by consensus. The author included previously
reported inter-rater reliabilities (0.88 to 0.99), but did not establish inter-rater reliability for this study. Riley reported that the TSOD score for active-inquiry approach (5.9) was significantly higher than the TSOD score for vicarious-inquiry approach (4.5). Thus, he concluded the treatments were different and employed as intended. Nonetheless, this report lacked certain descriptions of the treatment verification process that would add support to their claims. First, it was not indicated whether the analyzers of the videotapes had prior knowledge of the treatment group that was being analyzed. If they did, then the TSOD ratings could have been biased in favor of the appropriate treatment approach. Second, there was no report of observations of the control group. Observations of all groups should be provided to verify the treatments and non-treatment.

There were four criterion measures in the present study. The Science Process Measure for Teachers, from AAAS Science – A Process Approach, was used to measure preservice teachers’ abilities in science process skills. A Hoyt’s reliability had been previously reported as 0.89. This instrument contains seven subsets: (1) using number relations, (2) classification, (3) using space/time relationships, (4) observing, (5) inferring, (6) measuring, and (7) communicating and predicting. The Test on Understanding Science (TOUS) was used to assess the preservice teachers’ understanding of NOS (Hoyt’s reliability = 0.58). The Attitude Toward Science and Science Teaching Scales were used to assess attitudes (Hoyt’s reliabilities of 0.71 and 0.84 for the two instruments, respectively). All reported reliability information was from previously published reports and not established by this investigator. The Attitude Toward Method of Instruction Inventory, developed by the author, was used
to assess attitudes toward instruction employed in the present study. Hoyt’s reliability coefficient was calculated (0.93). The first three instruments had previously published results of validity establishments. Riley (1979) stated that a panel of experts established face validity of the *Attitude Toward Method of Instruction Inventory*. No further information regarding these instruments was provided. Even though validity and reliability measures were adequately established in previous studies, a more detailed report of the measurements should have been included to assess the relevant use of these instruments for the sample in the present study.

Analysis of covariance was suitably conducted to assess differences between group mean scores on the criterion variables for treatment, science grade point average, and treatment by science grade point average interactions. The covariate was the number of science semester hours the preservice teachers had completed. The Newman-Keuls multiple comparison was applied where statistical differences were found. The Newman-Keuls is a conservative multiple comparison technique, and thusly protects against Type I errors for analyses of multiple pairwise comparisons. For all hypothesis testing concerning cognitive measures (knowledge of process skills), significance levels were set *a priori* at 0.10 as an acceptable risk of Type 1 error. Significance levels for affective measures were set *a priori* at 0.05. Riley included understanding of NOS, attitude toward science, and attitude toward teaching science as affective outcomes.

Three sets of hypotheses were tested. The first hypothesis asserts that there is no difference in mean scores between the active-inquiry treatment group and the
control group, with the alternative hypothesis being that the active-inquiry treatment group's mean scores are lower than the mean scores for the control group. The second hypothesis declares no difference in mean scores between the vicarious-inquiry treatment group and the control group, with the alternate hypothesis being that the vicarious-inquiry group's mean scores are lower than the mean scores for the control group. A third hypothesis asserts that the mean scores for the active-inquiry treatment group are no different from the mean scores for the vicarious-inquiry treatment group, with the alternative hypothesis being that the mean scores for the two groups are significantly different. A directional test was employed for treatment effect on the cognitive measures for the first two hypotheses. This was justified by the assumption that treatment would not result in a decrease of mean scores for the active-inquiry and vicarious-inquiry groups relative to the control group. In all analyses, the intact treatment groups were appropriately used as the unit of analysis.

For the first criterion measure, knowledge of (abilities in) process skills, analysis of covariance revealed there was a difference in mean scores between treatment groups and the control group, with both treatment groups scoring higher on the process skills assessment than the control group (p<0.05). A positive relationship was indicated between achievement in process skills and science grade point average (p<0.10). Significant differences between the three treatment levels were determined by use of the Newman-Keuls method of multiple comparisons, although no data other than p-values were reported. Both the vicarious and active-inquiry groups scored significantly higher than the control (p<0.05), but the vicarious and active groups were not different from each other. It was also determined that students with
high science grade point average scored significantly higher than those in the low
grade point average group (p<0.01). Thus, preservice teachers in the active-inquiry
group and the vicarious-inquiry group gained better abilities in science process skills
than the students in the non-inquiry control group, as measured by *The Science
Process Measure for Teachers*. In addition, subjects in the two treatment groups did
not differ in achievement of science process skills as measured by this instrument.
The author did not remark on the practical significance of the outcomes or
differences among the groups.

*Analysis of The Science Process Measure for Teachers* subsets determined
that for five of the seven subsets, the active-inquiry and vicarious-inquiry groups
scored higher than the control group. The two subsets shown to be statistically
different for the treatment and control groups were *Classification* and *Using
space/time relationships* (p<0.001). For this measure, it was also determined that the
students in the high science grade point average group scored significantly higher
than those in the low grade point average group. Furthermore, significant differences
(p<0.05) were determined, by use of analysis of covariance and the Newman-Keuls
multiple comparison test, between the high grade point average group and the low
grade point average group and between the medium grade point average group and
the low grade point average group for the subsets of *Inferring* and *Using space/time
relationships*. No difference was found between the high and medium grade point
average groups. Significant interaction effects between treatment and grade point
average were obtained on the subsets of *Using space/time relationships* and
*Observing*. Again, the author did not offer practical implications of these results.
Actual mean scores for the different groups were not provided to allow proper assessment of the practical significance of the results.

An analysis of covariance revealed no significant differences between treatment groups' and the control group's mean scores on the TOUS. Thus, it was appropriately concluded that preservice elementary teachers in the active-inquiry, the vicarious-inquiry, and the control group did not differ in their understanding of NOS as a result of the treatments. Those with the high grade point average did have mean TOUS scores significantly higher than those with low science grade point average (p<0.05). Data for the adjusted cell and marginal means indicated, however, that the high grade point average group scored an average of only 2.8 points higher on the TOUS than the low science grade point average group.

No significant differences were detected due to treatment group or science grade point average in attitude toward science or attitude toward science teaching. Similarly, no differences were detected in the groups' mean scores on the *Attitude Toward Method of Instruction*. No mean scores were provided for further analysis.

Riley (1979), thus, concluded "training in the science process skills by either a vicarious-inquiry or an active-inquiry approach can be employed to improve preservice teachers' competence in selected process skills" (p. 383). These process skills were Classifying and Using space/time relationships. No evidence indicated one inquiry method was superior to the other in any of the criterion measures. No evidence was reported to indicate either inquiry approach was better at enhancing attitudes toward science and science teaching, attitudes toward method of instruction, or understanding of NOS. Undergraduate science grade point average appeared,
Riley appropriately asserted, to be related to acquisition of process skills as well as an understanding of NOS. Implications concerning the adequacy of understanding of NOS between the high and low grade point average groups, or any of the groups, can not be inferred due to the absence of mean TOUS scores for the groups. Not enough information was provided about the population from which the sample was derived to determine the generalizability of these results. Thus, the scope of inference is restricted to the present sample of preservice teachers involved in this study.

The author suggested that such findings imply that a non-inquiry approach to training of elementary teachers in science teaching methods was as effective as an inquiry, or hands-on, approach in developing teachers’ attitudes toward science and science teaching, and an understanding of NOS. Thus, the previous endorsement for the “doing science is necessary and sufficient for developing an understanding of NOS” campaign is not supported by this investigation. This is in contrast to reports by others, such as Barufaldi et al. (1977), who argue that gaining an understanding of NOS is obtainable through engagement in scientific inquiry.

More recently, qualitative investigations into the impact of science inquiry on teachers’ and students’ views of NOS have appeared in the literature. One such study by Shapiro (1996) examined the effects of doing an independent research project on prospective elementary teachers’ views and attitudes about science. Shapiro argued that the only exposure to science that prospective teachers get tends to be content-based, rather than process-based. Shapiro considered a publication by Latour (1987) who referred to the substantive body of knowledge of science as the “face of science that knows.” If students are to understand the complexities of the processes of
science, they need to learn about the “face of science that does not yet know.” Shapiro argued that the learning of science from texts and lectures alone results in an image of NOS lacking human engagement in the process of answering questions. Elementary preservice teachers especially tend to have limited or no exposure to personal scientific inquiry. They often express anxiety in teaching science as a result of such limited experiences. Shapiro indicated a “concern that unless student teachers are assisted in developing confidence and competence in science, the cycle may repeat itself with new teachers avoiding the topic, thereby providing the same limited encounters with science for students that they experienced” (p. 536).

In reviewing the literature, Shapiro (1996) discussed studies that suggest preservice teachers possess inadequate understandings of NOS. The literature suggests a need to focus on developing teachers’ conceptions. However, as Shapiro discussed, research studies have indicated that some approaches to teaching about NOS have little influence on learners’ views. The purpose of Shapiro’s study was to “clarify and characterize the kinds of changes that occur for student teachers concerning their thinking about NOS during participation in independent research projects” (p. 554). A second focus was to examine the implications of any of these changes for teacher education programs.

A case study from a larger, 4-year investigation was presented as an example of the ways in which one prospective elementary teacher’s views of science and scientific investigation changed and developed as a result engaging in an inquiry project. That is, philosophical changes are assessed in association with the acts of posing a simple research question, designing an investigation, performing the
experimentation, and synthesizing the results in a presentable format. A description of the larger study is provided below, followed by the details of the reported case study.

The large research project conducted by this author spanned four years. She collected data from over 210 preservice elementary teachers as they worked on the assignment, “Inviting Investigations” as part of the requirements in an elementary science methods course. The intention of the assignment was to enhance students’ understandings of science and the procedures of investigation in science. Over a 7-week period, students were to pose an original simple problem or question that might be asked by an elementary student. They were to design a procedure to address their question. Assistance in defining their questions was provided to the students by the author and other research assistants. Students were encouraged to consider the importance of defining and controlling variables in their investigations. Students kept detailed journals of all stages in the development of their problem, procedures, and analysis of results. They compiled their results in a display that was shared with class members and children. Student teachers used the displays to examine aspects of research design and implementation.

The first three years of the study utilized a survey and a repertory grid analysis. During the fourth and last year of the project, 38 students in the elementary science methods course were interviewed to examine their experiences with the project. The author described these students as being typical for an elementary science methods course in a North American teacher education program based in a university faculty of education. There were 34 females and 4 males in this cohort, 12
who held undergraduate degrees prior to entering the teacher education program. Four of the student teachers were majoring in science education. No further information on the sample was reported, although detailed information was collected for each participant.

Several data sources were used in this study: (1) A survey to gather information on participants' background, interest, and confidence in teaching science; (2) a statement by each participant explicating his/her definition of science written at the beginning of the course, before the students were introduced to the investigation project; (3) a statement written at the completion of the course indicating whether the initial definition was supported or changed by the investigation experience; (4) repertory grid charts developed at the beginning and again at the end of the investigation project; (5) notes made by the researcher during the course; (6) complete records of students' journals, notes, and reflections on the project; and (7) transcriptions of interviews with student teachers regarding changes they experienced as a result of doing the investigation project. Specific changes indicated by the repertory grid charts were examined. Six researchers conducted the interviews. No details of establishing and verifying consistency among the interviewers were provided.

Personal construct grids were used to document participants' views of the nature of scientific investigations. The personal constructs for the grid were developed from responses elicited during individual interviews with a pilot group of students from the third year of the study. No further details on the development of the constructs were provided.
Fifteen constructs, consisting of common terms or statements describing aspects of scientific investigation, comprised one side of the repertory grid. The provided constructs were:

1. Creating my own ideas/Just following directions
2. Challenging, problematic, troublesome/ Easy, simple
3. Shaping the investigation/ Conducting the investigation
4. Having some idea beforehand about the project outcome/ Having no idea what will result from the study
5. Using the imagination – spontaneous ideas/ Recipe-like prescriptive work
6. Frustrating experience/ Satisfying experience
7. Creating new knowledge/ discovering what exists – the way things are
8. Doing real science/ Doing things unrelated to science
9. Personally meaningful, interesting/ Not particularly meaningful or interesting
10. Rational, logical activity/ Affective – feelings and emotions involved
11. Experience with the phenomena/ Observing objectively
12. Theoretical work/ Practical work
13. Using the “scientific method” to solve the problem/ Not using any particular method
14. Important work in science/ Less important work in science
15. Process oriented/ Product oriented (Shapiro, 1996, p. 543)
The constructs were used to provide descriptive ratings for 12 elements of scientific investigation as described by the pilot group. Each of the 12 elements was rated separately. The 12 elements were:

1. A problem or topic of interest is selected for investigation.
2. The topic is developed into a testable question.
3. Factors and variables which may affect the outcome of the investigation are identified and defined.
4. An idea about how the investigation will turn out is developed.
5. Materials and equipment needed to conduct the investigation are collected.
6. Observations are collected and recorded to answer the problem question.
7. Improvements must be made to the original design of the investigation.
8. An unusual or unexpected result is produced.
9. The investigation does not run smoothly.
10. The results of the investigation are recorded.
11. Conclusions are drawn from the results of the investigation.
12. The findings and conclusions are organized for public presentation.

(Shapiro, 1996, p. 543)

The participants completed the grid, rating each of the 12 elements on each of the 15 personal constructs. The ratings were on a 5-point scale that spanned the opposite poles for each construct. Initially, Shapiro (1996) reported that 21 of the 38 students completed the repertory grids at the beginning of the science methods course and again at its completion. The author offered no explanation for why only
21 out of the 38 students completed both grids. There was no indication that the resultant 21 were representative of the original sample. Although the author indicated only 21 of the student teachers filled out both grids, she indicated evidence for changes in all 38 students. For example, Shapiro (1996) reported that 34 of the 38 participants made one or more complete reversals in ratings of elements on the personal construct charts. Thus, there are inconsistencies in regard to the number of participants whose pre and post-test repertory grids were actually analyzed.

For each participant, the pre and post repertory grids were compared and examined for changes in ratings for each of the elements. Significant, or pronounced, movement in rating of an element on a particular construct was defined by the author as a change of two or more steps in the 5-point scale toward one pole of the construct or the other. No further details of grid analysis were provided.

Following completion of the grid analysis, 21 students were interviewed. Interview discussions focused on the individuals’ identified changes in ratings of elements using the personal constructs. Students were asked to discuss the changes in relation to their individual investigations.

From interview transcripts, changes in thinking were coded by interviewers and categorized into “change themes” concerning the nature of investigations in science. Two graduate students, one research consultant, and the author then assessed and affirmed the categorization of the coded statements. This was an effort to establish inter-rater reliability. The author did not report the established level of agreement. No criteria were defined regarding the classification of statements. Twelve change themes were identified. Changes were seen in (1) ideas about the
nature of the steps and procedures of scientific investigations, (2) thinking about what science is, (3) ideas about the complexity of scientific investigation, (4) thinking about the value of establishing an hypothesis, clarifying variables and factors affecting the outcome of a study, (5) thinking about the importance of thinking logically while moving through the project, (6) thinking about the importance of personal commitment to a project, seeing it through to its conclusion, (7) the importance of a personal participation in the investigation experience, rather than a mindless following of steps, (8) views about self as successful science learner, (9) views about the importance of sharing knowledge with others, (10) ideas about the importance of taking a critical perspective on the knowledge claims of research studies, a realization of the power to manipulate results, (11) an expanded appreciation for finding the unexpected, and (12) ideas about the usefulness of independent investigations as a learning approach in the elementary science classroom.

The case study reported by Shapiro (1996) detailed three of the thematic changes that occurred for one student, Jan, during her independent investigation. Through the presentation of this case study, the author intended to exemplify the quality of changes that occurred for prospective elementary teachers as a result of the investigation project.

Jan was in her final year of the teacher education program. She was average to above average in academic achievement. Jan and her class partner, Laura, had similar background experiences. They were both representative of many of the students in the methods course because of their expressed lack of confidence in their
understanding of science. Neither Jan nor Laura was a science education major. At the inception of the investigation assignment, both Jan and Laura had difficulties designing their project, but they were willing to discuss their problems with the investigator.

Upon analysis of Jan's initial and final repertory grids, the author identified three change themes that described the manner in which Jan's views of the nature of science and science investigations shifted during the project. The first change theme regarded the changes in ideas about the nature of the steps and procedures of investigations in science. Jan and Laura had difficulties selecting a research question. The author stated that most students had similar difficulties. Jan and Laura's initial research question was, "Where do fruits ripen best?" They considered placing six different kinds of fruits in various locations to determine where a fruit would ripen best. After the importance of clarifying the change or variable conditions was stressed in class lessons and through discussions with the investigator, Jan and Laura came to realize the increased variability in their proposal due to use of different fruits in different stages of ripeness. They rephrased their question to, "Where do bananas ripen best?" They also realized that all the bananas would have to be at the same stage of ripeness. Furthermore, after more discussion, they decided they needed to clarify exactly where the bananas would be placed for the study. Again, they rephrased their question to, "Under what conditions do bananas ripen best?" Finally, they needed to define what was meant by "best." They eventually agreed on "best" to be "quickest," which could easily be measured. The final format of their research question was, "Under what conditions do bananas ripen most quickly?" All details
of the development of the research question, as well as all other aspects of the investigation, were documented in the students' journals.

The repertory grids and interview transcripts for Jan were analyzed in the aforementioned manner. The analysis indicated that Jan "made prominent shifts in movement on personal construct charts regarding areas associated with these first stages of the development of their investigation" (Shapiro, 1996, p. 549). Initially, Jan considered "problem selection" to be a "frustrating experience." She shifted, however, to view the process of problem selection as "satisfying." Such a shift was a complete reversal in her thinking. Similarly, she shifted her view of problem selection from one of "doing things unrelated to science" to "doing real science." Finally, she made a shift from "affective - feelings and emotions" involved to viewing the process of problem selection as a "rational, logical activity." In her interview, Jan reflected on the identified changes associated with the element of problem selection. She stated that she initially thought the activity was just something she had to do. However, once she got involved, the experience became quite satisfying. Additionally, she had previously considered science to be very methodical and systematic, but after participating in the project, she came to realize the need for creativity and imagination to solve a problem.

The second change theme identified for Jan was her thinking about "what science is." Initially, she defined science in four statements. First, she viewed science as factual information that had been established through rigorous testing. Secondly, she considered science to be a large part of our lives because of our use of technology. Thirdly, she believed the scientific endeavor would never end. Finally,
she viewed science as the study of both common and uncommon occurrences in everyday life. Jan was asked to review these statements upon completing the investigation. Jan indicated her beliefs about what science is had been altered somewhat. She saw science as more of a process of inquiry, rather than a collection of facts. She was consistent with her view of science being an ongoing study of everyday life occurrences, but she no longer equated science with technology. She indicated that through her experience, she realized science could be complicated by uncontrolled variables that can limit the conclusions.

The third change theme identified for Jan was change concerning ideas about the usefulness of independent investigations as a learning approach in the elementary science classroom. Jan reflected on how her experience might be useful in the classroom. She thought such an investigation would be enjoyable and worthwhile for her students. She developed an appreciation for science as a creative endeavor, by asking her own questions and finding her own answers. Such a process, she explicated, would be a valuable experience for elementary students as well.

One purpose of this present study was to characterize any changes that occurred in prospective teachers' understanding of NOS while conducting an independent research investigation. Shapiro combined themes related to NOS and scientific investigations in this study. Therefore, the change themes identified relate to aspects of NOS as well as inquiry. Shapiro (1996) identified 12 such change themes, three of which were detailed in the case study. Consistent for many of the participants in the study was an increased awareness of the commitments, appreciation, and values associated with the development of scientific knowledge.
Furthermore, students gained personal satisfaction in formulating problems to study, cooperating with a research partner, and developing original ways of communicating their results with others.

Shapiro (1996) described three aspects of the project worth consideration during development of such an inquiry-based program for a science methods course. First, adequate time is required for students to fully define their problem. She recommended four weeks for emphasizing the nature of scientific investigations and problem development. The second important aspect of the project was the potential for integrating several disciplines such as language, mathematics, and creative arts. The third aspect of the project important for teacher education programs, Shapiro suggested, was the development of a display to share findings with peers and children. Presentation of results required careful organization of procedures and findings in order to draw conclusions. Sharing of results appeared to be a motivational factor for the participants. Shapiro concluded that "this project provided the experience, insight, and support needed to show student teachers and, in the future, their own students the excitement and sheer intellectual adventure of turning to 'the face of science that does not yet know'" (p. 557).

Opportunity for reflection was perhaps the most important aspect of this study. Students were asked to reflect on the developments and processes in which they engaged throughout the investigation. Moreover, students were asked to reflect on their views of the nature of scientific inquiry and NOS. Shapiro (1996) noted that through reflective discussions, students were able to isolate specific examples from their own investigations to explicate their views of NOS. Students were also
provided with the constructs and elements associated with scientific investigations and NOS that otherwise might not have been apparent to the students. Thus, students were not just "doing science," but were thinking about how they were "doing science," what it meant to be "doing science," and what science is. This approach offered more explicit attention to NOS, especially through guided reflection, than other studies yet reported in this literature review.

Initial surveys for this study indicated 90% of the participants had no prior experience in scientific investigations. Shapiro (1996) argued, therefore, that such students are not given opportunities to develop an understanding of NOS. Thus, Shapiro assumed participation in inquiry investigations is necessary for developing an adequate understanding of NOS. Many participants indicated such an approach as used in the present study would be useful in the elementary classroom. The assumption is that to use these methods successfully in the classroom, the teachers must have experienced such investigations and reflections themselves. Thus, the results of the present study provide implications for undergraduate teacher education programs. However, not this study, nor any of the studies in this review, directly addresses this assumption. This study also did not address possible differences in outcomes based on the type of project the preservice teachers conducted.

Bianchini and Colburn (2000) explored the use of inquiry to teach NOS to preservice elementary teachers during an inquiry-based general science course. Their purpose was to identify and characterize explicit instructional episodes and discussions related to NOS during inquiry activities. They based their study on the work of Schwab (1962) and other advocates of teaching science as inquiry to
develop conceptions of NOS. Their literature review indicates the difficulties and deficiencies teachers have in approach and views of teaching NOS through inquiry. Instead of focusing solely on student outcomes, Bianchini and Colburn studied interactions that occur during an inquiry-based science course where the teacher purposefully introduces aspects of NOS in the context of the activities. An interesting component of this study was that the two investigators held different perspectives of NOS that they felt should serve as the basis for inquiry instruction. Colburn reportedly expressed greater emphasis on the constitutive values of science, or those values that are more internal to science, such as characteristics of acceptable scientific practice (e.g. science is empirical, creative, and supported through consensus). Bianchini tended to focus on the social and cultural context in which science is practiced. They viewed their perspectives as complementary as well as overlapping. They also saw an opportunity to utilize their differences in analysis of classroom practice.

The course was a general science course for fourth year elementary education majors (n=15) at an urban university in California. The purpose of the course was to (1) "encourage students to like science; (2) to help them better understand what science is and how scientists work, and (3) to foster the abilities to identify, define, and solve problems in discipline-appropriate ways" (p. 183). Colburn served as course instructor. The course comprised inquiry-based lessons in biology, chemistry, and physics, with topics typical of k-6 curricula. There was also science education instruction related to science teaching methods, aspects of NOS, developmental issues, and misconceptions. Three units were chosen for the present study. The
authors reported that the three units were presented in a style following the learning cycle, with students in small groups investigating the various phenomena. Students did not have written instructions or manuals, but they did receive guidance from the instructor on procedures as needed.

Bianchini and Colburn (2000) wanted to examine what aspects of NOS were addressed in the inquiry-based units and in what context such instruction arose. To address these questions, 20 hours of videotape were collected, transcribed, and analyzed based on teacher and students’ words and actions. No further information on data collection was provided.

Both researchers analyzed the data separately. Because of their different roles (Colburn as the teacher-researcher, and Bianchini as researcher), they were able to assume different perspectives with the analysis. Colburn “constructed his case of the inquiry classroom using both implicit and explicit discussions of the nature of science” (p. 185). He identified episodes where he and/or his students explicitly or implicitly addressed aspects of NOS and identified the contexts of such interactions. Even though Colburn reportedly looked for both implicit and explicit teaching of NOS, the authors did not describe either teaching approach or distinguish the two in the examples.

Colburn’s analysis resulted in six “ideas” about NOS that were addressed throughout the three units. These six ideas were those aspects of NOS he intended for his students to learn. The instructional episodes identified showed support of these ideas and the contexts in which they were presented. The ideas were:

1. No right way exists necessarily exists to solve a problem in science
2. Scientists and students of science can seemingly be doing the same thing, yet get different results. The differences are not necessarily because someone did something wrong. Examination of carefully recorded procedures often provide explanations for disparate results. Thus, careful record keeping is important when doing science:

3. Scientific knowledge is created and verified on the basis of evidence, rather than authority.

4. Other than the need for evidence, no single way exists to know definitively whether a conclusion is true.

5. The establishment of scientific truth ultimately includes a social component. Thus, part of science involves scientists persuading each other of the validity of their conclusions.

6. Knowledge is not created simply by collecting data. Knowledge also ultimately depends on how scientists interpret their results. This, in turn, depends on their already existing knowledge and beliefs.

The examples included in the report were of discussions between the teacher and students. The discussions related to specific activities (e.g. a Mystery Powders unit wherein the students performed various tests to determine what “knowns” were included in an unknown “mystery” powder.). Discussions were teacher lead with frequent questioning of students about their procedures, results, and verification of conclusions. For example, one discussion lead students to make comparisons among groups to point out the effects of different procedures on outcomes. After the comparison made apparent that one group had results different from the other
groups, Colburn asked, "Does that mean this group is wrong?" The discussion that followed reportedly:

explored how differences in groups' procedures (each group created and executed its own series of procedures) could produce different results — without any group's procedures or results being dismissed as wrong. Differences in results for Powder A, for example, could be related to the time heated or the closeness of the flame. (p. 189)

Some discussions included explicit comparisons on the classroom activities with the activities of real scientists. Unfortunately, the discussion of the results did not include a clear explanation of Colburn's meaning of "explicit" or "implicit" teaching episodes. Some of the examples he provides contain direct reference to aspects of NOS. These are assumed to be examples of explicit teaching of NOS based on the description provided in the introduction to this literature review. Other examples are not so clear.

Colburn does point out instances of missed opportunities he recognized only during the analysis and reflection on his teaching. In listening to classroom dialogue, he identified occasions where he wasn't clear in his explanations or guidance. He identified areas of weakness where he could have addressed certain NOS aspects. He also identified aspects which were less emphasized than others. For example, he reported that idea #6 was not presented as often as the others. He found that he rarely explored instances of theory-laden observations, and he found that students never brought up this idea on their own. He reports that since analyzing the data, he purposely looks for opportunities to teach about this NOS aspect. He looks for instances when students come to different conclusions based on the same data.
"When I notice this, I discuss with them how observations are theory-laden and perhaps offer an example from science..." (p. 195).

Bianchini, as an outside researcher, looked for explicit discussions of NOS found in classroom transcripts to construct her account. Like Colburn, she did not provide a clear explanation of her meaning of "explicit" either. She attempted to identify contexts in which the teacher and/or students made explicit reference to aspects of NOS. In this way, the authors stated, she attempted to minimize the influence of her own bias that would result by her inferring instances of implicit NOS episodes. However, her own bias regarding the meaning of "explicit" and NOS could not be avoided in such interpretive work. The authors claim their varying perspectives strengthened the study. Therefore, claims associated with minimizing the effects of one’s views of NOS and NOS teaching in such a study are inconsistent and irrelevant with the chosen methodology.

Bianchini divided her findings into two major categories, or contexts for NOS teaching: examination of scientific practices and exploration of scientists’ discourse and concepts. More common were episodes (discussions) related to the nature of scientific practices than related to scientific concepts. Included in the first category were instances of teaching about the following NOS aspects: Creative, tentative, social endeavor (students saw themselves as scientists within their own community), examination of evidence, varying procedures (yet refined and standardized), and the importance of careful documentation and communication. For aspects related to scientific concepts, Bianchini identified instructional episodes where ideas of about scientific language, the idea that scientific phenomena can be
placed on a continuum from concrete to abstract (e.g. physical and chemical properties vs. atomic structure), the idea that scientific concepts may have multiple applications (e.g. density can be used to explain floatation, movement of air masses, and movement of ocean currents), and the idea that scientists do not always achieve consensus on an explanation for a given phenomenon. The examples provided by Bianchini were supposedly “explicit” teaching episodes. However, the dialogue presented was often cut short and only descriptions of extended discussions were included. Bianchini’s analysis did not detail missed opportunities even though her examples provided evidence of times when Colburn could have been more clear or could have extended the discussion to target another NOS aspect. For example, she included the following excerpt as evidence of explicit teaching about how scientists do not always reach consensus on certain topics [the activity involved the study of chemical reactions]:

Marco: How about if you mix water and salt? And then you steam all the water out and you get all the salt at the bottom. [Is that a chemical reaction?]

Alan: You know, I’ve had this discussion with scientists before. I used to think it was an easy question and I don’t anymore. I think you make an excellent point. Technically, they would say, yes, you put the salt in the water and the chemical name for that is to dissolve the salt in the water...[Then you boil off the water] and you end up with salt still. Some would tell you, not a chemical reaction. Others would tell you when you put the salt in water, chemical reaction, that you then reverse. You get rid of the water and the salt comes back again. How’s that for a nice, clear succinct answer? (p. 201)

Based on this example, Colburn didn’t really address reasons for the discrepancy. The explicitness related only to the fact that discrepancies occur. However, discrepancies are a result. What about the reasoning behind that result? This is where
other aspects of NOS relate. For example, discrepancies result from different scientists interpreting data differently, but still consistently within their theoretical frameworks. This is the impact of subjectivity on science. Furthermore, interpretation is a creative act and different puzzles, or explanations, can be put together from the same pieces. As such, data can be interpreted differently, and interpretations can change given changing perspectives so long as conclusions remain consistent with data and are valid within the boundaries of scientific knowledge. Highlighting these aspects (subjectivity, creativity, empirical basis, and tentativeness) would have been appropriate during this discussion. From the excerpt provided, the occurrence of discrepancies in science was reaffirmed explicitly, but the rationale for such discrepancies was not. It is not known if such discussions took place or if either of the researchers recognized the opportunity.

The instances of explicit NOS teaching in Bianchini’s and Colburn’s study (2000) were primarily initiated by the teacher and discussions were lead by the teacher through directed questioning as well as direct explanation to relate NOS topics. They reported that students did not raise issues on their own, but responded to teacher questioning and comments by asking additional questions and offering examples from their own work. The authors state that the teacher plays a pivotal role in the teaching of NOS in such inquiry activities. Both researchers saw the teacher, Alan, as instrumental in raising, discussing, and demonstrating various aspects of the nature of science in this inquiry classroom—either as scripted pieces of his lesson plan or as responses to issues that rose organically from the inquiry activities themselves. (p. 203)

These findings, the authors suggest, offer support to the notion that engaging in inquiry activities does not necessarily lead to better understandings of NOS.
If instructors want students to explicitly explore, debate, and reach consensus on nature of science issues, not only must they offer hands-on inquiries, they must, like Alan, explicitly tell students for what conceptual purposes these activities are to be used and repeatedly engage students in discussions that connect the activities to ideas related to the nature of science. (p. 203)

Bianchini and Colburn use the discussions with students as evidence of student understanding, or lack thereof, of NOS aspects throughout the course. As Colburn initiated discussions, students appeared to lack clear understandings of the NOS aspects. Student comments during discussions as the course progressed indicated students were gaining in their understanding of the aspects addressed. Colburn reported using some of the discussions as formative assessments that guided his instruction later in the course. However, there is no mention of formal assessment of students’ understandings of NOS.

Bianchini and Colburn (2000) identify a strength of their study contained within their different perspectives of NOS and roles in the study. This is a unique aspect of the investigation, and they appropriately identify strengths of both positions. The teacher-as-researcher position was able to better identify contexts of NOS instruction and to critique his own teaching in a reflective manner. The researcher, Bianchini, presented a broader perspective of the NOS instruction based on her views of NOS. Their analyses overlapped in many areas, but their descriptions of NOS teaching in the inquiry-based classroom reflected their perspectives of NOS and roles in the investigation.

Although this study does not provide details of students’ final learning outcomes related to NOS, it does explore one teacher’s approach to teaching NOS through inquiry. The results suggest that if aspects of NOS are going to be made
explicit within the context of the inquiry activities, the teacher needs to take purposeful action to draw students into discussion about relevant aspects. If NOS is a valued learning outcome, it needs to be given priority within instruction. Finally, even with priority, this study provides evidence of the difficulties associated with recognizing opportunities in different inquiry contexts. This study showed that even if teaching about certain aspects of NOS is a goal of a course, and the teacher has knowledge of NOS and is attempting to explicitly attend to NOS instructionally, difficulties still arise. Opportunities are missed. Explanations and guided discussions may not be as clear as planned. The students in this sample showed that even in a course where NOS was a goal, they relied on the teacher to lead them to discussions about NOS aspects in the different contexts. The teacher was needed to make connections from the inquiry activities to NOS. According to Bianchini’s and Colburn’s analysis, the students did not make such connections on their own.

**Summary.** Of the five studies involving science teacher preparation programs, two examined effects of teaching NOS implicitly through inquiry-oriented activities (Barufaldi et al., 1977; Riley, 1979). In these studies, the understanding of NOS was viewed as an *affective* outcome, more like a feeling or attitude than science content. None of the treatment groups from these studies was provided any tools with which to develop a framework for their views of NOS. Barufaldi et al. reported those preservice teachers who participated in the inquiry-based course achieved higher understandings of the tentativeness of science. However, the approach to data analysis was questionable, leaving the results suspect. In contrast, Riley reported no differences in gains of understandings of NOS between students engaged in active-
inquiry, vicarious-inquiry (nonmanipulative), and control groups. Thus, the evidence towards the beneficial effects of implicitly teaching of teachers about NOS through inquiry activities is still rather inconclusive based on these two reports. The other three reports in this section offer some relief to the dilemma.

The studies by Shapiro (1996), Billeh and Hasan (1975), and Bianchini and Colburn (2000) addressed understanding of NOS as a *cognitive* outcome, requiring explicit instruction. Shapiro reported that participation in the inquiry investigations in conjunction with reflections on the relationship between the processes and aspects of NOS appeared to enhance participants' NOS conceptions. Bianchini and Colburn (2000) make a similar suggestion that the teacher plays a pivotal role in directing participants' focus on related aspects of NOS. They reported that discussions about relevant NOS aspects and the student activities were initiated and directed by the teacher, not the students. It should be noted that the inquiry activities in the Shapiro (1996) study were reported to be more open inquiries (with participants asking their own questions and carrying through an investigation on their own) as compared to the guided inquiries in Bianchini and Colburn (2000). Nonetheless, both reported the importance of providing guidance and purposeful opportunities for participants to develop conceptions of NOS within the context of open as well as guided inquiry activities. Billeh and Hasan (1975) provide further evidence that laboratory experiences alone are not sufficient to increase participants' views of NOS. Those participants who did not receive the lecture portion of the program that directly addressed aspects of NOS did not show gains in understanding of NOS on the NOST. However, differences related to teaching area (biology vs. physics,
chemistry, or physical science) were not adequately addressed. The results of Billeh and Hasan indicate the teachers in physical science and chemistry had lower initial scores, and thus showed the most gains. Yet the final scores for all four groups was quite similar overall. Given that the final scores for all groups were low (only about 58%), there is the possibility that the groups vary in their views. Information on how they may vary based on discipline area is not available from that study.

These three studies, along with the findings of Riley (1979), support the notion that explicit instruction on NOS may be required for the desired increase in teachers’ understandings of NOS through engagement in inquiry investigations. These studies suggest a need for teachers to relate appropriate NOS and inquiry aspects within a variety of inquiry contexts, such as investigative methods and content areas. None of the studies explored distinctions in explicit attention based on the contexts of methods or content.

Scientists’ Views of Nature of Science

The increased concern in the status of science and technology in the United States in the 1950s lead to targeting science teachers and scientists as leaders in advancing science education. As repeatedly demonstrated throughout this literature review, the prevailing assumption for many of the inquiry-based programs and courses has been that “doing science” will necessarily enhance conceptions of NOS, regardless of investigative approach or content area. This assumption likely stems, in part, from conflating NOS with scientific inquiry. This view is “If one is ‘doing science,’ one is ‘doing NOS.’” Because of the intuitive appeal of the assumption, the development of many of the instruments aimed to assess individuals’ views of NOS
have been based on the views of scientists, for they are doing real, or authentic, scientific inquiry. Furthermore, investigations involving students or teachers conducting scientific investigations for the purpose of learning NOS may be limited by the lack of authenticity of the inquiries due to constraints of school-based or teacher-education contexts. Does context and authenticity of the inquiry impact one’s views of NOS? It becomes apparent that the views of scientists who participate in authentic development of scientific knowledge within the context of the scientific community should be examined. Five such studies that examined scientists’ views of NOS are included in the present review.

Behnke (1961) surveyed science teachers and scientists to determine their relative views of NOS. The two assumptions that lead to this study were first, that scientists necessarily understand NOS, and second, that the views of science teachers with respect to NOS directly influence the quality of their science teaching. Behnke argued that if science teachers held views of NOS that contrasted those of scientists, there is a problem with the current science teaching as well as science teacher education programs. Furthermore, Behnke aimed to determine if differences existed in these views between physical science teachers and biology teachers, between science teachers from different geographical regions, and between science teachers of varying educational backgrounds and teaching experience.

A 50-item questionnaire was developed to measure the participants’ views of science and science teaching. Response choices were “favoring,” “opposing,” or “neutral.” Statements were developed from consideration of science publications, literature on science and philosophy, the investigator’s personal observations and
discussions, and interviews with specialists. No elaboration as to the credibility of these sources was provided. The questionnaire had four groups of statements that included NOS, science and society, the scientist and society, and the teaching of science. Presumably, in an attempt to establish validity, a pilot test was administered to a group of high school science teachers and participants in a conference on science and religion. Statements considered unsatisfactory were identified. However, this did not ensure validity of the instrument. First, it is not known whether the pilot groups were informed of the intent of the statements or asked for their opinions of the intent of the statements. In addition, the qualifications of these groups as experts on NOS and science teaching are not known. Due to the lack of appropriate validity or reliability measures of the developed instrument, any conclusions from this study are questionable.

The membership listing of the National Science Teachers Association was used to randomly select a stratified sample of physical science (N=600) and biology (N=400) teachers from three geographic regions (Northeast, South, and West-Midwest). The stratification was based on the ratio of physical science and biology teachers among the membership. Those selected were mailed the questionnaire. From the initial random selection, 621 teachers (200 from biology, 421 from physical science) returned the questionnaire. Seventy-two percent of the nonrespondents returned a second mailing of the questionnaire. Chi square analysis of the first and second groups of responses showed no difference in replies between the two groups. Thus the initial respondents were used as the final sample and appropriately considered to be representative of the total sample.
One hundred scientists, from all geographic regions, were chosen from lists of scientists in the Science Teaching Improvement Program of the American Association for the Advancement of Science and lists of scientists considered to be well known in academia and nationally. The selection process was not described in the present report. Of the 100, 52 were in the field of life sciences, and 48 in physical science. Twenty-eight were reported to be actively interested in science education, and 72 were active in research, academic life, or public affairs. Ninety-three held academic positions, and the remaining were from research institutes, foundations, or government positions. No description of these categories was provided nor information as to how the scientists were placed in these groups. No information was provided regarding how representative this group of 100 scientists was to the population of scientists. The author reported a response rate of 70 from the selected 100. There was no description of the extent to which the final sample of 70 was representative of the initial sample. Because the focus of the questionnaire toward beliefs about science and science education and the reliance on volunteer responses, the scientists who chose to respond may not have been representative of scientists in general. As such, no generalizations beyond this particular sample would be appropriate.

Percentages of choice responses were calculated for various groups of participants. The researcher employed \( t \)-tests to determine differences in participant groups based on the standard error of differences between two uncorrelated percentages for each item. The total teaching group was compared to the total scientist group. The teacher group was also divided into five subgroups for
comparison: subject taught, geographical region, type of undergraduate institution attended, ratio of hours in science to hours in education, and length of teaching experience. No specific data or levels of significance were reported.

Because Behnke (1961) chose to determine the percentage of responses for the different groups, Chi square analysis would have been a more appropriate statistical approach to determine differences between groups. Multiple linear regression could have been used to determine the significance of the effects of the teacher sub-categories on responses. The approach the author used involved many multiple comparisons with no mention of adjusting the significance level accordingly to minimize the occurrence of a Type 1 error. He compared the percentages of agreement, disagreement, and neutrality for each item on the questionnaire. This means that 150 individual comparisons were made for each group comparison! As such, for each group comparison, there was a $99.95\%$ chance of committing Type 1 error. Even so, the results reported were presented based more on tendencies rather than on statistical significance of the comparisons. There is no explanation given for not including the statistical results in the report.

Results indicated the teachers and scientists were most different in their responses to statements concerning NOS. There were 14 statements in the NOS category (no examples provided). These statements reportedly concerned the value of team research, the limitations of science, goals of science, the scope of scientific approaches to everyday problems, and the tentativeness of science. Scientists and teachers were reported to be in close agreement on only five statements. Both groups valued the team approach to science, recognized that science had limits, and felt that
scientists brought bias into investigations. Two thirds of the teachers and 75% of the scientists thought that man's place in the universe could only be established through scientific procedures. The areas of most disagreement dealt with the goals and limitations of science. For the scientists, 20% believed the goal of science was for the improvement of human welfare or solving practical problems, whereas 50% of the teachers held this view. A large majority of the teachers and only 14% of the scientists felt that religion would interfere with the work of a scientist. Regarding views of the tentative nature of scientific knowledge, over 50% of the teachers and 20% of the scientists felt that scientific findings were absolute. Furthermore, a large majority of scientists and 50% of the teachers stressed a belief that scientific work requires high intelligence.

Behnke (1961) suggested that the differences in views of NOS were related to a lack of understanding of NOS by both the scientists and the teachers. Neither group held views entirely consistent with the views accepted at the time of the study. For example, Behnke stated that the accepted view was that the goal of science is to increase the dependable knowledge about man and the universe, and its relations to the practical and to human welfare. Although the acceptability of the views was not a direct focus of the study, this study did not support the notion that scientists automatically understand NOS the way philosophers, historians, or even science educators promote.

On statements concerning science and society, the greatest difference was for statements referring to the impact of science on society. Twenty-five percent of the teachers and 50% of the scientists thought that man's religious doctrines should be
continually revised in the light of new scientific truths. Little difference was found in views of the role of scientists in society. Both groups saw scientists as important for creating public awareness of science and formulating public policies on scientific and technological matters. On the assertion that scientists should conform to social pressures, 80% of the scientists and 4% of the teachers agreed with the statement. However, only 30% of both groups felt scientists' work should be influenced by a governmental request to do something for the public welfare. Thus, there were apparent inconsistencies in responses.

In regard to the teaching of science, the majority of teachers and a smaller percentage of the scientists felt that science should not be taught as strict cause and effect relationships. The teachers tended to approve of including moral issues in teaching more than did the scientists. The largest difference in this category between the teachers and scientists was in their views of special attention for exceptional students and the competence of science teachers who reject evolutionary theory. A majority of scientists, but very few of the teachers, felt that special provisions should be made for students with exceptional scientific talents. A majority of scientists and only one third of the teachers thought that teachers who rejected evolutionary theory were unqualified to teach science. Behnke (1961) claimed that biology could not be taught as a science without the theories of evolution. Furthermore, Behnke proposed that many of the teachers may themselves reject evolution. Overlooked by the author was the indication that the teachers and many of the scientists do not have an understanding of the development and role of theories in science.
The intragroup comparisons of the responses of the science teachers revealed teachers within and among the subgroups were typically more similar than different in their views. Furthermore, the differences between the scientists and the teachers as a whole group were greater than the differences among the teachers in the various subgroups. The most effective category in discriminating among the teacher subgroups was NOS. Some differences were found in views of NOS in all comparisons of subgroups except those subgroups based on teaching experience.

Depending on the category of the statements, there were some differences depending on geographic region, subject taught, ratio of science to education courses, undergraduate institution attended, and teaching experience. Teachers from the Northeast tended to be in most agreement with the scientists. Teachers from the South differed most from the scientists in their belief that a teacher who rejected evolutionary theory should not teach biology. Furthermore, teachers from the South were more likely to consider the goals of science to be for practical benefits and human welfare. The subgroup with the highest science to education course ratio was, not surprisingly, in highest agreement with the scientists. Little difference was found among the subgroups based on undergraduate institution.

The subgroup with the most teaching experience (over 5 years) differed from the subgroup with the least teaching experience (less than 5 years) on the statements concerning the teaching of evolution. The higher teaching experience group tended to believe that a teacher who rejects evolutionary theory is not qualified to teach biology, scientific work takes superior intelligence, teamwork is more productive, a scientific approach can help solve personal problems, and a scientific approach
should be applied to questions of value. The teachers with the least experience tended to disagree more with statements concerning ill effects of science fiction and teaching science as strict causal relations, but agree more with introducing moral issues in science. For these two subgroups, the least discriminating statements were those dealing with NOS.

Behnke (1961) suggested several implications for teacher education based on the results of this study. First, “prospective science teachers should have the opportunity to acquire a comprehensive understanding of the nature of the scientific endeavor, including its goals, limitations, and relation to other ways of thinking and working” (p. 205). Behnke suggested including studies in the history and philosophy of science in both science content courses and specific seminars to address the inadequacies of both teachers’ and scientists’ views. In addition, the author argued for opportunities for prospective teachers to “study modern aspects of science, including the most recent developments in theory or conceptual schemes” (p. 206). Teachers should be encouraged to critically examine the significance of theories in science, and they should explore cultural influences of science. Finally, teachers should be trained in and encouraged to conduct research on their own science teaching. Overall Behnke claimed both teachers and scientists hold inadequate views of NOS. He stated, “Knowing a great deal about science is no guarantee that he will know what science is about” (p. 207).

Schmidt (1967) reported a comparison between teachers’, students’, and scientists’ views of NOS as measured by the Test On Understanding Science (TOUS). This study was in response to an earlier study by Miller (1963) that
reported many of the secondary teachers in the earlier study demonstrated less of an understanding of NOS than their students in the 11th and 12th grades. Schmidt raised a question as to the value of the TOUS and questioned whether it really measures “understanding of science as scientists themselves perceive the scientific enterprise” (p. 365). This study aimed to replicate Miller’s study, with the addition of examining views of scientists for comparison with teachers.

Very few methodology details were included in the present report. The author mentioned another investigator for this study, suggesting an additional reference, but did not provide a formal reference for a separate source of publication. A search of the publication databases did not reveal an additional publication for this study. The limited known details of this investigation should be considered when interpreting the implications of this report.

Students in grades 7 and 9, and high-ability students in grades 11 and 12 were the student participants. The teachers were only described as in-service teachers. The author did not elaborate as to the number of participants, their source, method of selection, background variables, or any further details about the teacher or student sample. No information was provided concerning test administration. There were 116 scientists from the state of Iowa that were administered the TOUS in an undescribed manner. In addition, 29 college seniors taking a secondary science methods course and 43 college sophomores from an elementary science survey course were administered the TOUS. Again, no further elaboration of the sample or test administration was provided.
Consistent with the study by Miller (1963), Schmidt (1967) found that 14% of the grade 9 students and 47% of the high-ability grade 11 and 12 students scored higher than 25% of the teachers on the TOUS. Furthermore, 5% of the grade 9 students and 9% of the high-ability grade 11 and 12 students scored higher than 50% of the teachers on the TOUS. It is assumed that these comparisons were based on individual student and teacher scores; however, only the mean TOUS scores are reported. The mean scores were 24.9 (range of 10 to 40), 34.0 (range of 12 to 48), 41.0 (range of 23 to 52), and 45.5 points (range of 33 to 55) for the 7th graders, 9th graders, high-ability 11th and 12th graders, and teachers, respectively. No statistical tests were reported to measure differences in mean TOUS scores between any of the groups. Nonetheless, Schmidt did claim that significant gains in understanding of NOS were made as students progressed through the science education sequence. This general claim is well beyond the scope of the reported study.

The mean TOUS score for the scientists in this study was 50.8 with a range of 36 to 59 points. The secondary science methods students had a mean score of 48.0 (range of 37 to 58) and the elementary science survey students scored 40.5 (range of 29 to 50). Schmidt (1967) suggested that the scientists' score of 50 out of 60 possible points (83%) may indicate an acceptable score on the TOUS is less than may have been previously thought. He did not offer an "acceptable" score based on these results, nor did he indicate previous claims of teachers holding inadequate views were necessarily invalidated by this study. He did suggest that to help in-service teachers develop better understandings of NOS, they should have opportunities to
work with scientists rather than simply taking more science courses. Lacking from the description of scientists’ views is any variation according to research area.

Perhaps the most interesting finding from this study, and one which was overlooked by Schmidt (1967), is that the scientists, college students, in-service teachers, and high-ability grade 11 and 12 students had some degree of overlap in their range of scores. Thus, some high school students, teachers, and college students seemed to have better understandings of NOS than some of the scientists in the sample, as determined by the TOUS. It would seem that even though the scientists are viewed as the group who should set the standard for what constitutes an “adequate” score on the TOUS, scientists do not necessarily have “adequate” understandings of NOS, or even more acceptable understandings than people who are not active participants in the scientific enterprise.

Another study that compared views of NOS held by scientists and science teachers was conducted by Kimball (1967-68). Kimball argued that previous studies that determined teachers lack adequate understandings of NOS used nonrepresentative samples of teachers. These samples comprised teachers assigned to teach science regardless of their qualifications. As such, Kimball claimed these samples could not necessarily be expected to possess adequate views of NOS because they did not all have the training to be qualified science teachers. Studies that did not control for teacher qualifications in science were not informative for the science teacher education programs in terms of engendering adequate conceptions of NOS in the science teachers. Kimball stated that NOS views of qualified science teachers had not been examined.
The aim of the study by Kimball (1967-68) was to assess qualified science teachers’ understandings of NOS in comparison to scientists’ views. The goal was to compare these views to determine the source of adequate conceptions of NOS through training or experience. In doing so, teacher education programs and science programs could be informed about ways to improve teachers’ conceptions of NOS.

Four questions were addressed in this study: (1) Do qualified science teachers express the same view of NOS as do practicing scientists of similar academic background? (2) Do views of the scientists change as they become farther removed in time from their undergraduate training? (3) Do the teachers display a different understanding of NOS as a function of their year of graduation? (4) Do philosophy majors, as a result of their explicit study of the structure of knowledge, understand NOS differently from science majors?

A model of NOS was developed in accordance with the literature on the nature and philosophy of science. Kimball’s model comprised eight assertions about NOS: (1) Curiosity drives science. (2) Science is a process-oriented, dynamic, ongoing activity. (3) Science aims at increasing comprehensiveness and simplicity. (4) There is not one “scientific method.” (5) Science relies on the senses, operational definitions, recognition of arbitrary definitions of organization, and the evaluation of scientific work by reproducibility and usefulness. (6) The physical universe is susceptible to human ordering and understanding. (7) Science is characterized by openness of mind and realm of investigation. (8) All science is tentative. In the report each of these assertions included descriptions and reference to philosophy of science literature.
A 29-statement, 3-point Likert scale instrument, the Nature of Science Scale (NOSS), was developed from an initial pool of 200 statements based on the NOS assertions. These statements were analyzed for content validity by a team of experienced science teachers, school science supervisors, science professors, and professors of science education. No clarification regarding these people's expertise in or knowledge about philosophy of science was provided. Statements chosen by this panel were arranged in a 3-point Likert scale instrument and administered to 54 college graduates, 32 of whom were science majors. Thirty-one items were then chosen based on the responses of the pilot group. Chosen items were those with few neutral responses, in agreement with the NOS model as determined by seven of the nine members of the initial team, and discriminated in favor of the science graduates. The 31-item form of the instrument was administered to 97 college graduates, half of whom were science majors. A split-half reliability was determined and corrected by the Spearman-Brown formula to be 0.72. Two items were eliminated based on poor discrimination. The resultant 29-item test, the NOSS, was the final instrument used in the study. Responses in agreement with the model were counted as 2 points. Neutral responses counted as 1 point, and no points were given for the opposite response.

The sample for this study consisted of 712 graduates who had majored in science (N=625) or philosophy (N=87). This sample represented 74% of the initial requested population who returned the original mailing. No further information regarding administration of the test was provided. The nonrespondents were polled and no significant difference (α=0.10) was determined between the respondents and
nonrespondents. The corrected split-half reliability of the NOSS for the 712 subjects was 0.54. The lower reliability measure was reportedly due to the greater homogeneity of scores of the test population compared to the instrument development group. Of the 625 science majors, 128 were discarded because they reported not being scientists or science teachers.

Group comparisons were made by testing the differences between the mean scores of the groups by t-tests. Kimball (1967-68) stated that because none of the results were significant at the 0.10 level, the use of t-tests was appropriate. Analysis of variance would have been a better statistical procedure for use in this study. Chi-square was then employed to compare groups per item response.

To compare mean NOSS scores for science majors who had become scientists and those who had become science teachers, eight group comparisons were made. Four comparisons were made based on year of graduation. Two comparisons were based on university attended (Stanford scientists vs. Stanford teachers and San Jose State scientists vs. San Jose State teachers). The mean NOSS score for all scientists in the sample was compared with the mean score for all teachers. Finally, nondoctorate scientists were compared with nondoctorate teachers. None of these comparisons showed significant differences between the mean NOSS scores ($\alpha=0.05$). Thus, it was concluded that there was no significant difference in understanding of NOS between science majors who became scientists and those who became teachers. Interestingly, and perhaps most importantly, the average score for all groups of science majors was around 35 out of a possible 59 points. The views of
NOS held by the science majors in this sample were deemed *inadequate*, regardless of occupation.

To test for differences in understanding of NOS as a function of time after graduation, analysis of variance (α=0.05) was utilized in a comparison among three groups of science majors who were or intended to become teachers. Similarly, three groups of science majors who were or intended to become scientists were compared. Groupings were established based on year of graduation, from newly graduated up to 12 years after graduation. No significant differences were found based on year of graduation. It was concluded that neither experience nor time since graduation was significantly related to these scientists’ or teachers’ views of NOS.

In a third set of analyses, mean NOSS scores for philosophy majors were compared to mean scores for all the science majors (both scientists and teachers combined). The mean score for philosophy majors was significantly higher than the mean score for science majors, suggesting the philosophy majors in this sample have more adequate views of NOS than the science majors, as determined by the NOSS. Further analysis determined the significant differences to lie in the mean scores between the philosophy majors and the scientists only. This difference was determined irrespective of university attended. For these analyses, no data were provided that would enable an estimate of the practical significance of the difference in NOSS scores. Because scores of the philosophy majors were not reported, it is not possible to determine if their views of NOS would be considered *adequate* or just *better* than the scientists’. In any event, the philosophy majors outscored the scientists on the NOSS.
Chi-square analysis of frequency of correct responses to individual items revealed no differences between scientists and science teachers on individual item responses. However, the philosophy majors differed from science majors with respect to understanding the methodology of science. This was the only area where philosophy majors were significantly more in agreement with the model. It seems, therefore, that even though scientists perform the processes of scientific investigations, they do not necessarily understand the nature of the inquiries in which they are engaged. Kimball (1967-68) suggested there is thus a need for a philosophy of science component in the undergraduate science curriculum. Furthermore, because philosophy majors do not typically have scientific research experience, the results suggest that active engagement in scientific inquiry is not essential for developing an understanding NOS.

Taking this argument a step further, this report presents evidence that adequate understanding of NOS is not an automatic result of the undergraduate science curriculum, and experience in the field as a scientist or a teacher does not change one’s conceptions. “Adequate” in this study was based on the NOS model Kimbell constructed from the philosophical stance of the time. Kimball (1967-68) argued that the undergraduate science curriculum is in need of revision to improve science students’ views. Consistent with the results of this study, he suggested instruction in the philosophy of science may increase understanding of NOS of science majors. He recommended further study into this area.

In addition, because views appear to be established by the completion of formal education, Kimball suggested teacher preparation programs may improve
teachers' views of NOS by offering additional instruction to achieve this goal. He recommended further investigation into specific ways to improve scientists' and teachers' views of NOS.

A limitation of this study, which was appropriately recognized by the author, is its lack of generalizability beyond the participants of this study. Further investigations of scientists' and teachers' views of NOS and the sources of those views are recommended. Furthermore, the potential bias of the instrument in favor of views held by scientists should be considered, although the results of this study were apparently not affected by such a bias.

In a more recent study, Pomeroy (1993) examined and compared the views of NOS among scientists, secondary science teachers, and elementary teachers. The interest in a possible relationship between the philosophy and history of science to science teaching stemmed from the literature reporting conflating views of the philosophy of science held by scientists and science educators. Describing the relative views of these groups, Pomeroy argued, held possible implications toward informing science teacher education. Specifically, the study aimed to identify differences between how scientists and teachers view NOS, scientific method, and related aspects of science education. Correlations were sought between beliefs about science and beliefs about science education. In addition, the author examined the views of logicoempiricism held by scientists and science teachers.

Pomeroy (1993) identified the logicoempiricist view of science as the "traditional" view, or the belief that scientific knowledge progresses only through inductive methods based on observation and controlled experimentation. This picture
of science, most often presented in science textbooks and science courses, suggests science is objective, based on logic alone and is removed from human interpretation.

In contrast, Pomeroy (1993) described the nontraditionalist view as that which recognizes that "dream, intuition, play and great inexplicable leaps as potentially part of scientific method" (pg. 262) are a part of science. This view also recognizes that objectivity is unobtainable.

It was this traditional/nontraditional dichotomy that led Pomeroy (1993) to question the source of her own nontraditionalist view of science. She raised the question of whether her nontraditionalist views evolved from experiences as a research scientist or as a science educator. In an effort to identify possible sources of one's views, in this study she examined the philosophies of NOS, scientific method, and science education held by scientists and science educators.

The survey developed for this study was a 50-item, agree-disagree, 5-1 Likert scale instrument. Items on the survey ranged from statements congruent with traditional views of science and science teaching to the nontraditional views. The items considered "roles of deduction, art, perception, attitude, judgment, community, and prior belief in shaping the work of the scientists and their knowledge of nature" (pg. 263). Also included were statements targeted toward beliefs of science education, K-12, such as the role of laboratory experiences, process vs. content, depth vs. breadth, and mastery vs. exploration. Finally, the survey requested demographic and biographic information. She reported that the instrument had not been validated or subjected to reliability tests that would be appropriate for more general use of the instrument. She recognized the limitations and potential bias
imposed by the length and philosophical nature of the survey in the resulting self-selected sample.

The sample was one of volunteers who filled out the survey by written request. The sampling region was narrow and limited to research scientists and secondary and elementary educators in Alaska. To the author's credit, the nonrandom sampling was recognized as a source of possible bias. Those who responded may have been more informed and interested in the philosophical and educational focus of the study. As all respondents would most likely be equally biased, scientists as well as science educators, Pomeroy (1993) argued in favor of their comparability. No details regarding the rate of return of the written requests was provided.

Seventy-one scientists participated in the study with an average age of 44.5 and a male to female ratio of 3:1. The disciplines of the scientists were 45% physical science, 49% biological or environmental science, and 6% social science. Thirty-four percent had no background in the history or philosophy of science; 7% reported having taken such a course; 43% reported having engaged in independent reading on the subject; and 16% reported having taken a history or philosophy of science course as well as engaging in independent reading. The author reported finding no significant differences between male and female scientists or between research specialties with respect to any of the background variables. No further description was provided.

There were 109 teacher respondents, with an average age of 37.3 and a female to male ratio of 2:1. Thirty-three percent of the teacher respondents were
primary teachers, 22% intermediate, 19% junior high, and 26% high school. For this study, the primary and intermediate teachers comprised one group and the junior high and high school teachers comprised the other teacher group. Of the entire teacher sample, two thirds reported no scientific research experience. One fourth of the sample were members of the National Science Teachers Association (NSTA). A significant number of more men were secondary teachers, had previous scientific research experience, had more science course education, and were members of NSTA. These differences, however, were not gender-based, according to multiple linear regression analysis. Rather, the differences could be explained by the greater number of male secondary science teachers in the sample. To this effect, the correlations could be based on the fact that the secondary science teachers would likely have more science experience.

Categories of statements were developed based on underlying themes identified within the statements. Statements were then analyzed as sets and described by summary variables. Two methods of categorization were employed, and the results then compared. The first approach was based on statistical correlations using a correlation matrix. Cronbach’s $\alpha$ greater than 0.5 indicated internal consistency. Six clusters of statements resulted from this initial method. The second method of categorizing statements was based on original intentions of being traditional or nontraditional statements. These two groupings were then compared and examined for internal consistency of the clusters generated. Three main clusters resulted from the combined groupings. The statements were grouped according to (1) a traditional view of science, (2) a traditional view of science education, and (3) a nontraditional
view of science. Items that did not contribute to the overall internal consistency were omitted from the analysis. The mean, variance, and distribution were calculated for each of the clusters, creating three summary variables.

Pomeroy (1993) conducted nine t-tests to compare responses of scientists to teachers, men to women, and secondary science teachers to elementary teachers. Reported in the present study were the results with respect to the clustered themes (traditional views of science, traditional views of science education, and nontraditional views of science).

Of the 50 statements on the survey, 8 were congruent with the traditional views of science. The internal consistency reported for the responses on these statements was moderate at 0.651, with an overall mean response of 3.02. For this cluster, the scientists scored significantly higher than all the teachers (p=0.022, difference in means of 0.22). The males showed a tendency toward more traditional views than females (difference in means of 0.16), and secondary teachers were not significantly different from elementary teachers (difference in means of 0.32).

There were 14 statements in the traditional views of science education cluster. These were statements that opposed those views held in the current national reform efforts for science education. Generally, the traditional view of science education emphasized the textbook and lecture format to transmit science content. The Cronbach's α for this cluster was strong at 0.809, with an overall mean score of 2.37. The scientists scored significantly higher than all the teachers (p<0.000; difference in means of 0.55). The males scored significantly higher on this cluster than the females (p<0.000; difference in means of 0.36). The secondary teachers
scored significantly higher than the elementary teachers (p<0.000; difference in means of 0.38). These results were consistent with those reported for the traditional views of science cluster.

There were nine statements in the cluster representing the nontraditional view of science. These statements had a rather low internal consistency, as measured by Cronbach’s α, of 0.592. The overall mean score of 4.03 indicated general agreement with these statements. The women in the sample tended to score higher than the men (p=0.074; difference in means of 0.14), and the elementary teachers scored significantly higher than the secondary teachers (p=0.009; means not given). There were no significant differences found between the teachers and the scientists for this cluster (means not given).

A determination of the correlative relationships between the clusters revealed a significant (p=0.0008) negative correlation (R=-0.2485) between the traditional science and nontraditional science cluster. This correlation, although weak, was expected. However, the authors do not comment on the low correlation value. A slightly stronger correlation (R=0.418) was found between traditional science views and traditional science education views. This correlation was also significant (p=0.0001).

The initial question regarding the differences in views held by scientists and science educators was addressed by these findings. The differences in responses based on gender and the relative responses of the sample groups were the focus of the discussion in the present report. First, the data revealed the men tend to hold more traditional views of science and science teaching relative to the women in the
sample. However, Pomeroy (1993) reported that multiple linear regression analysis provided evidence to suggest this finding was a result of most of the men in the teacher sample being secondary science teachers. However, there were no details of the analysis provided in the report to estimate the strength of the results. The data provided in the report indicate both men and women tended toward traditional views of science and science teaching. The difference in the mean scores for the two clusters suggested the differences between views of men and women were more traditional views of science and science education were most strongly held by scientists, followed by secondary science teachers, and then by the elementary teachers. Pomeroy (1993) discussed the consistency of this finding with other reports of the scientists and science teachers holding positivistic views of science. She suggested these views possibly arise out of the indoctrination of these groups into the norms of the scientific community. Additionally, Pomeroy suggested these views could arise from the training of the scientists and secondary science teachers that was more consistent with Kuhn’s “normal” science. “Normal” science, Pomeroy argued, is more representative of a positivistic form of scientific method than is revolutionary science. While scientists might acknowledge the subjective and creative nature of science in their work, they are required to report their findings in an “impersonal, empirically justified, reconstructed logic of traditional scientific method” (pg. 270). Furthermore, because secondary science education programs emphasize science content, science educators are not provided opportunity to reflect on the process or philosophy of knowledge development, and they rely on their science textbooks in their classrooms. To this effect, the relative lack of traditional
views held by the elementary science teachers might be due to their having less science experience. However, the author aptly indicated that such a claim is beyond the scope of this study.

An encouraging result from this study was the overall general high agreement with the nontraditional cluster by the entire sample (mean 4.03). This suggested the sample in this study held more nontraditional views of science. A substantial limitation due to the sampling method for this study was recognized and discussed by the author. The potential bias due to the self-selection of the participants could have led to the result of the sample, as a whole, holding more nontraditional views of science. Those who responded might have been more informed and interested in the philosophical aspects of science.

Pomeroy (1993) claimed that an implication toward science education resulting from this study is the importance of philosophy of science in science education. However, this study did not report a separate analysis of those participants who had backgrounds in the philosophy of science. Hence, the stated implication is not directly supported by the reported findings.

Pomeroy (1993) made several suggestions for the source of the reported views, even though the results of this particular study do not directly support her claims. First, the teachers in the sample might base their views of science on their understanding of constructivism. “An explanation of elementary teachers’ views of science that rests on their views of children’s construction of knowledge echoes Piaget’s notion of genetic epistemology, which proposes that evolution of science is recapitulated in cognitive development” (Pomeroy, 1993, p. 272). Pomeroy
suggested further studies involving teacher interviews and observations would help inform this claim. Second, the teachers' views of NOS could result from the construction of their own teaching knowledge from experiences and observations. If so, teacher education programs should focus on the processes of teachers' construction of their knowledge about children and their own practice of teaching. By making the connection between their own science learning and the processes of science, Pomeroy argued, teachers may develop beliefs about NOS that are more currently accepted and less foreign to them.

A third possible explanation for the views held by the teachers in this study considers the effects of possible in-service participation. It is suggested that the teachers may have participated in programs that introduced the constructivist, nontraditional nature of science. All stated possibilities were recommended emphases for further study.

Perhaps the most obvious limitation of this study is the lack of validity and reliability measures for the instrument. Pomeroy (1993) did report not taking these measures, and thusly did not make conclusive claims from her findings. The implications for teacher education were appropriately prefaced by the admission that they were not specifically explored in the reported study.

More recently, Glasson and Bentley (2000) investigated scientists' views of NOS and the views they portray through presentations of their research to science teachers. Their primary focus was to characterize how the scientists communicate the nature and relevance of their research to teachers. Comparisons were made between scientists' views of NOS, portrayal of NOS through their presentations, and the
positions adopted by national reform documents on NOS and STS for K-12 science education. The authors described NOS as presented in the NSES as comprising components of 1) the scientific world view, 2) scientific methods of inquiry, and 3) the nature of the scientific enterprise. They also provide descriptions of aspects of NOS promoted among philosophers and scholars. These aspects included the idea that science is “a theory-driven and empirical enterprise in which the historical development of science is characterized by periods of consensus and dissensus” (p. 471). Other aspects included science as dynamic, parsimonious, tentative, testable, comprising multiple methods, relying on peer review and community authority.

Glasson and Bentley (2000) also presented current goals for learning about STS issues. These included knowledge the relationship between the practice of science and society, and the interrelationships between science, technology, and society in the context of social issues.

The purpose for studying NOS as portrayed by scientists through presentations stems from the authors’ suggestion that teachers might attain a better understanding of the sociocultural view of NOS if they experience “science in the making” as opposed to “ready made science” as is typically presented in textbooks. “Science in the making” is that science which is under development and debate within the scientific community.

Six scientists served as the subjects in the present study. They were participants in a state conference for educators. The conference included presentations and workshops conducted by research scientists that attempted to connect scientific research to the K-12 classroom. The total number of scientists who
participated at the conference is not mentioned. The authors do not describe the manner of selection of the six participants. The sample consisted of four research scientists (chemist, botanist, ecologist, biologist) and two engineers (transportation research, materials science).

Data included field notes and handouts from scientists’ presentations and interviews of each of the six scientists. Presentation data were analyzed by comparing statements to categories of NOS as defined in the NSES as well as to categories relevant to the nature of science, technology, and society. These categories included 1) methods of inquiry, 2) sociopolitical connections, 3) technological connections, 4) historical connections, and 5) underlying values and assumptions. Authors searched the data for evidence of consistency and inconsistency with the positions advocated in reform documents relative to these five categories. Interviews were conducted to describe scientists’ views of NOS and STS in relation to their area of research. The authors asked scientists to describe how their own work reflected the NOS, as they understood it, and also to identify aspects of NOS that they felt were communicated to the teachers at the conference presentation. No further information is provided about the interview protocol.

Interview data were categorized into themes. The authors do not state that the themes were similar or not to those used in the analysis of the presentation data. The authors compared scientists’ views attained from interview data with their actions attained from the presentation data. They also compared the actions with the reported meaning of the actions (interview data). The authors do not describe their efforts to reduce researcher bias that could result from such dual analyses. They did report
efforts to find discrepancies between scientists stated views in the interview with the view portrayed through the presentations.

Results for two of the scientists were presented in detail, and the other four summarized. The chemist was involved in research on tropical plants for potential drug use. His presentation contained numerous examples to illustrate the connection of his research to society. He explained the importance of his negotiations with South American tribes in accumulating plant species for his studies. The focus of the presentation was on pharmaceutical application of the research. Problems of deforestation were raised as well as developments in social, political, and ecological issues within the South American countries. The interview with the chemist revealed what first appeared to be a different perspective of NOS and related societal issues. The scientist focused on his research having foundation in empirical data to study cause and effect relationships. He emphasized "objective truth" in scientific and ethical areas that could be achieved through experimentation. He did discuss the importance of teaching science as it relates to everyday life. The authors state that this scientist's views and presentations reflect a strong focus on connections between science and society.

The engineer in material science served as the other case study. The authors chose this engineer because of the contrast in research focus as compared to the chemist. He focused his presentation on ways teachers could incorporate materials into their instruction to teach various concepts. He also emphasized recycling and consumer land/energy use issues. Glasson and Bentley (2000) report that this
engineer emphasized problem solving, peer review, and community authority during the interview.

Overall, the authors report few discrepancies between NOS as portrayed by the scientists in interviews and through their presentations. Although they tended to impart societal impacts, they also implied their work to be objective and based on value-free assumptions. Differences did exist between the scientists and engineers. The scientists from so-called traditional disciplines (chemistry, biology, and botany) tended to view their science as empirical and based on experimental design. The ecologist emphasized the use of statistical analysis in his work. The two engineers focused more on problem solving through multiple pathways and the development of materials. The authors state that none of the scientists could necessarily be categorized as positivist or post-positivist based on this study. The use of terminology became an issue of interpretation. For example, the chemist described his work as the search for objective truth through experimentation. The authors recognize that such a view could be determined positivistic and inconsistent with the chemist’s presentation of his work as involving negotiations with drug companies and tribal communities. However, Glasson and Bentley suggest that the chemist uses the word “objective” as a part of the process of scientific research that results from empirical study, not from negotiating such results. They do not provide evidence to support their interpretation. They add that “objectivity” may be seen by members of the scientific community to relate to the strength of support within the community for a particular concept. This argument is faulty based on the data presented. At best, the authors can claim terminology use and meaning needs to be considered when
comparing views of NOS. Others have emphasized the importance of asking
participants to elaborate on their meaning of certain terms such as “proof” and
“experiment” (e.g. Lederman & O’Malley, 1990; Schwartz, Lederman, &
Thompson, 2001).

Interestingly, Glasson and Bentley (2000) acknowledge possible disparity
between epistemological assumptions of the scientist and those of people outside of
the scientific community.

[From within, the] epistemological assumptions may require no explication
because theory, methods, and aims in their fields are closely interrelated and
developed within the context of the scientific community. From the
epistemological perspective of an outsider, however, these assumptions may
be considered subjective and value-laden. (p. 481)

Although the NOS views of the teachers who heard the scientists’ presentations were
not assessed, the authors rightly suggest that their views would not necessarily be
enhanced by their participation. They suggest a need to “facilitate the
epistemological development of teachers if teachers are going to know how to think
about what they hear about ‘cutting-edge’ scientific research” (p. 483). Just as the
scientist may not recognize elements of the nature of the scientific discipline in
which he is engaged, an outsider won’t necessarily make appropriate connections
without a framework from which to interpret the research.

Summary. Within the limitations of these reports, the five studies reviewed
here indicate scientists do not necessarily hold views of NOS in congruence with
currently accepted views. Little is understood about how views relate to research
specialty. Only Glasson and Bently (2000) made direct comparisons. The engineers
and scientists in their study held somewhat different views of inquiry, and the
ecologist differed from the other scientists in terms of use of statistics. The small
group limits the generalizability of the results. However, their results suggest that
scientist who participate in different research methods may possess varying views of
NOS and inquiry.

The study by Behnke (1961) had numerous problems with instrument validity
and data analysis, but the report of the trends of the scientists and their comparison to
the teachers were supported by the other reports in this section. That is, neither the
scientists nor the teachers in these studies were consistently different in conceptions
of NOS. In addition, neither group consistently held accepted conceptions of NOS
based on current stance advocated in science education. The report by Schmidt
(1967) suggested the scientists’ scores on the TOUS were lower than expected, and
as such, acceptable scores on the TOUS may be lower than previously thought.
Schmidt suggested teachers be provided opportunities to work with scientists in
order to develop their understandings of NOS. However, Schmidt failed to remark on
the overlap of TOUS scores for all the participant groups in the study. The
implication of the overlap is that all the scientists did not have better understandings
of NOS compared to the high school students, college students, and teachers in the
sample. Thus, not all of those who “do science” understand NOS better than all of
those who do not “do science.” The nature of any differences between groups of
participants, however subtle, was not explored in either of these studies. The findings
of Behnke and Schmidt are consistent with those of Kimball (1967-68) and Pomeroy
(1993). Moreover, Pomeroy’s findings revealed the teachers with the least
experience in science, the elementary teachers, held more nontraditional views of
NOS. Included in the reasons for her findings was the nontraditional views held by those with little or no science background may be due to their lack of exposure to, or indoctrination into, science and scientists. Glasson and Bentley (2000) suggest that scientists maintain views of NOS and STS that might not be completely congruent with the descriptions promoted within reform documents because they are an integral part of the scientific community. As such, scientists do not typically examine NOS or STS issues from the perspective of one outside of that community. This argument suggests that membership within the scientific community could affect one's perspective of NOS.

Three of the reports suggest the need for instruction in the history and philosophy of science in undergraduate science and/or teacher education programs (Behnke, 1961; Kimball, 1967-68; Pomeroy, 1993). Included in the report by Kimball were comparisons of the views of philosophy majors with those of scientists. The findings suggest philosophy majors possessed better understandings of NOS than scientists. Without a comparison of the views of teachers and scientists with the views of individuals who have the suggested educational experiences in the history and philosophy of science, Behnke (1961) does not have support for his recommendation to include a history and philosophy of science component in science education. Although the idea is intuitively appealing, research does not show history of science courses are necessarily effective in enhancing understandings of NOS (Abd-El-Khalick, 1998). The additional requirement for explicit instruction on NOS is indicated for individuals to gain understandings of NOS in history and philosophy of science courses.
The findings of each of the five reports reviewed here bring into question the accuracy of the assumption that scientists possess views of the nature of the discipline and processes in which they actively participate that are in agreement with the currently advocated views that are stressed in reform documents and by science educators. Moreover, these findings suggest that engaging in scientific inquiry does not automatically translate into understanding the nature of the scientific enterprise. Indeed, these reports bring into question the engagement in scientific inquiry as an essential element in the development of NOS views. Other factors, either in addition to or separate from inquiry experiences, contribute to an individual’s conceptions of NOS. Lacking from all of these studies is how experiences in science might be related to NOS and inquiry views. We still have little understanding of scientists perspectives of NOS and inquiry, and how these views may correlate with research practices.

Discussion and Conclusions

The research reviewed here identifies disparities between the intentions and outcomes of scientific inquiry experiences and views of NOS. Taken as a collection, these studies identify flaws in the assumption that people learn NOS implicitly by engaging in inquiry-based instruction alone. Review of the literature examining the effects of the inquiry-based curricula of the 1960s indicates that students did not consistently achieve higher gains in understanding of NOS, nor in understanding the processes of science. The investigations by Crumb (1965), Trent (1965), Jungwirth (1970), Tamir (1972), and Meichtry (1992) compared traditional science instruction,
which relies heavily on the textbook and “cook book” style laboratories, with
inquiry-based instruction. These studies reveal that students do not significantly
differ in changes in their conceptions of NOS as a result of engaging in the designed
inquiry-based activities. Neither approach in the context of biology or physics was
consistently effective in enhancing students’ views. The curricula examined in these
five studies were developed with the intention for students to gain an understanding
of NOS by engaging in activities whereby students were required to draw
conclusions based on observations (Shulman & Tamir, 1973; Sund & Trowbridge,
1973). This approach, as previously explicated, is implicit in nature due to the lack of
planned opportunity for students to learn about and overtly consider connections
between aspects of NOS and their investigations. From this group of authors, only
Meichtry suggested direct and consistent attention to NOS is needed to develop
adequate student conceptions. Limitations to several of these studies due to lack of
classroom observations to characterize the type of inquiry in which students were
engaged (Crumb, 1965; Trent, 1965; Jungwirth, 1970; and Tamir, 1972) reduce the
usefulness of these results because it is not known just how the curricula were
implemented. The lack of classroom observations brings into question the subject
matter and level of the inquiry activities in which the students were actually engaged.
Tamir’s results suggesting the more effective biology curriculum may be due to
teachers’ familiarity with the curriculum and subject matter rather than differences in
the nature of the disciplines. Khishfe and Abd-El-Khalick (2002) show that
familiarity with subject matter may lead to different portrayals of the nature of the
subject matter. There is a clear need to examine views of NOS and inquiry with participants who have similar expertise in subject matter and inquiry skills.

The level of inquiry promoted by the early BSCS and PSSC curricula is generally "low" and not necessarily comparable to the work of scientists (Herron, 1971; Shulman & Tamir, 1973). The majority of activities within these curricula have been described as demonstration or verification activities wherein students are aware of the "right answer" and are conducting an already outlined investigation to reinforce the concept. The lack of considerable opportunities for students to engage in high-level scientific inquiry (Herron, 1971; Schwab, 1962) may have contributed to the failure of the programs to achieve their stated objective of enhancing students' understandings of NOS. Even though inconsistencies between philosophy and practice, as demonstrated by the contrasting intentions and outcomes of the inquiry-based curricula of the 1960s, were recognized, Shulman and Tamir (1973) argued that "as long as the conception of science conveyed to the students reflects the proper role of inquiry, perhaps the relative infrequency of inquiry activities in the lessons need not be cause for alarm" (p. 1113). It would seem that even at this point, Shulman and Tamir questioned the role of active engagement in inquiry activities in the learning about NOS. Nonetheless, evaluation of the effectiveness of the curricula demonstrated the objective of engendering an understanding of NOS in the students was not achieved (Crumb, 1965; Jungwirth, 1970; Lederman, 1992; Ramsy & Howe, 1969; Tamir, 1972; Trent, 1965; Yager et al., 1969; Yager & Wick, 1966). Was this the result of the inquiry-based activities being guided rather than open, or lower level instead of higher level? Was this the result of the reliance on implicit messages to
teach about NOS? Perhaps the result was due to limiting the context of the inquiry experiences, such as focusing only on experimental investigations or only in one subject area. The reports lack sufficient description to address these issues.

The question of differential impact based on guided or open inquiry (or low vs. high level inquiry) was addressed in three of the reports in this review (Haukoos & Penick, 1983, 1985; Spears & Zollman, 1977). Spears and Zollman (1977) found no differences in achievement on the SPI between students engaging in unstructured laboratories and students engaging in structured laboratories in a physics course. The unstructured approach is considered “open” inquiry, and the structured is considered “guided.” Similarly, Haukoos and Penick (1983, 1985) investigated the effects of doing inquiry-based biology activities in a discovery classroom climate (open) versus a nondiscovery classroom climate (guided). The two studies by Haukoos and Penick achieved inconsistent results. The first study reported higher gains in understanding of NOS for the discovery group. In contrast, the replication study reported no difference between the discovery and nondiscovery group in achievement of understanding NOS. Thus, the level of structure or direction provided students during independent investigation did not consistently result in differential gains in conceptions of NOS.

The study reported by Yager and Wick (1966) looked at the effects of various curriculum emphases on students’ understandings of NOS as measured by the TOUS. Three approaches were used to teach the BSCS-Biology curriculum. Of the three approaches, the multireferenced laboratory with ideas (MRLI) approach led to the greatest gains in understanding of NOS. The MRLI approach utilized the same
materials as the multireference-laboratory (MRL) and textbook-laboratory (TL) groups (BSCS-Biology) with the addition of open discussion of the history and development of the scientific knowledge. Student ideas were frequently incorporated into the lessons. They concluded that increased emphasis on ideas and developments in science resulted in increased understandings of NOS for the students in this study. Importantly, the MRLI approach utilized by Yager and Wick is not considered here to be an explicit approach to teaching NOS. There was no clear indication that overt attention was paid to specific aspects of NOS during these classes. However, students were given opportunities to reflect on and discuss scientific developments. Such exercises, it would appear, influenced students' responses on the TOUS. However, as previously explicated, the degree of adequacy of the students' conceptions is questionable. Nonetheless, this study suggests a role for soliciting student ideas for use in discussions about science.

Yager et al. (1969) offered one of the first studies to compare different approaches to an "inquiry theme." This study differed from the others in that the level of active student engagement in the biology inquiry activities varied, rather than the amount of guidance provided the students. They found no differential gains in students' understandings of NOS due to participating in laboratory activities, observing demonstrations, or only discussing the inquiries. The results of this study indicated that "doing science" as done in the laboratory activities did not enhance students' views of NOS. Yager et al. suggested including explicit attention to NOS would aid in achieving this objective.
An important consideration from the results of these studies (Haukoos & Penick, 1983, 1985; Spears & Zollman, 1975; Yager et al., 1969; Yager & Wick, 1966) is that even though different approaches were employed, they were still considered implicit in reference to NOS in that none of the approaches offered clear discussion or instruction relating the inquiry activities to aspects of NOS. Yager & Wick (1966) reported that the MRLI group experienced more opportunity to examine and reflect on their own ideas, but the extent to which relations to NOS were examined is not known. Moreover, none of the approaches or emphases in these studies was particularly influential in heightening students’ conceptions. That is, as presented in these studies, inquiry, in any approach, was not effective in fostering adequate conceptions of NOS in the participants. These results were common for students in the biology settings and the physics settings. It should be noted that “adequate” is used tenuously here as NOS conceptions were determined by a variety of instruments (e.g. TOUS, WSPI, VS, MNSKS) that do not necessarily address the same aspects of NOS. No comparisons can be made about NOS or inquiry views held by those learners who experienced inquiry in a biology context vs. a physics context. Nonetheless, the effectiveness of implicit teaching of NOS through inquiry, regardless of context, is severely questioned based on these results.

Various attributions have been placed towards the failure of the inquiry-based curricula to achieve their NOS goals, including Hodson’s claim in a lack of consistent and contemporary philosophy and intentions within the curricula (Hodson, 1988). “...modern science courses have failed to achieve their goals because of inadequacies in the philosophical stance underpinning course design and in the
implicit philosophies of science teachers” (Hodson, 1988, p. 35). Despite encouraging reactions from students, educators were recognizing similar problems even at the time of rapid development and implementation. There was yet no consideration for how curricula for NOS might need to differ based on discipline or investigative approach.

Teachers are observing more enthusiasm, independence, and responsibility among the children. In those instances where the teachers were disappointed, it was evident that the teacher was using the newer materials without fully understanding the philosophy and intent of the programs. (Katagiri, 1964, p. 48)

Other reasons for the continued failures include the curriculum developers, who were primarily scientists rather than science educators (Duschl, 1985), the lack of authentic inquiry engagement (Herron, 1971), influential differences in teachers and teaching approach (Katagiri, 1964; Schwartz & Lederman, 2002; Trent, 1965; Yager, 1966; Yager & Wick, 1966; Yager et al., 1969), the reliance on implicit messages of NOS during the prescribed activities (Khishe & Abd-El-Khalick, 2002; Meichtry, 1992; Yager et al., 1969; Yager & Wick, 1966), lack of teacher experience conducting scientific inquiries (Gallagher, 1991; Harms & Yager, 1981; Welch et al., 1981), and the lack of teachers’ understanding of NOS (Hodson, 1988; Lederman, 1992).

Science educators have suggested that to effectively teach NOS and the nature of inquiry through inquiry experiences, the teacher must have experience in doing scientific inquiry (Barufaldi et al., 1977; Harms & Yager, 1981; Riley, 1979; Schmidt, 1967; Shapiro, 1996). Five of the studies reviewed here focused on the enhancement of prospective teachers’ views of NOS through the provision of inquiry
experiences in science teaching methods courses (Barufaldi et al., 1977; Bianchini & Colburn, 2000; Billeh & Hasan, 1975; Riley, 1979; Shapiro, 1996). This collection of studies further examines the impact of various styles of inquiry-based activities on development of NOS, and it reinforces the need for explicit NOS instruction in the context of inquiry activities. Again, lacking from this literature is a consideration for differential influences based on type of inquiry investigation or content area. Only the study by Billeh and Hasan (1975) suggests that content area may differentially affect a teacher’s views. The fact that the chemistry and physical science teachers came into the program with the lowest scores in the NOST, and the fact that the physics teachers showed the least gains, may be related to their subject matter knowledge. The authors did not address this possibility, as content difference was not a focus of their study. Too little information is provided in the report to address this particular issue.

Barufaldi et al. (1977) assessed the impact of elementary science methods courses on students’ conceptions of the tentativeness of scientific knowledge. The authors concluded that students in the courses employing hands-on, inquiry-oriented activities achieved significant gains in their understandings of the tentative nature of science. They suggest that an elementary science methods course that “stresses inquiry methods and procedures, emphasizing a hands-on approach integrated with individual problem solving, develops, alters, and enhances students’ philosophical view of science” (Barufaldi et al., 1977, p. 293). The results of this study, however, are suspect given the lack of validity measures for the instrument, lack of classroom observations to verify the implicit treatment, and the high chance of Type 1 error. In
contrast, a similar study by Riley (1979), reported no differences in gains of understandings of NOS between students engaged in active-inquiry, vicarious-inquiry (nonmanipulative), and control groups. This study by Riley was similar to the one conducted by Yager et al. (1969) in the level of active engagement of the learner varied, but learning of NOS did not. Not only are differences in developments of NOS conceptions minimal between groups of learners who engage in various levels of inquiry, but the overall conceptions of NOS held by the learners in all these studies is considered "inadequate" or uninformed based on the instruments employed in the studies. Inquiry alone, as experienced by the participants in the studies thus discussed, was not sufficient to develop informed views of NOS, as measured by the instruments in these studies.

Only four studies presented in this review purposely taught NOS and prompted participants to directly consider NOS in relation to their investigations. After describing common change themes for all participants in the larger study of the use of individual scientific investigations to enhance prospective elementary teachers' views of NOS, Shapiro (1996) elaborated on the results of one participant she treated as a case study. She incorporated personal constructs relating to specific aspects of scientific investigations and reflective interviews in an effort to discern conceptual changes that occurred as a result of doing the investigations. Because the participants were to design and conduct their own investigation, the level of inquiry in this study is considered more open than guided. The case study subject, Jan, demonstrated changes in her thinking about the nature of the steps and procedures of investigation in science, changes in her view of what science is, and changes
concerning ideas about the usefulness of independent investigations as a learning approach in the elementary science classroom. Clear from the report is that Jan’s views of NOS and the nature of scientific investigations improved throughout the course of doing an independent investigation. Shapiro claims the discussions and act of reflection served to challenge Jan’s original ideas about science and scientific investigations. The requirement of Jan to explain and support her ideas with examples from her own experiences resulted in Jan showing a shift in her conceptions. The study by Bianchini and Colburn (2000) makes a similar suggestion that the teacher plays a pivotal role in directing participants’ focus on related aspects of NOS. They report that without the teacher initiating discussion and guiding reflections through questions, the participants would not have considered NOS in relation to their activities. Furthermore, the reports by Shapiro (1996) and Bianchini and Colburn (2000) demonstrate similar progress in NOS learning even though their participants were engaged in different levels of inquiry. As described in the reports, those in the Shapiro study experienced more open inquiry, and those in Bianchini and Colburn experienced guided inquiry.

Considering the study by Billeh and Hasan (1975) provides further evidence that laboratory experiences alone are not sufficient to increase participants’ views of NOS. In their study, all groups participated in the laboratory component of the summer program, but those participants who did not receive the lecture portion of the program that directly addressed aspects of NOS did not show gains in their scores on the NOST. Therefore, just participating in the guided discovery-type laboratories of this program was not sufficient to significantly improve the teachers’
understandings of NOS. The instruction on aspects of NOS appeared to be the critical influential component of the program that enhanced participants’ views of NOS. It must be mentioned again, however, that the resulting views of these participants would not be considered acceptable or in congruence with currently accepted views. The content of the NOS instruction participants received was not described. However, because the NOS instruction was reported to be lecture format, it is assumed that the instruction was direct and explicit.

Khishfe and Abd-El-Khalick (2002) directly compared the effects of implicit and explicit NOS instruction within an inquiry context. They confirmed the suggestions from Billeh and Hasan (1975), Bianchini and Colburn (2000), and Shapiro (1996) that explicit instructional attention to NOS results in greater advances in NOS conceptions when an inquiry-based curriculum is employed.

The reports of Yager et al. (1969), Haukoos and Penick (1985), Spears and Zollman (1977), and Riley (1979) demonstrate that simply engaging in inquiry activities, regardless of level or degree of active participation, does not seem to impact conceptions of NOS. In contrast, the reports of Shapiro (1996), Bianchini and Colburn (2000), and Billeh and Hasan (1975) and Khishfe and Abd-El-Khalick (2002) demonstrate that engagement in various levels of inquiry activities can lead to enhanced views of NOS. Therefore, it does not seem to be the type of inquiry, high versus low or active versus vicarious, that determines if NOS learning can occur. The difference between the instructional approaches in the former and latter studies is the degree to which NOS is explicitly addressed in the learning environment. When NOS was not purposely addressed by the instructor, participants did not
automatically develop the desired conceptions of NOS. Positive development in NOS understanding is more likely to occur through inquiry activities when the learners are instructed about NOS and guided to make appropriate connections. This conclusion is supported by other studies that compared NOS learning outcomes resultant from explicit and implicit teaching of NOS in the context of scientific inquiry (Schwartz et al., 2003; Westerlund, Schwartz, Lederman, & Koke, 2001). [Note: These two studies are directly relevant to the issue, but they did not meet the publication requirement for inclusion in the critical review.]

Scientists who are actively creating scientific knowledge through authentic acts of scientific inquiry do not necessarily hold or portray conceptions of NOS in agreement with the currently accepted views as described by philosophers of science (Behnke, 1961; Glasson & Bentley, 2000; Kimball, 1967-68; Pomeroy, 1993; Schmidt, 1967). Further support is evident in a study by Crawford, Bell, Blair, and Lederman (1999) who examined high school students’ experiences in a science apprenticeship and views of NOS. These students spent eight hours a day for eight weeks working with a research scientist, designing and conducting their own investigation. Despite their immersion within the scientific community and their role as apprentice scientists, the students in this study did maintained somewhat naïve NOS views. More recently, Bell (2000) compared views of NOS and decision-making processes of scientists, philosophers of science, science educators, and other professionals with equal levels of education. He reported that those subjects who had experiences in philosophy of science held more informed views of NOS, regardless of participation in authentic scientific research.
Glasson and Bentley (2000) offer a rationale for scientists’ views:

The overriding view among practicing scientists is that science is essentially experimental and empirical; however, the important role of theory, the multiplicity and complexity of science methods, and the value-ladenness of science require that scientists examine the assumptions underlying their own research and what goes into the decision-making that affects research design. (Glasson & Bentley 2000, p. 483)

This is a powerful statement that suggests scientists are such an integral part of the scientific endeavor that they do not necessarily take a step back from that role to reflect upon their discipline (Schwartz et al., in press). They must assume a different, a philosophical, perspective. Until they do, they do not necessarily recognize the values and assumptions that are an inherent part of the scientific discipline itself. The importance of such a shift in perspective from “working within” to “reflection from the outside” in enabling individuals to develop informed views of NOS through engagement in scientific inquiry has been illustrated by others (Schwartz et al., 2000; in press). How, then, can students be expected to “discover” the nature of the scientific enterprise through engagement in inquiry-oriented activities? If those who develop the knowledge do not instinctively view their discipline with a philosophical eye, it is unreasonable to expect students to become intuitive philosophers during the course of an inquiry investigation.

These findings could be sufficient for science educators to rethink the role of inquiry in the science classroom as a method for teaching NOS. Although there were various methodological problems that differed among the studies, the collective findings are consistent. Such consistency is compelling and telling. These studies suggest similar outcomes but with different instruments and through different designs. They point to the same conclusion. They indicate a flaw with the
aforementioned assumption inherent to the implicit approach of teaching NOS through inquiry. Perhaps the most compelling implication from this present literature review is that engagement in authentic-scientific or classroom-based scientific inquiry is *insufficient* for developing and/or enhancing adequate conceptions of NOS in the participants. This report has presented abounding evidence in support of this claim. The missing, critical, factor needed to effectively use inquiry to teach NOS is suggested in two of the papers reviewed here (Meichtry, 1992; Yager et al., 1969) and supporting evidence is provided by four of the studies reviewed (Bianchini & Colburn, 2000; Billeh & Hasan, 1975; Khishfe & Abd-El-Khalick, 2002; Shapiro, 1996). Explicit attention to NOS, perhaps in conjunction with, and in direct reference to, the inquiry activities in which the students are engaged may be the critical pedagogical component required for success. Furthermore, guided reflections aimed to help learners make concrete connections between their ideas of NOS and the inquiry activities are suggested as an important addition to just “doing science.” Without explicit attention to NOS, learners’ NOS conceptions appear to remain unchallenged and unchanged (Barafaldi et al., 1977; Bianchini & Colburn, 2000; Billeh and Hasan, 1975; Crumb, 1965; Haukoos & Penick, 1983, 1985; Khishfe & Abd-El-Khalick, 2002; Jungwirth, 1970; Meichtry, 1992; Riley, 1979; Schwartz et al., 2000, in press; Schwartz et al., 2001; Spears & Zollman, 1977; Tamir, 1972; Trent, 1965; Westerlund et al., 2001; Yager et al., 1969; Yager & Wick, 1966).

**Explicit Teaching of NOS through Inquiry: Necessary... but content-dependent?**

Descriptions of explicit teaching of NOS through inquiry have been scattered throughout this report. However, given its importance, it seems appropriate to offer
further description and clarification. An explicit approach to teaching NOS through inquiry involves engagement in inquiry-based activities with the added instructional component that provides specific attention to aspects of NOS. Additionally; explicit teaching of NOS encourages reflection on learners' views of NOS and their relation to the inquiry activities. In this approach, NOS is viewed as science content, or a cognitive learning outcome, that requires instructional planning, attention, and assessment. In contrast to an implicit approach, explicit teaching of NOS provides overt instructional priority to relevant aspects of NOS, rather than relying on the activity itself to influence learners' views of NOS. This description of explicit teaching of NOS is supported in the studies reviewed. The information provided within the studies does not offer further details, but additional clarification is gleaned from assuming the stance that NOS is "science content."

Some readers may find a contradiction to the idea of explicitly teaching about NOS or the nature of inquiry through inquiry. If one views explicit teaching as strictly adhering to a direct instructional approach and one also views inquiry as a nontraditional instructional approach whereby the learner "discovers" for him/herself the concepts, the contradiction is apparent. It should be noted that such an idealistic, and unrealistic, view of inquiry instruction is commonly held among teachers struggling to implement inquiry-based activities (NRC, 2000). However, as used and described here, explicit teaching does not limit instruction to direct, or deductive, methods, nor is inquiry as a pedagogical approach meant to be applied as "true discovery."

In an inquiry approach to instruction there are times when it is appropriate to use reception learning techniques and expository teaching. It is not always
possible or realistic for students in a science course to figure things out for themselves. There will be many occasions when information must be conveyed to inform and interest the learner. Providing answers and ideas at the right time during instruction can enhance the learning process. (Collette & Chiappetta, 1989, p. 77)

Just as a teacher should not expect a student to come to understand gene regulation simply by doing an investigation with bacteria and antibiotics, a teacher should not expect a student to come to understand NOS and the nature of scientific inquiry simply by doing a scientific investigation. Once NOS is viewed as science content, teaching and assessing NOS becomes much less arduous (Schwartz & Lederman, 2002). The same can likely be said for views of nature of scientific inquiry. Importantly, NOS and the nature of scientific inquiry are considered science content in current national science standards (AAAS, 1990; NRC, 1996).

Within this framework, explicit teaching of NOS (and the nature of inquiry) through inquiry experiences takes form. The bottom line is that, like any science content, the concept of NOS requires instructional priority. The literature on effective teaching strategies offers some insight into the meaning and purpose of “explicitness” as intended here. Teacher effects research has resulted in a compilation of characteristics of effective teachers (Good & Brophy, 1991) that includes “Active teaching” and “Teaching to mastery.” “Active teaching” means that teachers “teach students rather than expecting them to learn mostly on their own from interacting with curriculum materials” (Good & Brophy, 1991, p. 443). Teachers provide demonstrations, explanations, and opportunity for student engagement in activities that are relevant to the content being taught. “Teaching to mastery” involves teachers providing opportunities for students to practice and apply
their newly acquired knowledge. Teachers “monitor each student’s progress and provide feedback and remedial instruction as needed, making sure that the students’ achieve mastery” (Good & Brophy, 1991, p. 443).

“Active teaching” and “Teaching to mastery” describe the essence of explicit teaching of NOS within any context, including inquiry. Like the instructors in the Shapiro (1996), Bianchini and Colburn (2000), and Khishfe and Abd-El-Khalick (2002) studies, the teacher serves as a source of information and as a facilitator to guide students’ thoughts, monitor student understanding, and model the skill of reflection. Explicitness takes the form of instruction (e.g. teacher talk, readings, videos), guided discussions (teacher guides the discussion in the desired direction), and focused questions (teacher focuses students’ thinking on NOS topics).

By considering specific aspects of NOS among a lesson’s content objectives, the teacher gives instructional priority to NOS. The teacher can explicitly teach NOS through deductive or inductive means. For example, Rosenshine and Stevens (1986) describe “guided practice” as a critical element of effective teaching through deductive strategies: In guided practice the teacher:

- Asks a large number of questions
- Guides Ss in practicing new material, initially using prompts to lead students to the correct response and later reducing them when Ss respond as desired
- Checks for Ss understanding
- Provides feedback
- Corrects errors
- Reteaches if necessary
- Provides for a large number of successful repetitions (Rosenshine & Stevens, 1986).
In this approach the teacher doesn’t just tell the students what NOS aspects he/she wants the students to learn and expect them to memorize a list. Similarly, in inductive activity, “the teacher serves as a guide to channel thinking, attempting to get students to construct appropriate labels for the relationships they have just discovered” (Collette & Chiappetta, p. 87). In terms of helping students apply newly learned concepts, “the teacher encourages the students to find examples to illustrate the concept they have just acquired” (Collette & Chiappetta, p. 88).

The context of inquiry provides concrete experiences for the learners so that meaningful learning can occur (Schwab, 1962). Yet the teacher must still provide learning opportunities that include questions and reflection exercises that help students form connections and deeper understanding of relevant aspects of NOS and inquiry activities. Feedback and correctives provide students with prompts or hints that lead them in the desired direction. If students are not recognizing examples and making genuine connections among aspects and inquiries, the teacher needs to guide them further through questioning, prompts, and reteaching. This is “Active teaching” and “Teaching to mastery” targeting NOS and inquiry within the context of inquiry-based experiences. This review of the literature highlights deficiencies in NOS learning that occur in the absence of active teaching and teaching to mastery with regard to NOS.

If teachers are to reinforce and guide learners to connect appropriate NOS and nature of inquiry concepts with various inquiry activities, they need knowledge of what aspects are appropriate based on the science content area and investigative approach employed. NOS views related to contexts of subject matter and
investigative approach (e.g. experimental vs. descriptive) is not addressed by any of the reviewed studies. Based on this literature review, "explicit teaching" is generalized across contexts as much as the concepts of NOS and the nature of inquiry. Science disciplines and inquiry methods vary considerably (Knorr-Cetina, 1999; Mayr, 1988; 1997; Ruse, 1998; Speiker, 1972). Differences in relevant NOS and inquiry aspects, as well as differences in meaning of NOS and inquiry aspects, may differ based on science contexts.

**Recommendations**

To better understand how the concepts of NOS, scientific inquiry, and the practice of scientific inquiry may relate, it seems appropriate to examine the views of scientists. The studies reported here did not characterize scientists' views beyond the limits of the instruments utilized. Use of open-ended questionnaires and interviews would provide scientists opportunity to describe in their own words how they view NOS, scientific inquiry, and their own research practice. What are the views of those who define the scientific community? What are the similarities and differences among scientists of various disciplines in how they view NOS, scientific inquiry, and their own research practices? What are the sources of scientists' views? How do their views of their research practice compare with their actual practice? These questions need to be addressed.

Furthermore, the results of Tamir (1972) and Khishfe and Abd-El-Khalick (2002) suggest that familiarity with subject matter may relate to one's views of NOS. Abilities to describe differences in conceptions of NOS and inquiry within various
contexts may be limited if the participants are students or teachers. Both of these groups may have limited subject matter knowledge or may consider different science contexts when portraying their NOS and inquiry views. Effects of the respondent’s framework need to be explored. By examining scientists’ views of NOS and the nature of inquiry, limits to level of subject matter knowledge and expertise with scientific inquiry are not conflating factors. Experienced and successful scientists possess expertise in specific content knowledge and research specialty. Exploring scientists’ epistemological views of science enable a purposeful comparison of views based on science discipline and investigative approaches that the studies reviewed here were not able to conduct. How scientists’ view the nature of their science and inquiry is a study of views that develop through experiences within authentic science. Are they all the same? Variances and constraints based on context will inform educators of explicit instructional targets. The appropriateness of the generalized treatment of NOS and inquiry across science settings and classrooms is addressed in such a study.
CHAPTER III
DESIGN AND METHOD

This study is exploratory in nature. Through intra- and cross-discipline comparisons, this study explores scientists’ epistemological views of science. These comparisons lead to descriptions of possible context-based conceptions of NOS and scientific inquiry deemed important for scientific literacy. There are two primary research questions:

1. What are practicing scientists’ views of nature of science and scientific inquiry?
2. What are the relationships, if any, among practicing scientists’ views of nature of science, views of scientific inquiry, approach to scientific inquiry, and the scientific discipline in which the scientist participates?

Method

The framework of NOS and NOSI relevant for K-12 science education guided the data collection and analysis, while seeking emergent themes and description. The open-ended elements of this study are consistent with methods that are "useful for exploring a phenomenon, for understanding it, and for developing an understanding of it into a theory. These methods humanize situations and make them come alive" (Krathwohl, 1997, p. 243). In this investigation, the phenomenon under study is scientists’ epistemological views of science.

There are four phases of the investigation. The design and schedule of data collection and analysis is presented in Figure 1. The design of the study was based
partly on a phenomenological perspective of qualitative research (Bogden & Biklen, 1992; Creswell, 1998), in that the perspectives of the scientists are sought through open-ended questionnaires and interview techniques.

<table>
<thead>
<tr>
<th>Phase 1.</th>
</tr>
</thead>
<tbody>
<tr>
<td>• Soliciting volunteers</td>
</tr>
<tr>
<td>• Collection of scientists’ vitae &amp; research descriptions.</td>
</tr>
<tr>
<td>• Selection of participants.</td>
</tr>
<tr>
<td>• Preliminary classification into discipline groupings</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Phase 2.</th>
</tr>
</thead>
<tbody>
<tr>
<td>• VNOS-Sci &amp; VOSI-Sci questionnaire administration</td>
</tr>
<tr>
<td>• Semi-structured interviews</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Phase 3. Analysis of Data</th>
</tr>
</thead>
<tbody>
<tr>
<td>• Final categorization of participants within discipline and approach groups</td>
</tr>
<tr>
<td>• Generation of Views of NOS and scientific inquiry profiles</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Phase 4.</th>
</tr>
</thead>
<tbody>
<tr>
<td>• Answering Research Questions</td>
</tr>
</tbody>
</table>

Figure 1. Study Design

Participants

The sample consisted of 24 practicing scientists (6 female, 18 male) from across the United States and representing four primary scientific disciplines and a variety of subdisciplines (Table 1). With an average of 25 years research experience since receiving their Doctoral degree, the selected scientists are experts in specific
subject matter of their discipline as well as experts in specific skills of scientific inquiry. By focusing on experts within the scientific community, this study allows a clear exploration of the research questions without conflating factors that may be associated with a novice. Suggested impacts include inability to articulate views of NOS due to limited science content knowledge (Abd-El-Khalick, 2001); confusion with NOS and inquiry concepts due to overwhelming influence of the investigative context (Lederman, Schwartz et al., 2002); and minimal response due to anxiety or negative attitude toward science (Schwartz et al., 2001).

Scientists from disciplines of physics, life science, Earth and space science, and chemistry were sought. These disciplines are consistent with the main science content areas in reform documents (AAAS, 1993; NRC 1996). These are the disciplines in which teachers are expected to teach about scientific inquiry and NOS through inquiry-based activities. Given the diverse areas of specialization within each discipline, attempts were made to obtain a diverse participant pool, while still allowing for multiple participants within each broad grouping.

The goal of this selection process was to obtain scientists who are discipline-specific experts in terms of subject matter knowledge and practices of authentic scientific inquiry. The selection criteria stem from expert/novice literature and situated cognition literature related to communities of practice (Chi, Glaser, & Rees, 1982; Samarapungavan, 1992; Wenger, 1998). This study employed four selection criteria that are considered evidence of expertise in science and scientific community membership: (1) earned Doctoral degree in science; (2) currently conducting scientific research (investigation(s) in process) in an area that can be classified into a
broad discipline of physics, life science, chemistry, or Earth and space science; (3) at least two research-based publications in peer reviewed science journals based on current research area within the past two years. Criterion (3) is consistent with a typical tenure requirement for university faculty. A fourth criterion for university-based scientists was that they be tenured or within two years from being considered for tenure. With the exception of one participant, all held tenured academic positions at universities. The exception was an aquatic ecologist with 22 years post PhD research experience, employed by an independent not-for-profit research institution. All were educated and currently employed within the United States.

Scientists from areas of specialization that integrate more than one broad discipline category were considered. For example, two astronomers were in the final pool. These were classified in the “Earth and space science” group based on their primary subject matter, the exploration of stars and stellar systems. However, they reported a strong physics background and use of physical principles in their work. There were additional non-physics participants with similar overlaps between subject matter of their primary discipline and use of physics within their research. Likewise, the molecular biologists within the life science group reported applying and exploring chemistry principles such as molecular bonding and biochemical reactions.

Volunteers were sought through university and industry contacts. Phone calls and email contacts (Appendix A) were sent to faculty in science departments on the campus of Oregon State University and additional research scientists on other campuses and research laboratories in the United States. Contacts were both in mass (sent to department faculty lists) and personalized (sent to individual scientists). The
mass mailings were unsuccessful in securing volunteers. Identifying individual scientists to contact stemmed from the researcher’s personal contacts, referrals provided to the researcher, and information gained from reading the scientists’ work or about the scientists themselves. The general intent of the study, type of data desired, and general time commitment were included in this initial contact. Potential participants were told that their opinions about science and their work is a focus of the research, but they were not told the specific research questions. The personal contacts also included information specific to the scientist’s research area and/or how the researcher had come to request their individual participation. For example, the initial message sent to the forest ecologist included, “I am trying to gain perspectives from a range of scientific disciplines. The insights you would bring from forest ecology would be a valuable addition to this study.”

Scientists who responded favorably to the initial contact were sent a second message and informed consent form (Appendix B) to confirm their willingness to participate. This mailing also contained a request for volunteers’ curriculum vitae and description of current research. The vitae provided specific details of scientists’ years, locations, and topics of professional experiences, and publication records. The current research descriptions provided details of the type of research the scientists’ conduct. The research documents and vitae were used in preliminary classification into discipline groupings (physics, life science, chemistry, Earth and space science).

Fifty-six personal contacts were made. Thirty-seven responded favorably. Of these, thirty-four eventually committed to participate. Twenty-four completed the data collection process. The most commonly reported reason for not participating or
not completing the process was time commitment. Table 1 presents information on the individual participants. Table 2 summarizes the sample, grouped by discipline areas.

The participant codes are identifying in terms of discipline, research approach, and research base. The first letter of the code is an individual identifier for the specific participant. The second letter(s) indicates the research approach. “E” is experimental. “ED” is combination of experimental and descriptive. “D” is descriptive. “T” is theoretical. The next letter(s) identifies the participant’s primary research base. “L” is laboratory. “F” is field. “FL” is both field and laboratory. “FC” is both field and computer/mathematical. “C” is computer/mathematical. The final number signifies the participant’s discipline area. “1” is life science. “2” is Earth and space science. “3” is chemistry. “4” is physics. Based on this coding system, participant “OEL1” engages in experimental research (E), in the laboratory (L), in the context of life science (1).
Table 1. Description of scientist participants

<table>
<thead>
<tr>
<th>Participant</th>
<th>Discipline</th>
<th>Approach</th>
<th>Research base</th>
<th>Gender</th>
<th>Years post-doc</th>
</tr>
</thead>
<tbody>
<tr>
<td>OEL1</td>
<td>molecular biology</td>
<td>E</td>
<td>lab</td>
<td>m</td>
<td>19</td>
</tr>
<tr>
<td>UEL1</td>
<td>molecular biology</td>
<td>E</td>
<td>lab</td>
<td>m</td>
<td>18</td>
</tr>
<tr>
<td>SEL1</td>
<td>cell biology</td>
<td>E</td>
<td>lab</td>
<td>m</td>
<td>32</td>
</tr>
<tr>
<td>BEFL1</td>
<td>forest ecology</td>
<td>E</td>
<td>field/lab</td>
<td>f</td>
<td>11</td>
</tr>
<tr>
<td>NEFL1</td>
<td>marine ecology</td>
<td>E</td>
<td>field/lab</td>
<td>f</td>
<td>26</td>
</tr>
<tr>
<td>KEDF1</td>
<td>/evolutionary development</td>
<td>E/D</td>
<td>lab</td>
<td>f</td>
<td>20</td>
</tr>
<tr>
<td>MEDF1</td>
<td>Community ecology</td>
<td>E/D</td>
<td>field</td>
<td>m</td>
<td>25</td>
</tr>
<tr>
<td>PEDF1</td>
<td>aquatic ecology</td>
<td>E/D</td>
<td>field</td>
<td>m</td>
<td>22</td>
</tr>
<tr>
<td>SEDF1</td>
<td>Entomology</td>
<td>E/D</td>
<td>field</td>
<td>m</td>
<td>24</td>
</tr>
<tr>
<td>mDF1</td>
<td>wildlife ecology</td>
<td>D</td>
<td>field</td>
<td>f</td>
<td>20</td>
</tr>
</tbody>
</table>

Earth and Space Science (ESS)

<table>
<thead>
<tr>
<th>Participant</th>
<th>Discipline</th>
<th>Approach</th>
<th>Research base</th>
<th>Gender</th>
<th>Years post-doc</th>
</tr>
</thead>
<tbody>
<tr>
<td>GDF2</td>
<td>fluvial geomorphology</td>
<td>D</td>
<td>field</td>
<td>m</td>
<td>17</td>
</tr>
<tr>
<td>eDF2</td>
<td>atmospheric science</td>
<td>D</td>
<td>field/comp</td>
<td>m</td>
<td>35</td>
</tr>
<tr>
<td>cDFC2</td>
<td>atmospheric science</td>
<td>D</td>
<td>field/comp</td>
<td>m</td>
<td>31</td>
</tr>
<tr>
<td>hEDFC2</td>
<td>astronomy</td>
<td>D</td>
<td>field/comp</td>
<td>m</td>
<td>31</td>
</tr>
<tr>
<td>pEDFC2</td>
<td>astronomy</td>
<td>E/D</td>
<td>field/comp</td>
<td>f</td>
<td>28</td>
</tr>
</tbody>
</table>

Chemistry (Ch)

<table>
<thead>
<tr>
<th>Participant</th>
<th>Discipline</th>
<th>Approach</th>
<th>Research base</th>
<th>Gender</th>
<th>Years post-doc</th>
</tr>
</thead>
<tbody>
<tr>
<td>gEL3</td>
<td>organic chemistry</td>
<td>E</td>
<td>lab</td>
<td>m</td>
<td>16</td>
</tr>
<tr>
<td>fEF3</td>
<td>environmental analytical chemistry</td>
<td>E</td>
<td>field</td>
<td>f</td>
<td>13</td>
</tr>
<tr>
<td>wEL3</td>
<td>analytical chemistry</td>
<td>E</td>
<td>lab</td>
<td>m</td>
<td>26</td>
</tr>
<tr>
<td>bEL3</td>
<td>mass spectrometry</td>
<td>E</td>
<td>lab</td>
<td>m</td>
<td>28</td>
</tr>
<tr>
<td>Code</td>
<td>Specialty</td>
<td>Approach</td>
<td>Research Base</td>
<td>Gender</td>
<td>Years post PhD</td>
</tr>
<tr>
<td>-------</td>
<td>--------------------------------</td>
<td>----------</td>
<td>---------------</td>
<td>--------</td>
<td>----------------</td>
</tr>
<tr>
<td>kEL4</td>
<td>nuclear physics</td>
<td>E</td>
<td>lab</td>
<td>m</td>
<td>33</td>
</tr>
<tr>
<td>ITC4</td>
<td>computational physics</td>
<td>T</td>
<td>comp</td>
<td>m</td>
<td>33</td>
</tr>
<tr>
<td>jTC4</td>
<td>High Energy Theoretical physics</td>
<td>T</td>
<td>comp</td>
<td>m</td>
<td>36</td>
</tr>
<tr>
<td>sTC4</td>
<td>theoretical planetary physics</td>
<td>T</td>
<td>comp</td>
<td>m</td>
<td>27</td>
</tr>
<tr>
<td>pTC4</td>
<td>relative astrophysics</td>
<td>T</td>
<td>comp</td>
<td>m</td>
<td>32</td>
</tr>
</tbody>
</table>

Approach: E: Experimental; E/D: both experimental and non-experimental; D: non-experimental (descriptive, correlational); T: Theoretical

Research base: Lab: research primarily conducted in a laboratory setting; Field: research primarily conducted in a field setting; Comp: research primarily conducted with the use of a computer (utilizes simulations) and/or mathematics

Gender: M: male; F: female

Years post PhD: the number of years since the participant earned a Doctoral degree in the science area
Table 2. Summary of scientists grouped by discipline

<table>
<thead>
<tr>
<th>Discipline</th>
<th>Total</th>
<th>Life Sciences</th>
<th>Earth &amp; Space Sciences</th>
<th>Chemistry</th>
<th>Physics</th>
</tr>
</thead>
<tbody>
<tr>
<td># of Participants</td>
<td>24</td>
<td>10</td>
<td>5</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>Average # Years post PhD</td>
<td>25.2 (avg)</td>
<td>21.7</td>
<td>28.4</td>
<td>21.3</td>
<td>32.2</td>
</tr>
<tr>
<td></td>
<td>26 (median)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>7.2 (st. dev)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>#male</td>
<td>18</td>
<td>6</td>
<td>4</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>#female</td>
<td>6</td>
<td>4</td>
<td>1</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>Research Approach</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Experimental</td>
<td>10</td>
<td>5</td>
<td>0</td>
<td>4</td>
<td>1</td>
</tr>
<tr>
<td>Descriptive</td>
<td>5</td>
<td>1</td>
<td>4</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Combination E/D</td>
<td>5</td>
<td>4</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Theoretical</td>
<td>4</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>4</td>
</tr>
<tr>
<td>Research base</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Laboratory-based</td>
<td>8</td>
<td>4</td>
<td>0</td>
<td>3</td>
<td>1</td>
</tr>
<tr>
<td>Field-based</td>
<td>6</td>
<td>4</td>
<td>1</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>Lab/Field</td>
<td>2</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Field/computer</td>
<td>4</td>
<td>0</td>
<td>4</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Computer/mathematics</td>
<td>4</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>4</td>
</tr>
</tbody>
</table>
Data Sources and Analysis

Consistent with the phenomenological tradition of research (Creswell, 1998), the research questions are not based upon a priori hypotheses that scientists of particular disciplinary and/or methodological fields adhere to any predetermined views or differ in particular ways. Rather, the investigation sought first to describe scientists’ epistemological views. Consistent with the reform-based framework of the proposed study, data collection and analysis focused on explicitly seeking scientists’ conceptions of the features of NOS and scientific inquiry deemed important and appropriate for a scientifically literate individual. As such, predetermined aspects of NOS and scientific inquiry served to guide the instrument design and initial coding of data from questionnaires and interviews (Miles & Huberman, 1994). Additional features of scientists’ epistemological views emerged through the analysis. After profiles were generated, discipline- and approach-based comparisons of scientists’ conceptions were conducted.

Multiple sources of data serve to address the questions of interest. Sources include questionnaires, semi-structured interviews, and researcher journal.

Questionnaires

Views of Nature of Science Questionnaire – Sci

Participants’ views of NOS were assessed through analysis of responses to a slightly modified version of the Views of Nature of Science Questionnaire, form C [VNOS-C] and follow up semi-structured interviews (Lederman, Abd-El-Khalick, Bell, & Schwartz, 2002). The NOS aspects initially targeted in this investigation
include that science is a) tentative, b) based on empirical observation, c) influenced by subjectivity (both personal subjectivity and theory-laden), d) comprised of both observation and inference, e) the product of human inference and creativity, and f) comprised of theories and laws that are fundamentally different types of knowledge, based on different types of data. These are among the aforementioned agreed upon NOS aspects that are generalized across scientific disciplines.

Exploring the manner in which practicing scientists in different disciplines understand and describe these aspects was a goal of this investigation. As such, the VNOS-C was modified to reflect the specialized focus of this study. Eight of the 10 original VNOS-C items remain. One item from VNOS-B (Lederman et al., 2002) was added. This item intended to address the role of creativity in science by asking for a comparison between art and science. One item from VNOS-C' (Lederman, Schwartz et al., 2002) was added to address respondents’ views of scientific models, their source and role in science. This item also targets respondents’ views of inference, subjectivity, creativity, and tentativeness. The modifications include asking respondents to justify responses and include specific examples from his/her own research endeavors that exemplify their views on the VNOS-Sci. This additional probing enables further description of the depth of respondents’ views of the targeted NOS aspects. The request for relevant examples also intends to explore connections between the respondents’ stated views and their specific scientific context. The final 10-item questionnaire, VNOS-Sci, is presented in Appendix C. Face and construct validity for the previously used VNOS items was established elsewhere (Bell, 2000; Lederman, Abd-El-Khalek et al., 2002). Participant interviews following the VNOS-
Sci administration further established validity of item interpretations and responses for the current sample.

The Views of Scientific Inquiry Questionnaire – Sci

Participants’ views of scientific inquiry were assessed through analysis of responses to a modified version of the Views of Scientific Inquiry Questionnaire [VOSI] (Schwartz et al., 2001). This is an open-ended questionnaire designed to elicit respondents’ views of five aspects of scientific inquiry. These aspects were identified from science education reform documents, reports from science educators, and reports from science studies. Aspects of scientific inquiry targeted on the VOSI-Sci include a) multiple methods of scientific investigations, b) multiple purposes of scientific investigations, c) the form and role of argumentation in the development and acceptance of new knowledge, d) recognition and handling of anomalous data, and e) sources, roles of, and distinctions between data and evidence. These aspects are not necessarily mutually exclusive from each other or from NOS aspects.

The revised and newly developed questions for the VOSI-Sci are open-ended in order to probe scientists’ views of the targeted aspects (Appendix D). A request for the respondents to justify their responses and include specific examples from his/her own research endeavors that exemplify their views on the VOSI items has also been included. A life scientist and panel of science educators established validity of the instrument. The scientist reviewed early versions of the questionnaire and provided feedback on scientists’ perceptions in general and potential interpretations of items. Modifications included deleting two items considered irrelevant, changing the order of items, modifying the wording, and including
additional examples related to questions. A panel of experts within science education and science established content validity of the VOSI-Sci items. The panel reviewed the questionnaire for alignment between items and intended inquiry aspect(s) addressed through the items. As with the VNOS-Sci, instrument validity was further established through participant interviews following the VOSI-Sci administration.

Administration of the Questionnaires

Questionnaires were either mailed or emailed to participants. Participants were told there are no right or wrong answers to any of the VNOS-Sci or VOSI-Sci items. The questionnaires could be completed in any order. Participants were asked to return an electronic version of their responses within three weeks. In a few cases, electronic return was not convenient or preferred, and these participants returned a hard copy. Notices of receipt and appreciation were sent to those who responded by the requests. Follow up emails or phone call reminders needed to be sent to scientists who do not respond by the requested deadline. On most of these occasions, their responses were fairly quick and data were either received or alternate procedures agreed upon (discussed below). However, as the attrition rate shows, not all volunteers completed the process. If they did not reply after repeated requests, they were eliminated from the pool. Thus, the final pool was 24 scientists.

Seventeen of the 24 final participants returned both completed questionnaires. Two only returned one completed questionnaire (VNOS-Sci) with indications that they preferred to provide all of their responses to the VOSI-Sci during the interview. They indicated time constraints as the reason for their not
completing the second written response. In the essence of time, the remaining five preferred to only give their responses through the interview.

Semi-structured interview

After completion of the two questionnaires, the scientists were interviewed to ensure validity of item interpretation and responses, and to further probe views of NOS and scientific inquiry. Interviews are “used to gather descriptive data in the subjects’ own words so that the researcher can develop insights on how subjects interpret some piece of the world” (Bogden & Biklen, 1992, p. 96). Regarding views of NOS and scientific inquiry, interviews have been suggested as an effective means of further validating written responses and probing deeper into respondents’ views and rationale for written statements (Leach et al., 2000; Lederman & O’Malley, 1996; Lederman, Abd-El-Khalick et al., 2002).

The initial stage of the interview involved the scientists describing their background, research agenda, and current investigations. Their vitae and website information served to probe questions during this part of the interview. Typical questions included:

- I see some of your earlier work was in ______. Can you describe what you did in this work?
- Are there other areas of science where you have conducted research?
- What is the purpose of your current line of research?
- For how long have you been working in this area?
- Please describe the research projects you are currently conducting?
• Please describe your research group, including how many graduate students and post-docs you have. Do you hold group lab meetings? If so, who participates? What is the purpose of these meetings?

• Are you involved in collaborative projects? Can you describe them?

The completed VNOS-Sci and VOSI-Sci questionnaires served to guide the portion of the interviews designated to probing scientists’ NOS and inquiry views, as previously described (Lederman et al., 2002; Schwartz et al., 2001). After discussing details of background and research agendas, participants were either shown or reminded of their written responses to the VNOS-Sci and VOSI-Sci and asked to clarify or elaborate on their responses, provide an example, and/or explain their intended use of certain terminology. For those who did not formally respond to the items, the scientists were shown the item or asked the item specifically. Follow-up questions had two aims. First, they sought examples and to probe more deeply into participants’ views related to the topic. Second, they sought clarification on language use. This technique was very important to ensure that the intentions of the respondents were aligned with the interpretations of the researcher. Prior studies indicate problems with making assumptions about common language use. For example, words such as “theory” and “experiment” hold varied meanings depending on the respondent (Leach et al., 2000; Lederman & O’Malley, 1996; Lederman et al., 2002; Schwartz et al., 2001). The general protocol for this phase of the interview was as follows:

• I see on question # ___ you said, “_______.” Can you explain a bit more what you mean by __________?
• Could you explain what you mean when you say _________?
  [responses, written or verbal]
• Can you provide an example of what you mean by _________?
  [responses, written or verbal]
• How does your response on #X relate to what you said on #Y?

It was requested that all the participants review the questionnaires in preparation for the interview, and all but one of the scientists did so. For the seven who did not provide written responses to all the questionnaire items, it was requested that they write down or at least consider their ideas related to the items in advance of the interviews. It was emphasized that their ideas did not have to be formalized in advance, but they would have the opportunity to elaborate and provide examples during the interview. This flexibility in data collection was necessary to minimize attrition and alleviate frustration for the participants with severe time constraints. Several indicated appreciation with such flexibility. With the exception of one interviewee, all other interviewees (with and without formal written responses) indicated they had reviewed the questions and their initial ideas before the interview. As such, the interviews were fairly smooth with regard to establishing purpose and focusing the participants on reflection about their practice. Indeed, many indicated having thought about the questions in the interim between filling out the questionnaires and the interview. All but one of the participants was able to be interviewed. This participant was unresponsive to email requests to arrange a telephone interview. He was maintained in the final sample because (1) his profile based on written responses did not vary from others within his discipline group any
more than the typical intra-discipline variation, and (2) for those with both written
and interview data, the interview data corroborated the written responses. Thus the
written responses were considered a valid indication of these scientists’ NOS and
inquiry views. This participant had previously sent a curriculum vita, two
descriptions of research, and a signed informed consent form.

Interviews were conducted as soon after completion of the questionnaires as
possible, based on the scientists’ availability. Interviews were either in person or via
telephone, depending on location of the scientist. The in-person interviews were
conducted either in the scientist’s office, a local coffee shop, or home. Interviews
averaged 2 hours each. All but two of the interviews were conducted in a single
session. One of these others was conducted in three 1-1.5 hour sessions (4 hours
total) and included a detailed tour of the scientist’s laboratory. These details,
although informative for understanding the specific research of the scientist, were
not particularly informative for understanding the scientist’s NOS and inquiry views.
As such, the addition of the laboratory tour for this one participant did not
compromise the study. The other interview was conducted in two 1-1.5 hour sessions
(2.5 hours total) due to schedule limitations. Both of these were participants who did
not have formal written responses. All interviews were audiotaped and transcribed
for analysis.

Analysis of VNOS-Sci and VOSI-Sci questionnaires and interview

Views of NOS and Scientific Inquiry

The open-ended VNOS-Sci and VOSI-Sci questionnaires and interviews
were used to generate descriptive profiles of respondents’ views of the initially
targeted categories. Evaluation of responses as adhering to specific philosophical positions were avoided in the analysis. First, eight NOS-related categories targeted on the VNOS-Sci, and six inquiry-related categories targeted on the VOSI-Sci served as a “start list” of codes (Lederman, Abd-El-Khalick et al., 2002; Miles & Huberman, 1994). The start list “comes from the conceptual framework, list of research questions, hypotheses, problem areas, and/or key variables that the researcher brings to the study” (Miles & Huberman, 1994, p. 58). These NOS-related codes were: tentativeness; empirical basis; creativity; subjectivity; sociocultural embeddedness; difference between theory and law; difference between observation and inference; and models. The scientific inquiry-related codes were: meaning of experiment; purpose of inquiry; methods of inquiry; identification and resolution of anomaly; justification of claims; and difference between data and evidence. It should be emphasized that these categories are not intended to represent mutually exclusive aspects related to either NOS or scientific inquiry. The category of “models” for example, relates to both scientific knowledge as well as the process of developing scientific knowledge. As was found in the analysis, both questionnaires served to inform views of categories of NOS and scientific inquiry. The utilization of multiple questionnaires provided a rich source of information. Nonetheless, the individual questionnaires were relevant for the targeted start codes. Using the written responses and verbal clarification/elaboration, a profile of NOS and inquiry conceptions was generated for each participant. Data for each participant were reviewed multiple times for confirmatory and contradictory statements until the data were sufficiently reduced and organized.
The initial codes (main categories) were supplemented with emergent main categories and subcodes (Bogden & Biklen, 1998). Additional main categories were identified as repeated occurrences of themes not included in the initial 14 start codes. Individual profiles were compared and contrasted to enable subcode generation and revision. For example, from the category of “tentativeness,” subcodes represented the range of participants’ conceptions of “tentativeness.” They included “change affirmed” (inherent tentativeness of scientific knowledge), “approaching certain knowledge” (science progresses toward understanding the world as it actually is), and “attain certain knowledge” (science succeeds in knowing absolute truth about the world). Also included as subcodes were the scientists’ views of tentativeness as they relate to “levels of certainty” and science as “self correcting.” These will be described in more detail in the results section. Several rounds of analysis were conducted to check against confirmatory or contradictory evidence in the data, making modifications as necessary until satisfactory categories and subcodes emerged. This process served to reduce and organize the data from individual’s transcripts and questionnaires, to individual’s descriptive profiles based on all main categories (represented in a 16 x 24 matrix of participant profiles with embedded quotes and comments) and finally to emergent subcodes within categories (represented in a table of 16 categories, multiple subcodes, and tallies for each scientist).

It was recognized that even though the two questionnaires were examined separately during the interview, respondents could make statements relevant to views of NOS or scientific inquiry throughout the interview. As such, the entire interview
was included throughout the analyses. Furthermore, it was recognized that some verbal and/or written statements may be representative of more than one category within the NOS or scientific inquiry analyses, or representative of both NOS and scientific inquiry. Consistency among the respondent’s responses on both questionnaires and interview was sought in generating the final profile. In the event of discrepant statements within a participant’s written and/or verbal responses, the data were reviewed again to clarify the participant’s position. In the rare case of unresolvable discrepancies, a notation was made that the participant demonstrated an inconsistent view.

Researcher journal

A researcher journal was kept during this study in order to reduce and identify episodes of researcher bias. Entries included time, extent, and nature of contacts with participants. The purpose was to ensure equal opportunity to collect the same type of information from all participants. Comments about interviews and ideas to consider further were also included.

Analysis

The researcher journal was reviewed throughout the study. As needed, interactions with scientists were modified to ensure adequate data collection needed for the comparisons targeted in the study. The entire journal was reviewed at the completion of data collection to identify evidence of bias in interviews. Bias was a recognized limitation of the study.
Final categorization of participants

Vitae, research descriptions, research examples from questionnaires, and interview data were examined to generate final classification of participants into discipline groups and classify approaches to scientific inquiry.

Answering of Research Questions

The analysis required multiple coding and searching for connections among multiple variables. Such analysis was complex and involved arranging and rearranging participant results into appropriate groups to address the specific research questions.

Research question #1. What are practicing scientists' views of nature of science and scientific inquiry?

The results of the VNOS-Sci and VOSI-Sci questionnaire and interview analyses were used to describe participants' views. Trends across participants' NOS and scientific inquiry profiles were sought to generate a description of the entire sample based on targeted and emergent aspects of NOS and scientific inquiry. Results are reported in descriptive fashion with representative quotes, as well as with percentages of the sample falling within the categories and subcodes.
Research questions #2. What are the relationships, if any, among practicing scientists' views of nature of science, views of scientific inquiry, approach to scientific inquiry, and the scientific discipline in which the scientist participates?

The results of the VNOS-Sci and VOSI-Sci questionnaire and interview analyses, descriptions of participants' approach to scientific inquiry, and discipline classifications were used to address this research question. Trends were sought (a) based on discipline classification, and (b) based on approach to scientific inquiry. Participants were divided into their respective groups (e.g. discipline groups). Tallies within each category and subcode were counted, converted to percentages of the respective groups (e.g. 40% of the life scientists purported a view of "subcode" in relation to "category"), and compared with the results for the total sample (e.g. 20% of the total sample). Because the groups were of different sizes, converting tallies to percentages of the groups enabled more meaningful comparisons. However, due to small sample sizes, statistical measures were not employed. Rather, relationships were identified by examining the trends within the main comparisons. Tendencies were sought such as "Those participants within B group are more likely to hold Y views." For example, a suggested relationship exists if there is a tendency to hold a particular view (of inherent tentativeness of scientific knowledge, for example) among scientists within a certain discipline (e.g. Earth and space science) or conducting a particular type of inquiry (e.g. descriptive research). Data were compiled based on percentages (e.g. ___% of participants in B group also hold Y views). Trends and patterns were sought based on discipline and approach to inquiry.
The Researcher

This investigation explores practicing scientists' epistemological views of science. The science background, experiences, and views of the researcher relative to authentic science practice and science epistemology allows initial understanding and access to participants’ contexts and views, but may also bias her procedures and interpretations during this study. There is a need to recognize the researcher’s background and views about NOS and scientific inquiry.

The researcher holds a bachelor’s degree in genetic biology and a master’s degree in molecular biology. She conducted original research in molecular biology as an undergraduate. For a year and a half, she explored the SOS DNA repair mechanism in chlorellophage. For nearly four years as a master’s student, she explored translational frameshifting and nucleotide sequence bias in E. coli. Both of these research experiences utilized molecular research techniques within a laboratory setting. Both were experimental in approach. Both involved presentations of the research to peers and scientists. The master’s work resulted in a publication in a peer reviewed science journal.

Despite formal schooling and experience with scientific research, she only began considering aspects of NOS and the nature of scientific inquiry during her doctoral work in science education. She initially thought about NOS during a science education course soon after completing her master’s degree. Her initial reaction was to resist thinking about science in philosophical ways. The purpose of considering epistemological issues of science was not clear from within her framework of a working scientist. She felt philosophical issues were not relevant to the practice of a
scientist, and thus not relevant to the teaching of science. Through readings, reflections, and discussions with science educators, she became aware of the relevance of NOS to K-16 science education as a means to teach about science and where the scientific knowledge comes from. Understanding NOS and inquiry provides meaning to the knowledge base of science. The transition from seeing NOS and the nature of scientific inquiry as irrelevant to scientific knowledge to seeing their understanding as a necessary component of scientific knowledge, without which one merely knows isolated subject matter, marked the transition in the researcher’s identity from scientist to science educator.

To the researcher, NOS is a multifaceted concept. Central to this concept is recognizing the tentativeness of all scientific knowledge. This inherent characteristic is best explained through other elements of NOS. Science is based on observations and inferences within the empirical world. Empirical data provide scientific knowledge stability, but not infallibility. Like pieces of a fluid puzzle, data provide a glimpse into nature, but these pieces are collected and interpreted within a conceptual framework to find meaning. Meaning is temporally based, and change can occur through perspective changes as well as through the gathering of additional observations. Current understanding, values, and personal background effect what is investigated, how investigations are designed, and how data are interpreted and utilized. All facets of the scientific endeavor are influenced by the biases of the practitioner as well as the culture and society in which the science is conducted. Scientific knowledge is the product of a human endeavor that involves creativity and imagination to develop understanding of the natural world. Creativity and
subjectivity are unavoidable and, moreover, desirable aspects of science. Without creativity, scientists would not be able to ponder questions, conduct inquiries, or infer meaning from data. Without subjectivity, science progress would halt. Subjectivity in terms of the theory-laden NOS relates to the consistency, building, and acceptance of scientific knowledge. Current knowledge directs the scientific endeavor. This is the theory-laden NOS. Observations are not blindly gathered and meaning revealed. Observations depend on the senses: what we see; what we hear; what we feel; or extensions of those senses: what we can measure. Inferences place meaning on the observations. Scientific theories and laws develop from observations and inferences, but are two different types of knowledge. Scientific theories offer explanations for observed phenomena. Theories are well supported empirically and address the question of “why” or “how.” Scientific laws provide a description of relationships among observed phenomena. Laws are also well supported empirically, but address the question of “what.” Neither provides absolute knowledge. Both are subject to change. Distinctions between observation and inference, and theory and law, relate to the empirical NOS. All scientific knowledge has some basis in the empirical world. This is what distinguishes science from other ways of knowing, such as religion or even mathematics. These other ways are creative and subjective and culturally bound, like science, but they do not necessarily have firm foundation in empirical observations. Because of the theory-laden NOS; personal subjectivity of the scientists, culture, and society; and role of creativity in observation and inference, scientific knowledge is inherently tentative. As new data are gathered and as existing data are reexamined, changes in our understanding of
the world occur. Change in knowledge base, change in societal values and expectations, and change in personal experience all contribute to change in scientific knowledge.

The researcher describes scientific inquiry as the processes and activities involved in the development and acceptance of scientific knowledge. The processes take many forms, but common to all inquiries is a central question. The origin of the central question varies. The researcher does not consider there to be one Scientific Method. Investigations may be descriptive or exploratory, aiming to describe a phenomenon as it exists. Investigations may aim to identify relationships within the phenomenon or among phenomena. Describing relationships does not necessarily involve variable manipulation or controls. Examples of fields of science that often conduct descriptive and correlational studies include astronomy, anatomy, anthropology, and field ecology, among others. Experimental science involves identification and manipulation of variables and utilization of controls to suggest cause/effect relationships. The researcher’s experiences with authentic scientific inquiry typically followed experimental methods. Scientific inquiry involves the gathering of data relevant to the question of interest. Data are observations. Data interpreted in light of that question, which support a conclusion, are considered evidence. The justification of a conclusion depends on the strength of the evidence. Strength of evidence depends on validity of the data collection and consistency of interpretation with current knowledge. Herein lies an overlap between inquiry and NOS. Argumentation is an inferential, subjective, and creative process. The researcher believes that expectations for valid data and justified conclusions depend
on the science field and existing paradigms within fields. Identification of anomalies varies with investigation. Expectations of the scientist (subjectivity) influence what data are considered anomalous. How scientists choose to deal with recognized anomalies can take several forms including ignoring the data, explaining the data as experimental error, investigating to find an explanation within the current framework, investigating to find a novel explanation, incorporating the anomaly within the framework without explanation, recognizing the anomaly as unexplained and disregarding for current time. The researcher acknowledges the possibility for other approaches to dealing with anomalous findings. The purpose of scientific inquiry also varies. Some investigations aim to develop knowledge that will help living beings have healthier and longer lives. Some are conducted to inform technology advances. Some aim to better understand and conserve the natural world. The researcher believes most investigations have multiple purposes. The purpose of the researcher’s work with translational frameshifting was to better understand the relationship between sequence bias and translational accuracy. This research contributed to understanding of the evolutionary basis for sequence bias, stability of base-pairing during protein synthesis, and potential identification of programmed frameshift sites (such as those within HIV) that could be targets for health-related research and development.
CHAPTER IV
RESULTS

The purpose of this study was to explore the epistemological views of scientists representing the science disciplines of life science, physics, chemistry, and Earth and space science. The study also aimed to explore possible contextually-based variances in scientists’ epistemological views. The specific research questions addressed are:

1. What are practicing scientists’ views of nature of science and scientific inquiry?

2. What are the relationships, if any, among practicing scientists’ views of nature of science, views of scientific inquiry, approach to scientific inquiry, and the scientific discipline in which the scientist participates?

This chapter first presents the results for the categorization of the participants into research approach groups. Second, the results of the analysis of all the data for identifying emergent categories and subcodes are presented. Third, results of the analysis of questionnaire and interview data are presented for the total sample (Research question #1) and for the sample when grouped according to science discipline and then when grouped according to research approach (Research question #2). The inclusion of the cross-approach comparison provides additional information to aid the overall exploration as well as offer insights into discipline-based trends. This was especially useful given the fact that 4 of the 5 physicists were theoretical researchers. Their views, and the views of all in the sample, may be reflective of
their research approach as opposed to, or in association with, their discipline area. These results are discussed in terms of suggested patterns within this sample of scientists.

Research Approaches

Four research approaches emerged from review of the data. They were “experimental,” “descriptive,” “experimental/descriptive combination,” and “theoretical” (Tables 1 and 2). Participants were classified as “experimental” if their primary research practice involved traditional manipulative investigations with controlling variables and assessing cause/effect relationships. Participants were classified into the “descriptive” group if their research was primarily void of direct manipulative features. The five researchers within this group conducted primarily correlational studies and/or observational studies. The atmospheric scientists in this group emphasized modeling systems (e.g. clouds) through use of computer simulations developed from data they collect in the field as well as satellite data. Those who worked primarily in the realm of mathematical computations to derive explanations for the natural world, and classified themselves as theoretical scientists, were classified as such. There were five scientists who were involved in combinatorial programs of experimental and descriptive investigations. One of the astronomers, for example, indicated she conducted experiments, but not in the traditional direct-manipulative way. She selected the stellar systems she wanted to explore and compare based on their composition and characteristics. She considered this work experimental because of her purposeful selection of systems that differed
by the one variable of interest to her. In this way she identified causal factors. Her work also involved description of the stellar systems and identifying correlations among system features. Similarly, four of the life scientists participated in combinatorial research. The community ecologist, for example, maintains fields of a host plant (a pitcher plant) that he uses to study the ecological relationships among the organisms that live within the pool of water the pitcher plant holds. He not only describes natural relationships, he also manipulates the micro-environments within the host plants (e.g. adding or removing a particular organism) to examine the effects of the introduced changes.

Scientists were also categorized according to their primary base or location of research practice. There were scientists who conduct laboratory research, such as the molecular biologists, cell biologist, and analytical chemist. There were those who conduct field-based research, such as the community ecologist and wildlife ecologist. There were two scientists who conduct a combination of both field and laboratory research, such as the forest ecologist who is doing genetic studies of forest canopies. There were four who reported conducting a combination of field and computational (computer simulation/mathematical) research, such as the astronomers and the atmospheric scientists. Finally, there were the theoretical researchers who work strictly within the computational world of mathematics and computer simulations. These classifications were identified according to the scientists’ descriptions of their research. They are included here as descriptors of the sample only, and did not serve a role in the present analysis.
The scientists were forthcoming with their views. They articulated their positions with supporting examples and anecdotes from which multiple categories of NOS and inquiry emerged. There were 14 initial NOS and NOSI start codes used in the analysis. These start codes were based on aspects of NOS and NOSI shown to be relevant to K-12 education. They served to guide the initial round of analysis. Two additional themes emerged, to make a total of 16 main categories related to the sample’s epistemological views of science. The two added aspects were “reproducibility” and “prediction.” After repeated rounds of coding and analysis, multiple subcodes emerged within each main category to represent the scientists’ views. Individuals who affirmed a view aligned with an emergent subcode received a tally for that code. Subcodes were not mutually exclusive. As such, individuals could have more than one occurrence within the subcodes for a main aspect. Likewise, not all individuals’ responses yielded information relative to the main aspects. It is important to note that the subcodes emerged from the qualitative data. That is, these are the ideas offered forth by the participants through questionnaire and interview responses. Not being represented within a subcode means the participant did not make a statement aligned with that position which was offered by another participant. Table 3 presents the list of 16 main aspects and respective subcodes for each.
Views of NOS and NOSI

Table 3 presents the results for the total sample for the 16 main aspects and subcodes. The discussion following identifies features of these results, with representative quotes, and presents results for discipline and approach groups. The responses included here demonstrate representative views of the main aspects. They also demonstrate the interconnections among the aspects and subcodes. The reader may find the quotes demonstrate multiple categories. The open-ended qualitative approach to this exploration enables such coding and identification of connections.

Results for each main aspect are explored with respect to the whole sample (Research question #1) and discipline-based and approach-based comparisons (Research question #2).

Table 3. Scientists' Views of NOS and NOSI

<table>
<thead>
<tr>
<th>Aspect</th>
<th>Subcode</th>
<th>Total</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>#</td>
<td>%</td>
</tr>
<tr>
<td>tentativeness yes</td>
<td>affirm change</td>
<td>11</td>
<td>45.8</td>
</tr>
<tr>
<td></td>
<td>levels of certainty</td>
<td>8</td>
<td>33.3</td>
</tr>
<tr>
<td></td>
<td>complexity</td>
<td>4</td>
<td>16.7</td>
</tr>
<tr>
<td></td>
<td>discipline</td>
<td>4</td>
<td>16.7</td>
</tr>
<tr>
<td></td>
<td>method</td>
<td>2</td>
<td>8.3</td>
</tr>
<tr>
<td></td>
<td>approaching certainty</td>
<td>4</td>
<td>16.7</td>
</tr>
<tr>
<td>tentativeness no</td>
<td>attain certain knowledge</td>
<td>5</td>
<td>20.8</td>
</tr>
<tr>
<td>other: self correcting</td>
<td></td>
<td>2</td>
<td>8.3</td>
</tr>
<tr>
<td>empirical yes</td>
<td>empirically grounded</td>
<td>17</td>
<td>70.8</td>
</tr>
<tr>
<td></td>
<td>theoretical yet empirically grounded through confirmation of predictions</td>
<td>2</td>
<td>8.3</td>
</tr>
<tr>
<td></td>
<td>theoretical/mathematic: changing what science is... (string theory)</td>
<td>3</td>
<td>12.5</td>
</tr>
</tbody>
</table>
### Subjectivity

<table>
<thead>
<tr>
<th>Description</th>
<th>Count</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Personal only</td>
<td>9</td>
<td>37.5</td>
</tr>
<tr>
<td>Theory-laden (guiding framework)</td>
<td>15</td>
<td>62.5</td>
</tr>
<tr>
<td>Positive</td>
<td>8</td>
<td>33.3</td>
</tr>
<tr>
<td>Negative</td>
<td>4</td>
<td>16.7</td>
</tr>
<tr>
<td>Inconsistent response</td>
<td>3</td>
<td>12.5</td>
</tr>
<tr>
<td>Differs with discipline</td>
<td>4</td>
<td>16.7</td>
</tr>
<tr>
<td>Differs with approach: qualitative &gt; quantitative</td>
<td>5</td>
<td>20.8</td>
</tr>
<tr>
<td>Not subjective</td>
<td>4</td>
<td>16.7</td>
</tr>
<tr>
<td>Scientific Method = objectivity (ideal or practice)</td>
<td>2</td>
<td>8.3</td>
</tr>
</tbody>
</table>

### Creativity

<table>
<thead>
<tr>
<th>Description</th>
<th>Count</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Yes affirmed</td>
<td>16</td>
<td>66.7</td>
</tr>
<tr>
<td>Finding patterns/build connections</td>
<td>7</td>
<td>29.2</td>
</tr>
<tr>
<td>Epiphany only examples</td>
<td>3</td>
<td>12.5</td>
</tr>
<tr>
<td>Progress</td>
<td>5</td>
<td>20.8</td>
</tr>
<tr>
<td>No data to be clear</td>
<td>3</td>
<td>12.5</td>
</tr>
<tr>
<td>Inconsistent responses</td>
<td>1</td>
<td>4.2</td>
</tr>
<tr>
<td>Other</td>
<td>1</td>
<td>4.2</td>
</tr>
</tbody>
</table>

### Socio/Cultural

<table>
<thead>
<tr>
<th>Description</th>
<th>Count</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Process/conclusions (how done, reasoning)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Internal: influence on reasoning through processes, assumptions</td>
<td>4</td>
<td>16.7</td>
</tr>
<tr>
<td>Not s/c influenced: international community, technology</td>
<td>5</td>
<td>20.8</td>
</tr>
<tr>
<td>External: political and economic pressures</td>
<td>3</td>
<td>12.5</td>
</tr>
<tr>
<td>Questions/problems studied (What)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Human endeavor</td>
<td>1</td>
<td>4.2</td>
</tr>
<tr>
<td>Soc/c influenced: political, economic, societal pressures</td>
<td>15</td>
<td>62.5</td>
</tr>
<tr>
<td>Not s/c influenced</td>
<td>4</td>
<td>16.7</td>
</tr>
<tr>
<td>Products</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Universal application/generalizability</td>
<td>3</td>
<td>12.5</td>
</tr>
<tr>
<td>Not universal use</td>
<td>1</td>
<td>4.2</td>
</tr>
<tr>
<td>SM: Not s/c influenced</td>
<td>2</td>
<td>8.3</td>
</tr>
<tr>
<td>Other</td>
<td>4</td>
<td>16.7</td>
</tr>
</tbody>
</table>

### Theory/Law

<table>
<thead>
<tr>
<th>Description</th>
<th>Count</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hierarchical (theories become laws)</td>
<td>13</td>
<td>54.2</td>
</tr>
<tr>
<td>Different</td>
<td>8</td>
<td>33.3</td>
</tr>
<tr>
<td>Levels of confidence differ: laws &gt; theories; foundations</td>
<td>5</td>
<td>20.8</td>
</tr>
<tr>
<td>Theories more general &amp; complex/laws simple</td>
<td>3</td>
<td>12.5</td>
</tr>
<tr>
<td>Theories more likely to change</td>
<td>2</td>
<td>8.3</td>
</tr>
<tr>
<td>No difference</td>
<td>1</td>
<td>4.2</td>
</tr>
<tr>
<td>Differs with discipline</td>
<td>6</td>
<td>25.0</td>
</tr>
<tr>
<td>No laws in field of work</td>
<td>6</td>
<td>25.0</td>
</tr>
<tr>
<td>Obs/inference</td>
<td>affirmed role of inference</td>
<td>14</td>
</tr>
<tr>
<td></td>
<td>not clear from responses</td>
<td>10</td>
</tr>
</tbody>
</table>

| Models | explain or organize obs/predict/test | 17 | 70.8 |
| | understanding of system/complexity made | 9 | 37.5 |
| | simple/abstract made visual | 9 | 37.5 |
| | mathematics | 3 | 12.5 |
| | physical system | 1 | 4.2 |
| | analogy | 1 | 4.2 |
| | mental construct | 1 | 4.2 |
| | representation of reality | 1 | 4.2 |
| | more specific than theory | 2 | 8.3 |
| | Directing framework | 3 | 12.5 |
| | other | 4 | 16.7 |
| | N/A | 2 | 8.3 |

| Experiment | traditional: controls/variables/manipulation | 17 | 70.8 |
| | hypothesis driven | 10 | 41.7 |
| | hypothesis not required | 0 | 0.0 |
| | test of models | 2 | 8.3 |
| | inclusive: any test of idea, with or without manipulation | 2 | 8.3 |
| | variable by context: lab versus field | 6 | 25.0 |
| | requires replicas | 8 | 33.3 |
| | required for science | 3 | 12.5 |
| | not required for science | 13 | 54.2 |

<p>| Purpose | Basic | 3 | 12.5 |
| | understanding | 14 | 58.3 |
| | curiosity | 2 | 8.3 |
| | Applied | 6 | 25.0 |
| | improve quality of life | 4 | 16.7 |
| | answering questions | 1 | 4.2 |
| | Predict | 3 | 12.5 |
| | Discovery/Serendipity | 1 | 4.2 |
| | differs among disciplines | 3 | 12.5 |
| | no response | 3 | 12.5 |</p>
<table>
<thead>
<tr>
<th>Methods</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Single Scientific Method</strong></td>
<td>6</td>
</tr>
<tr>
<td>Multiple methods ...</td>
<td>16</td>
</tr>
<tr>
<td>all hypothesis driven</td>
<td>7</td>
</tr>
<tr>
<td>not all hypothesis driven</td>
<td>8</td>
</tr>
<tr>
<td>reductionist</td>
<td>2</td>
</tr>
<tr>
<td>expt, descriptive</td>
<td>11</td>
</tr>
<tr>
<td>expt, observational, theoretical</td>
<td>7</td>
</tr>
<tr>
<td>methods influenced by complexity of systems</td>
<td>5</td>
</tr>
<tr>
<td><strong>Hierarchy:</strong></td>
<td></td>
</tr>
<tr>
<td>Validity/confidence of claims influenced by methods:</td>
<td></td>
</tr>
<tr>
<td>expt &gt; descriptive</td>
<td>9</td>
</tr>
<tr>
<td>Maturity of the science: Expt more mature than desc</td>
<td>3</td>
</tr>
<tr>
<td>Theoretical sciences: pushing the limits of science/mathematics</td>
<td>4</td>
</tr>
<tr>
<td>Discipline differences: complexity of systems, ability to repeat</td>
<td>9</td>
</tr>
<tr>
<td><strong>Anomaly</strong></td>
<td></td>
</tr>
<tr>
<td>inconsistency with expectation</td>
<td>17</td>
</tr>
<tr>
<td>error in measurement: repeatable?</td>
<td>14</td>
</tr>
<tr>
<td>excitement/progress</td>
<td>10</td>
</tr>
<tr>
<td>natural variation?</td>
<td>3</td>
</tr>
<tr>
<td>Expand existing model</td>
<td>11</td>
</tr>
<tr>
<td>develop new model/theory to explain</td>
<td>7</td>
</tr>
<tr>
<td>discovery</td>
<td>2</td>
</tr>
<tr>
<td>ignore/discard</td>
<td>0</td>
</tr>
<tr>
<td>set aside for rainy day</td>
<td>3</td>
</tr>
<tr>
<td>depends on context</td>
<td>2</td>
</tr>
<tr>
<td>report finding/no explanation necessary</td>
<td>1</td>
</tr>
<tr>
<td>other</td>
<td>4</td>
</tr>
<tr>
<td><strong>Justification</strong></td>
<td></td>
</tr>
<tr>
<td>reproducibility</td>
<td></td>
</tr>
<tr>
<td>internal: statistics</td>
<td>11</td>
</tr>
<tr>
<td>external: multiple</td>
<td></td>
</tr>
<tr>
<td>researchers/same results</td>
<td>3</td>
</tr>
<tr>
<td>consistency with others</td>
<td>3</td>
</tr>
<tr>
<td>experiments over description</td>
<td>10</td>
</tr>
<tr>
<td>peer review</td>
<td>9</td>
</tr>
<tr>
<td>address alternatives</td>
<td>5</td>
</tr>
<tr>
<td>predictions/tests</td>
<td>8</td>
</tr>
<tr>
<td>model/ predict/test</td>
<td>2</td>
</tr>
<tr>
<td>scientific method/hypothesis test</td>
<td>3</td>
</tr>
<tr>
<td>differs with discipline/context</td>
<td>14</td>
</tr>
<tr>
<td>not clear from response</td>
<td>2</td>
</tr>
<tr>
<td>other</td>
<td>4</td>
</tr>
<tr>
<td>Data/evidence</td>
<td>1</td>
</tr>
<tr>
<td>-----------------------------------</td>
<td>---</td>
</tr>
<tr>
<td></td>
<td>15</td>
</tr>
<tr>
<td>data: measurements/observations</td>
<td>10</td>
</tr>
<tr>
<td>evidence: interpreted data in relation to a question</td>
<td>2</td>
</tr>
<tr>
<td>same</td>
<td>2</td>
</tr>
<tr>
<td>no evidence in field</td>
<td>5</td>
</tr>
<tr>
<td>no response</td>
<td>3</td>
</tr>
<tr>
<td>other</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Reproducibility</th>
<th>8</th>
<th>33.3</th>
</tr>
</thead>
<tbody>
<tr>
<td>statistics</td>
<td>9</td>
<td>37.5</td>
</tr>
<tr>
<td>requirement to be scientific/acceptable</td>
<td>9</td>
<td>33.3</td>
</tr>
<tr>
<td>depends on method/discipline</td>
<td>5</td>
<td>20.8</td>
</tr>
<tr>
<td>no mention</td>
<td>2</td>
<td>8.3</td>
</tr>
<tr>
<td>other</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Prediction</th>
<th>21</th>
<th>87.5</th>
</tr>
</thead>
<tbody>
<tr>
<td>affirmed important</td>
<td>6</td>
<td>25.0</td>
</tr>
<tr>
<td>requirement to be scientific</td>
<td>10</td>
<td>41.7</td>
</tr>
<tr>
<td>through models</td>
<td>2</td>
<td>8.3</td>
</tr>
<tr>
<td>through conclusions</td>
<td>2</td>
<td>8.3</td>
</tr>
<tr>
<td>testable</td>
<td>3</td>
<td>12.5</td>
</tr>
<tr>
<td>no mention</td>
<td>3</td>
<td>12.5</td>
</tr>
<tr>
<td>differs with discipline</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Aspect</td>
<td>Subcode</td>
<td>Total</td>
</tr>
<tr>
<td>-----------------------------</td>
<td>-----------------</td>
<td>-------</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>24</td>
</tr>
<tr>
<td>tentativeness</td>
<td>yes affirm change</td>
<td>11</td>
</tr>
<tr>
<td></td>
<td>levels of certainty</td>
<td>8</td>
</tr>
<tr>
<td></td>
<td>complexity</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td>discipline</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td>method</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>approaching certainty</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td>no attain certain knowledge</td>
<td>5</td>
</tr>
<tr>
<td></td>
<td>other self correcting</td>
<td>2</td>
</tr>
</tbody>
</table>
### Table 5. Scientists' Views of Tentativeness: Grouped by Approach

<table>
<thead>
<tr>
<th>Aspect</th>
<th>Subcode</th>
<th>Total</th>
<th>Grouped by Approach</th>
<th>Number of scientists</th>
<th>Percentage of group</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Total</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>E</td>
<td>E/D</td>
<td>D</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>%</td>
<td>%</td>
<td>%</td>
</tr>
<tr>
<td>yes</td>
<td>affirm change</td>
<td>11</td>
<td>3</td>
<td>3</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>levels of certainty</td>
<td>8</td>
<td>3</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>complexity</td>
<td>4</td>
<td>16.7</td>
<td>0</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>discipline</td>
<td>4</td>
<td>16.7</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>method</td>
<td>2</td>
<td>8.3</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>approaching certain knowledge</td>
<td>4</td>
<td>16.7</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>no</td>
<td>attain certain knowledge</td>
<td>5</td>
<td>20.8</td>
<td>0</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>other; self correcting</td>
<td>2</td>
<td>8.3</td>
<td>0</td>
<td>1</td>
</tr>
</tbody>
</table>

E: experimental; E/D: combination of experimental and descriptive; D: descriptive (non-experimental); T: theoretical
Tentativeness

Tables 4 and 5 present the results of the scientists' views of tentativeness of scientific knowledge. The columns with the “yes” and “no” indicate the conviction of the responses toward tentativeness. Those under the “yes” subcode viewed all of scientific knowledge as inherently subject to change. Those under the “no” subcode considered absolute knowledge of reality attainable through scientific endeavors.

The other subcode characteristics were more variable, as explained below.

Scientists' Views

_Inherent tentativeness._ Eleven of the 24 (45.8%) scientists affirmed that scientific knowledge is inherently tentative. Five of these explicitly made reference to not being able to prove anything absolutely in science, only disprove.

Usually an experiment only disproves something, a theory. So physicists tend to have an open mind. You postulate a theory. You do all the experiments you can think of. You predict new things and then you measure that. Then you say, ok, those experiments have proved that the experiment is valid. You haven't proved that it is right. It can always be wrong and superceded by another theory. But an experiment can prove that a theory is wrong. It can't prove it is right. So if it disagrees with an experiment it is wrong. If it agrees, it just means it isn't wrong. [ITC4, int]

You can't prove anything absolutely because the range of your observational powers is limited. It is limited in both quality and quantity. All Right. Just for example we cannot count to infinity. I don't care how big a number you pick, I can always pick a number that's bigger. There is no end. We can't conceive of an end. So we cannot possibly perform an infinite number of experiments. That is the only way we could get absolute knowledge. If there is an omniscient force out there, only that omniscient can do that. But you and I can't possibly do that. Okay? So A. we can't count to infinity. So because of that we cannot prove anything. [bEL3, int]

When this same participant went on to say _how_ one disproves in science, he foreshadowed his views on anomaly:
We can only disprove things. Disproving is relatively easy. All you have to do is find one exception. If the exception is repeatable, then the theory is gone.

In relating his ideas about “disproving” in science, one of the molecular biologists revealed that his views had changed since becoming a scientist:

S: [in explaining hypothesis] The concept of hypothesis is actually very important because you can never prove yourself right. You can only prove yourself wrong. ...But you collect data and you can have an idea and make a prediction as to what it might be and that is what your hypothesis is, and you test that. If it doesn't work, you know something is wrong with your hypothesis. If it does work, uh...there still could be other ways to explain it and you try to go through those in as many ways as you want and can to try to disprove the hypothesis. Even then it is not finally proven.
R: Is the hypothesis then like your prediction?
S: Yes...Sometimes you get really lucky and you can predict, based on umbers, we can calculate something, we can measure it, and it falls spot on...So that is nice, but it is not absolute proof that that is the way it happens...My view has changed a lot from when I started out. I couldn’t believe in Kuhnsian paradigm shifts.
R: When you first started?
S: Yeah. Well ‘of course there is truth in knowledge’, you know. Everybody knows what is right and wrong and it’s absolutely true.
R: That is a very interesting statement. As a graduate student you held these views?
S: No, more as an undergrad. [Goes on to tell a story and discuss the misrepresentation of science in the media] [OEL1, int]

Others related their view of inherent tentativeness to their view of scientific theories and laws. Their understanding of laws as confirmed absolutes were reconciled with their being unattainable within science, because all knowledge is subject to change:

Law somehow applies infallibility, which is impossible given the bizarre way that science works. You have to remember that science works only by disproving alternative theories, not “proving” any one theory. Thus, nothing is a certainty in science, making ‘laws’ unrealistic. [MEDF1, vnos]

S: There have been several points in the history of science, interestingly, where scientists absolutely believed that was it. There was nothing more to be discovered. And so anything that was ruled at that time would be laws. We
now know that there is hardly anything we know that probably there isn’t something else that eventually what we know now will be a special case of a broader overarching theory or concept of nature. So I don’t find a lot of distinction between these two things.

R: So do you think reaching this law status is unachievable in the strictest sense?

S: I believe so. The evidence of the past is that very few things that we might consider laws now might survive as such. Even if we don’t find that overarching theory or law that these belong, that might simply be the limitation of our own intellect or our ability... We can’t really know that things are irrefutable or engraved in granite and came down to Moses like the Ten Commandments. Things just don’t work that way. [bEL3, int]

Another view related to conventions of the scientific community:

I think a lot of scientists take themselves too seriously. It is really hard to get scientists to stop thinking 'statistics prove' because they don't. I think people take the numbers too seriously and sometimes their own hypotheses too seriously. Scientists fight over things... And of course, just the uncertainty. We've got all these uncertainty laws... ocam's razor, Heisenberg’s uncertainty. If people think about it there are all these reminders that people have thought about in the past that we don’t know what we are measuring in all cases, or we have changed what we are measuring. We may be accepting the simplest answer, the most pragmatic possibility, but that doesn't mean it is always the right one. There are just conventions that we have all agreed to use more or less. That doesn't mean they are right. [SEDF1, int]

**Approaching certainty or attaining certainty.** There were 9 scientists (37.5%) who held views that science attains certain knowledge (5 scientists) (that being absolute knowledge of the reality separate from the observer) or that science progresses nearer and nearer to certain knowledge (4 scientists). One scientist, an atmospheric scientist, responded when asked about the certainty of the model of the atom, “Certain. It's the way nature is.” [cDFC2, vnos] This is in contrast to one of the theoretical physicists who stated in reference to scientists understanding of the atom, “As certain as we can be.” [pTC4] Others related their views of certainty to their views of scientific laws. These included,
Scientists do not develop laws. They discover them! [KEDF1, vnos]

If you mean valid in the sense of true, I don’t think that is knowable, *at least not immediately.* You get more and more certain with passing time. [SEL1, int] [emphasis added]

This latter participant also indicted patterns in nature are “there to be discovered.”

Four scientists indicated that science progresses toward better approximations of the truth:

I don’t believe in certainty at all in science. I don’t believe certainty exists. I believe we approach certainty as a limit, but we don’t get there. [BELF1, int]

The aquatic ecologist commented on approaching certainty through the self-corrective nature of science in his VNOS written responses:

Scientific knowledge changes as better approximations of nature are realized while religious knowledge is dependent on established (or accepted) elements….Theories change as better approximations of nature (models) evolve. Theories encompass new complexities….Laws may change because they are shown to be wrong. All scientific knowledge is subject to question, doubt and criticism (a further distinction from religion)….Nonetheless, someone will eventually challenge an accepted scientific finding and take a fresh look at it. ….That is the self-corrective nature of science. Does science lead to universal truths? It leads to close approximations of universal truths. [PEDF1, vnos]

*Levels of certainty.* Eight participants (33.3%) indicated scientific knowledge can vary in terms of certainty. That is, some types of knowledge are more certain than others.

I think there are levels of certainty. I think that there are certain principles or insights that have very high levels of certainty attached to them. I think that some of that certainty transcends the scientific method but lends itself to it…. [example of Albert Einstein and his insight into the theory of relativity] As a scientist it is his job to bring everyone else along, to develop a language of mathematics and so forth. It is very similar in this case to religion. To have a mystical experience that is absolutely true and it is absolutely unquestionably true. The question is what to do with it. Do you hold it within, do you become a teacher? I am trying to get at this idea that there are certain
insights that have a very high level of certainty. They cannot be questioned or challenged. They don’t come about necessarily because you do a set of experiments and the experiments then … you say, “ah hah” They come about because of something more subterranean mechanism. Some of the best science does. Maybe the best science comes about this way. [GDF2, int]

There were only a four scientists who indicated there were differences in the certainty of the knowledge that was attributed to differences in complexity of the system involved (16.7%, 4 participants) and/or associated with the discipline (16.7%). These two features are not necessarily mutually exclusive. In response to questions about certainty of the model of the atom, one scientist, a theoretical physicist, said:

Very certain, in the sense that we have a theory that does a spectacular job explaining an extraordinary diverse set of observations and experiments. The mental picture we use of electrons orbiting and so on is perhaps less certain but it does not matter! There is a profound difference between a description that works and literal belief in the picture - you don’t need the latter in order to have a spectacular success of the former. [sTC4, vnos]

R: Do you think there are some areas of science that may be more certain than others?
S: Yes, certainly. There are some areas of physics especially; I think chemistry which is close to physics these days, where the level of certainty becomes very high because the system that you are studying is very simple. So when a chemist, a physical chemist, says they have a very complete understanding of what happens when two hydrogen molecules collide or when two simple molecules collide and a reaction takes place. I believe that because that is a repeatable experiment and it is also an experiment that can be matched with a computational machinery that comes from quantum mechanics... That doesn't mean it is absolutely true. That means there is a very high level of confidence .... In that kind of science, a lot of physics, not all of physics, but a lot of physics is in that category. A lot of chemistry. In that kind of science there is a very high level of certainty. In other areas, for example climate change or...can we decide whether we can predict Earthquakes or do we know when the Earth's magnetic field is going to reverse next... In those kinds of areas there is a low level of certainty and in some cases even the question is wrong, meaning that the correct question might be can we predict, not when.... So the question has to be removed somewhat to a level where you say, well, this is the kind of complex system
where you cannot even make a prediction, so you will never have perfect knowledge. Weather is also in that category. [sTC4, int]

In discussing his use of models and a specific example of resistance to model change within his community, a chemist also indicated complexity of systems as a reason for levels of certainty of the knowledge produced.

R: what leads to the model change?
S: well the very simple explanation is that the simplistic model is just wrong. .... so here is the model system and here is the real system, there has to be a continual feedback of you getting ideas out of the model system to explain what is going on out in the real world. And you identify problems, things that you don’t understand in the real world to give you an idea of what sort of model systems to set up. So this feedback is going back and forth. This particular issue is one that....just because of the complexity I think that the feedback got stuck and people got so comfortable with the model systems and forgot that it really doesn’t explain what is going out here in the complicated world. One of the reasons why I think it did get stuck, is that the issue of gathering experimental data for the complicated systems is also a large and expensive problem. It has made us all more comfortable with the model systems...You get social sciences and I don't think...social science is...the answers don't stay the same. People change their mind in the social sciences whereas in the physical sciences, the third law is still the third law...The idea is you start off with systems that are very simple and very well controlled and you rely on the fundamentals. Then there are good models and everyone will more or less agree on the results. Then you make systems increasingly more complicated and there is less and less agreement on what you are doing and less and less agreement on the results. It is a continuum. Some things are certain. Some things are not. Most things are sort of in the middle. In particular things that people care about are in the grey zone. [wEL3, int]

Discipline-based Comparisons

Table 4 presents results for discipline groups. The interesting finding from this category lies within the life sciences (LS) group and their tendency toward absolutist views in comparison to the group as a whole (Table 4). Four of the five total appearing in the “attain certain knowledge” subcode were LS. When comparing just group percentages, 40% of the LS fell within the “attain certain knowledge”
subcode as compared to 21% of the total group. In contrast, none of the physicists or chemists appeared in this subcode. Three of the four in the "approaching certainty" subcode were LS. There was not a noticeable difference in scientists who "affirmed change." That is, even though the LS showed more tendency toward absolutist views (attain certain knowledge), they were, nonetheless, as typical as the total sample with regard to describing scientific knowledge as inherently tentative. Three of the five physicists, and two of the four chemists fell within the "affirm change" category. The other two physicists indicated different levels of certainty determined by complexity of the system under investigation. These results show the LS were split in their views of tentativeness, demonstrating a range from absolutist views to views of inherent tentativeness with fairly equal frequency. In comparison, the physicists and chemists tended to make statements more demonstrative of "affirm change." The ESS participants were variable, but not to the extent as the LS. The ESS tended toward "affirm change" and "levels of certainty," with one ESS participant demonstrating absolutist views.

Approach-based Comparisons

When comparing by approach (Table 5), four of the five in the "attain certain knowledge" and all four in the "approaching certainty" were either E or E/D. These results suggest the life scientists within this sample who engaged in experimental or mixed E/D programs were more likely than the other groups to view scientific knowledge as absolute truth or progressing toward knowledge of that reality. The theoreticians were lower than the total group with respect to "attain certain
knowledge" (0% versus 21%), and higher than the total group with respect to “affirm change” (75% versus 21%).

Empirical

Tables 6 and 7 present the results for the scientists’ views of the empirical basis for scientific knowledge, compared by discipline groups and approach groups.

Scientists’ Views

Empirically grounded. Seventeen of the scientists (70.8%) acknowledged scientific knowledge as requiring an empirical basis. What constitutes empiracy may differ by science discipline:

Evidence for a chemical synthesis is far different from evidence for the evolutionary development of a biological feature; however, both areas demand the development of sequence of logical connections between the observable phenomena and the predictions of the hypothesis." [gEL3, vosi]

In discussing the difference between art and science, a life scientist voiced her view of the empirical nature of science as it relates to scientific method and certainty of the knowledge produced.

As a scientist, I would say that we use the scientific method to explore and understand the natural world...from the data we collect we attempt to develop a picture of that world...a picture that is as accurate as possible. Art, on the other hand, is sometimes accurate, but need not be so. [NEFL1, vnos]

This statement suggests that because science demands observations (data), the knowledge produced is more accurate than that produced void of observations, such as art. Not surprisingly, this participant also fell within the “approach certainty” category of tentativeness.
### Table 6. Scientists' Views of Empirical Basis: Grouped by Discipline

<table>
<thead>
<tr>
<th>Aspect</th>
<th>Subcode</th>
<th>Total</th>
<th>Grouped by Discipline</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>#</td>
<td>LS</td>
</tr>
<tr>
<td>empirical</td>
<td>empirically grounded</td>
<td>17</td>
<td>8</td>
</tr>
<tr>
<td></td>
<td>theoretical yet empirically grounded through confirmation of predictions</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>theoretical/mathematical: changing what science is... (string theory)</td>
<td>3</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>%</td>
<td>80</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>10</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>0</td>
</tr>
</tbody>
</table>

### Table 7. Scientists' Views of Empirical Basis Grouped by Approach

<table>
<thead>
<tr>
<th>Aspect</th>
<th>Subcode</th>
<th>Total</th>
<th>Grouped by Approach</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>#</td>
<td>E</td>
</tr>
<tr>
<td>empirical</td>
<td>empirically grounded</td>
<td>17</td>
<td>9</td>
</tr>
<tr>
<td></td>
<td>theoretical yet empirically grounded through confirmation of predictions</td>
<td>2</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>theoretical/mathematical: changing what science is... (string theory)</td>
<td>3</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td></td>
<td>%</td>
<td>90</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>0</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>10</td>
</tr>
</tbody>
</table>

E: experimental; E/D: combination of experimental and descriptive; D: descriptive (non-experimental); T: theoretical
Participants’ views of an empirical basis for scientific knowledge emerged in their comparisons of science and religion or philosophy:

[Religion] is very much based on faith. Science hopefully has less faith involved and the idea, is that it is more empirical and less bias. It is all based on questioning and testing your own thinking...Religion and philosophy are purely mental constructs. There is no data needed or wanted. Don't confuse me with the facts, you know...creation versus evolution. [fEF3, int]

The general practice of religion seems rooted more in faith and belief and some sort of super natural which is beyond the realm of science. It can't be observed. It can't be manipulated or tested. In that sense the practice of religion cannot be scientific. [UEL1, int]
Science deals with tangible/observable (cause/effect) things, while disciplines like religion and philosophy deal with intangible things that have little to do with cause/effect. [mDF1, vnos]

*Varies: theoretical/mathematical valid basis.* Three of the 24 (12.5%) indicated valid scientific knowledge can be acquired through purely theoretical or mathematical means. This position was related to the scientists’ view of mathematics or reliance on mathematics within their work.

Science and math are more or less the same. So if you’d said physical science is where I am, biological science is another branch, mathematical science is another. Three branches of science [wEL3, int]

[Comparing science and art] There are certainly elements of taste. There are elements of subjectivity. There are elements of fashion. There are those elements in physics too, especially recently with cutting edge theoretical physics. You can string along lots of ideas. Some of the ideas can be appealing because of the aesthetic elements... [In science] if something that is mathematically inconsistent or if something is inconsistent with a great wealth of data that is known about the physical world, you just don’t give it any further consideration. It is wrong. So that really is the difference. It is difficult to really pin down the way good science and art are different, except of course, the mathematical content could be brought in, the restrictive structures of mathematical consistency. But you can say there is bad science, but I don’t think you could say there is bad art. Right? ....[pTC4, int]

String theory was provided as an example of a non-empirical scientific field.
S: One of the things that characterizes physics is this enormous reliance on the definitiveness of mathematics. The clearest distinction I can make is that in mathematics, in mathematical research, there is no external motivation. You are fascinated by the mathematics in and of itself.

R: You use mathematics in your work. Is there ever a blend?

S: Oh yes, and sometimes it is not clear what you are doing. String theory is the best example. Sometimes the mathematics itself becomes the driving force, the beauty of the mathematics. String theory is much more mathematics at present than it is physics. They hardly ever talk about, they never talk about data. They are looking for mathematical consistency in patterns of symmetry of the theory. In the cutting edge of modern theoretical physics, it is getting more and more difficult to get data. You can’t do the experiment. String theory for example and particle physics beyond the standard model, require energies that are just not available in terrestrial machines. One is that the data are not...the data will be very difficult or in principle impossible to get, after all if some conditions only existed at the beginning of the universe, you just can’t duplicate that. Secondly the mathematics is so difficult that it may be extraordinarily difficult or maybe impossible to use mathematical restrictions, mathematical considerations themselves to nail down sufficiently the field of possibilities. Right now theoretical physics is in a strange state. String theory is like nothing else seen before...They may have it right, but how will we ever know. [pTC4, int]

**Discipline-based Comparisons**

The participants who explicitly stated mathematics and theoretical models generated valid, yet non-empirically based knowledge were theoretical physicists (2) and one analytical chemist. This is the only noticeable feature, as the majority of the scientists conform across the groups.

**Approach-based Comparisons**

Two of the three participants who reported science could be based on mathematics and theoretical models were theoreticians. This is the only noticeable feature, as the majority of the scientists conform across the groups.
Subjectivity

Tables 9 and 10 present results for the scientists’ views of subjectivity. The analysis identified different levels of understanding within this main category, from the individual’s personal bias to more sophisticated conceptions of theory-laden influences on scientific reasoning.

Scientists’ Views

Theory-laden. Fifteen of the 24 (62.5%) participants indicated a view of subjectivity that went beyond personal differences (e.g. reading the instrument differently). These scientists suggested their theoretical framework, or that of other scientists, guides the questions they ask, the investigations they conduct, and the interpretation of their data.

Without models, observation would amount to cataloging data. [models] Motivate the questions asked of the data, and thereby determine what data are going to be taken. [pTC4, vnos]

I don’t think there is any purely observational program. There were such programs 100 years ago, maybe even 50 years ago. But mostly now we have a framework of ideas because we don’t just say ‘let’s point the telescope there and see what we find” but we have some basic idea of what we are trying to find out about that is guided by more theoretical knowledge of what is going on. [pEDFC2, int]

Usually we are trying to get some estimate of the abundance of the arthropods. It doesn’t really matter to some extent if you are doing a comparative study. What you have to be careful of is that some of the traps have biases, like activity….You try to minimize the biases, but you know there are always biases because what you choose to minimize the biases is sort of bias too. [SEDF1, int]
<table>
<thead>
<tr>
<th>Aspect</th>
<th>Subcode</th>
<th>Total</th>
<th>Number of scientists</th>
<th>Grouped by Discipline</th>
<th>Percentage of group</th>
</tr>
</thead>
<tbody>
<tr>
<td>Subjectivity</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>personal only</td>
<td>9</td>
<td>10</td>
<td>4</td>
<td>40</td>
</tr>
<tr>
<td></td>
<td>theory -laden (guiding framework)</td>
<td>15</td>
<td>5</td>
<td>6 4 3 2</td>
<td>60</td>
</tr>
<tr>
<td></td>
<td>positive</td>
<td>8</td>
<td>5</td>
<td>5 1 1 1</td>
<td>50</td>
</tr>
<tr>
<td></td>
<td>negative</td>
<td>4</td>
<td>4</td>
<td>1 1 0 2</td>
<td>20 20</td>
</tr>
<tr>
<td></td>
<td>inconsistent response</td>
<td>3</td>
<td>4</td>
<td>1 1 0 1</td>
<td>20 20</td>
</tr>
<tr>
<td></td>
<td>differs with discipline</td>
<td>4</td>
<td>5</td>
<td>2 1 0 1</td>
<td>20 20</td>
</tr>
<tr>
<td></td>
<td>differs with approach: qualitative &gt; quantitative</td>
<td>5</td>
<td>4</td>
<td>1 1 0 3</td>
<td>10 20 0 75</td>
</tr>
<tr>
<td></td>
<td>not subjective</td>
<td>4</td>
<td>2</td>
<td>1 1 2 0</td>
<td>10 20 40 0</td>
</tr>
<tr>
<td></td>
<td>Scientific Method = objectivity (ideal or practice)</td>
<td>2</td>
<td>2</td>
<td>1 1 0 0</td>
<td>10 20 0 0</td>
</tr>
</tbody>
</table>
Table 9. Scientists' Views of Subjectivity: Grouped by Approach

<table>
<thead>
<tr>
<th>Aspect</th>
<th>Subcode</th>
<th>Total</th>
<th>Grouped by Approach</th>
<th>Number of scientists</th>
<th>Percentage of group</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>24</td>
<td></td>
<td>10 5 5 4</td>
<td></td>
</tr>
<tr>
<td>total</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>#</td>
<td>%</td>
<td></td>
<td>E E/D D T</td>
<td>E E/D D T</td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>37.5</td>
<td></td>
<td>4 2 2 1</td>
<td>40 40 40 25</td>
<td></td>
</tr>
<tr>
<td>15</td>
<td>62.5</td>
<td></td>
<td>5 4 4 2</td>
<td>50 80 80 50</td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>33.3</td>
<td></td>
<td>3 2 2 1</td>
<td>30 40 40 25</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>16.7</td>
<td></td>
<td>3 0 1 0</td>
<td>30 0 20 0</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>12.5</td>
<td></td>
<td>2 0 1 0</td>
<td>20 0 20 0</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>16.7</td>
<td></td>
<td>1 2 1 0</td>
<td>10 40 20 0</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>20.8</td>
<td></td>
<td>3 0 2 0</td>
<td>30 0 40 0</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>16.7</td>
<td></td>
<td>1 0 1 2</td>
<td>10 0 20 50</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>8.3</td>
<td></td>
<td>1 0 1 0</td>
<td>10 0 20 0</td>
<td></td>
</tr>
</tbody>
</table>

E: experimental; E/D: combination of experimental and descriptive; D: descriptive (non-experimental); T: theoretical
R: You say data is usually not evidence. Much data is ignored. How do you know what to ignore?
S: Well, we understand, for example, the atmosphere, we understand with the satellite observations where we have just a single layer of clouds present. When there are multiple layers of clouds present, you can’t make much sense of it. It is like unscrambling eggs. But we do know when we have multiple clouds present and we also know when we have a single layer. So we throw away the data for multiple clouds...and focus on those that we do think we understand. [cDFC2, int]

R: What makes people weigh evidence differently?
S: That is a very good question. Basically training and experience. It is your particular bias. Biases in science are typically best represented by how you weigh the evidence. For example, if you are a sedimentologist, you are very used to looking at the sedimentary record. You will believe evidence that is sedimentological over that which, for example, might be a computer simulation of the climate. On the other hand, if you are a computer modeler or a climatic modeler and spend a lot of time modeling the climate, you really believe this model...you might think the sedimentological data is not that important. [GDF2, int]

Scientists saw both positive and negative implications based on guiding frameworks. Positive features related to progress in science. Expectations and assumptions help scientists recognize anomalies:

R: You talked about the purpose of a model...You said simplification allows the user to focus on particular factors of interest, while of course, ignoring or holding other factors constant. My question is how do you decide what to focus on and what to ignore
S: There is the human bias. You focus on things that interest you. You focus on things you thought were going to be interesting or have some intuition or some prior information to know which parts of the system you can ignore.
R: Do you ever think you ignore something and come back later and say, "shoot, I shouldn't have ignored that."
S: Absolutely! All the time.
R: How do you realize you should not have ignored it?
S: Usually it is because the system starts to exhibit behaviors that you can’t understand anymore. That, you know, by making those types of assumptions, you say, "oh, if I tweek it this way, it should do that." You tweek it that way and it doesn't do that at all, you realize that it is sometimes because of assumptions you have made in part of the model you use to make those predictions. [MEDF1, int]
Negative implications related to funding sources and blocking abilities to identify and address alternatives:

R: Do you see the type of research being done, as in long extended projects vs short stint projects changing?
S: Definitely it is changing. I've had the benefit of being on long term projects. And now they are being pushed to short term projects. So [my current] study, I am still waiting to hear from NASA to see if there will be continued funding for it. I think it should be continued. I think there should be 10 of us studying it, not just me. But I don't know whether it will be funded. I certainly can't work on it without funding. I'm supposed to be working on other things with what funding I have and so if I show up working on [this other project]...that is where the budgets can get you. There is just not the freedom to explore areas that seem critical.

[in talking about the coral reef debate] They believe strongly that their process is the more important one. I sometimes wonder if it is a function of our system today in that people have their areas of research and scientists get funding. And I don't know if sometimes where you have this kind of debate, or whatever, or grouping of people into different camps is due to the fact that they have to make the case to receive funding for their research. If you get locked into thinking only one way, you can sometimes be blinded to other possibilities. [NEFL1, int]

Theories represent our model of the world and universe. Collectively, they define the world and probably restrict some of our thinking. For example, an observation that doesn't fit our collective model might be ignored, rather than being seen as the key to a new set of experiments to redefine our world view (for example, insects as regulators). [SEDF1, vnos]

R: What leads to dismissing ideas?
S: We tend to get very territorial about our ideas and we sort of go through this process of rationalization that makes us believe our ideas are right. It makes it very difficult to accept observations and data that are contradictory to our ideas. [pEDFC2, int]

Qualitative methods are more subjective than quantitative methods. Five of the 24 stated qualitative methods are more subjective than quantitative methods, and made reference to statistics as a means to reduce subjectivity.

S: We are all human so I don't think you can remove completely subjectivity or social and cultural values from science...Some science can be very political. There is a fine line there between collecting the data, interpreting
the data, and in my area...risk assessment. This is a very problematic area. It is very hard to be quantitative and not subjective about that. Find measures that communicate risk but do not communicate bias. It's a hard one.

R: Do you think there are some types of science that are more subjective than others?

S: I will bring my bias to the table. Some...forestry practices, for example, have the appearances of being supported by data, yet it doesn't seem to match. It is the idea that you can have the same data set and two different interpretations. It is the interpretation of the data set, the risk, the consequences, the ecological outcomes...the same set can go in two different directions....Without numbers, there is more bias. [IEF3, int]

_Science is not subjective._ Four suggested science is not subjective at all. One of these, a theoretical physicist, suggested theory-laden observations could lead to competing explanations of the same data, yet with sufficient data, the controversy would be resolved. In response to the question about different explanations for dinosaur extinction, a physicist responded:

They are essentially not using the same data. Because I suppose either group cannot explain all the data....One viewpoint...that is a good thing when you have two explanations for the same thing. The conclusion is that both of them need more work. One has to go into the other's territory and look. .....Part of the problem is that you tend to look under the light when you lose your keys. You look under the light because that is where it is brightest. So you tend to look for the explanations for things where you are most comfortable looking, rather than where the keys actually are. My impression there is that the fields are a little different and therefore the outlooks are different and they are not weighting all the data in the same way. They are focusing on certain aspects of it. I presume they both are. I will explain this because I can explain it this way and your theory can't touch it. [jTC4, int]

This participant went on to say, “The truth will be revealed with enough data.”

Two participants indicated that the “Scientific Method” was useful for maintaining objectivity within science. Both of these included a caveat that this was an ideal situation, and not always attainable.

I think a lot of science...the essence of the scientific method is that you shouldn't care. It is very transcendental. No ego. Very few people in science,
including myself, are able to do that. And so that is one of the things that defines how you weigh evidence. It may be subtle reasons. It fits with something more interesting. Your advisor likes it that way. Someone you don't like in science likes the other story. ...[GFD2, int]

This approach [scientific method] is supposed to provide objective unbiased answers. In my mind the conduct of science isn't as objective as many people would say. Sometimes there is room for interpretation and clearly scientists come from different backgrounds and experiences and therefore have different perspectives. [NEFL1, vnos]

In describing similarities and differences between art and science, a participant responded,

Science and serious art are both searches for truth. Both are cumulative in the sense that they are strongly influenced by what has gone before. They differ in terms of the tools that they use. Mathematics and technology are essential tools of science; art does not require either. Successful art need not be objective, quantitative, or reproducible. [eDF2, vnos]

**Discipline-based Comparisons**

Eighty percent (4) of the Earth and space scientists (ESS) held views of theory-laden NOS. This is greater than the 62.5% of the total group. Three of the 4 chemists indicated qualitative methods as more subjective than quantitative methods. That is, 3 of the 5 of the total within this subcode group were chemists. None of the physicists voiced this view.

**Approach-based Comparisons**

With regard to approach differences, 80% of the scientists within both the E/D and D groups fell within the theory-laden subcode. These two groups comprise over half of the total (8 of the 15). The distribution of the few participants within the other subcodes appeared similar, with no evident clusters.
Creativity

Tables 10 and 11 present results of the scientists’ views of the role of creativity in science. A majority affirmed a role of creativity, but they faltered somewhat in further explications.

Scientists’ Views

*Affirming creativity.* Sixteen (66.7%) of the total sample affirmed the involvement of creativity in the development of scientific knowledge. This was the most common response within the creativity theme. The following are typical responses related to creativity when participants compared art and science.

Both are creative; both have a mixture of individuality and community; both are conducted within the mind of the individual. The difference lies in repeatability...[PEDF1, vnos]

Noteworthy scientists and artists bring to bear superior imaginative and creative resources that leave the rest of us standing in awe. Whereas science often leads to new science, new understanding, applications that change lives, etc., good art simply leaves us standing in awe. [cDFC2, vnos]

In response to a question about where they saw creativity in science, several indicated “interpretation.”

Absolutely. This is what distinguishes genius from pedantic activities in science.....the one [stage of inquiry] that has probably the most imagination is interpreting the data. [bEL3, int]

You just look at the data and, you know, ...interpretation...I think I answered in some of your questions that there is a lot of art...in the interpretation there is a lot of creativity in how you choose to interpret the data, as well as in how you choose to design the experiment in the first place. [SEDF1, int]
Table 10. Scientists' Views of Creativity: Grouped by Discipline

<table>
<thead>
<tr>
<th>Aspect</th>
<th>Subcode</th>
<th>Total</th>
<th>Grouped by Discipline</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>total</td>
<td>LS</td>
</tr>
<tr>
<td>creativity</td>
<td>yes affirmed</td>
<td>16</td>
<td>6 4 3 3</td>
</tr>
<tr>
<td></td>
<td>epiphany only examples</td>
<td>3</td>
<td>0 2 1 0</td>
</tr>
<tr>
<td></td>
<td>progress</td>
<td>5</td>
<td>3 0 1 1</td>
</tr>
<tr>
<td></td>
<td>no data to be clear</td>
<td>3</td>
<td>1 1 1 0</td>
</tr>
<tr>
<td></td>
<td>inconsistent responses</td>
<td>1</td>
<td>0 0 0 1</td>
</tr>
<tr>
<td></td>
<td>other</td>
<td>1</td>
<td>0 0 1 0</td>
</tr>
</tbody>
</table>
Table 11. Scientists’ Views of Creativity: Grouped by Approach

<table>
<thead>
<tr>
<th>Aspect</th>
<th>Subcode</th>
<th>Total</th>
<th>Grouped by Approach</th>
<th>Percentage of group</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>creativity</td>
<td>yes affirmed</td>
<td>16</td>
<td>10</td>
<td>70</td>
</tr>
<tr>
<td></td>
<td>finding patterns/build connections</td>
<td>7</td>
<td>5</td>
<td>80</td>
</tr>
<tr>
<td></td>
<td>epiphany only examples</td>
<td>3</td>
<td>5</td>
<td>60</td>
</tr>
<tr>
<td></td>
<td>progress</td>
<td>5</td>
<td>4</td>
<td>70</td>
</tr>
<tr>
<td></td>
<td>data to be clear</td>
<td>3</td>
<td>4</td>
<td>20</td>
</tr>
<tr>
<td></td>
<td>inconsistent responses</td>
<td>1</td>
<td>2</td>
<td>25</td>
</tr>
<tr>
<td></td>
<td>other</td>
<td>1</td>
<td>1</td>
<td>10</td>
</tr>
</tbody>
</table>

E: experimental; E/D: combination of experimental and descriptive; D: descriptive (non-experimental); T: theoretical
Finding patterns and connections. Only seven (29.2%) voiced a view of creativity being specifically involved in finding patterns or building connections.

[Science involves] putting things together in ways not previously connected. [GDF2, int]

You get deluged with a bunch of facts, you have to sit back, you don’t worry about laws, theories, or principles, or anything. You day dream. You say, well...as I go through I sometimes find two different papers in the messy office and say, “wait a minute. There is an interesting connection.” That is largely what you doing as a scientist. So what you are talking about with theories and principles and hypotheses, is how I can convince my colleagues that I have thought through this well and it might be a physically meaningful principle. So basically what it comes down to is that science is a way of simplifying and expressing patterns and a way of testing whether or not those patterns make predictions in a way that other people can apply it and understand it. So that is what the real divergence is between how you do science versus what being in the profession of science is. Doing science is playing. It is a lot of fun. The profession of it is convincing other people that you’ve really done something, that it is not an artifact. That is important and they ought to pay attention to what you’ve done even though you are not paying attention to what they have done. It is a difference in creativity versus marketing and establishing something as a fact. [OEL1, int] [emphasis added]

Progress/Success. Five participants connected creativity to science progress or success as a scientist:

Not all scientists use creativity and imagination in their investigations, but it is my assessment that the best scientists do. Science only progresses if it moves beyond what is already known; a creative person is needed to perform novel research and to look at things in a fresh light. [mDF1, vnos]

Good scientists tend to be very creative and very imaginative and are able to make leaps far beyond what people can calculate. Some of these are leaps of faith. Then they start seeking the evidence. Kind of like knowing the solution and working backward to the starting point. [He talks about the ozone data a scientist had for 30 years; relates to cloud data; curiosity lead her to “take this data nobody else wanted to look at for 30 years and ah-hah there it was...all this information on the problem with the ozone”] [cDFC2, int]
Ego drives you to be creative. Science would be really boring if everybody just sat back and were truly logical and dispassionate about it. [He continues with an example of an epiphany he had.] [pTC4, int]

There was evidence of connecting views of subjectivity and creativity in the following excerpt. This participant uses a story about his young son’s creation of a train from puzzle pieces to describe how science progresses.

So you have to weigh the evidence and experiments and consider what...does this really disprove what I am thinking of or is there another explanation? If there is another explanation, what experiment can be done to test that? I view it in like building a puzzle. It is hugely complex and you don't have a picture of what the overall thing is by looking at individual pieces. You push them together and you build different things. ...[tells story of his son linking puzzle pieces together and calling it a choo-choo train]...So when you've got a lot of different pieces, you put them together, you build a theory. You build your own choo-choo train and you are pretty enthusiastic about it and your grants get funded. There is a lot of momentum behind this train. The idea that you have to take the pieces apart and look at them, and maybe rearrange them a bit to make them look a little different is very hard to do. You have so much involved intellectually and financially in getting this train moving in the first place. You've to hope it's not a blind track you are running into in a train wreck where all the pieces are going to fall off when you take it apart. But if you can take off one piece and bring it around to another and get a different view of the puzzle. That is perfectly fine. I think that is the way science evolves... In biology in particular we've got to be particularly careful in being willing to take the pieces apart and put them back together in different directions. [OEL1, int]

No role in making meaning. There were only three participants to suggest creativity does not play a role in making meaning of data. Typical among these responses were that creativity is involved in the initial stages, such as design, but not in interpretation.

S: Creativity periods are in the beginning, when we had these great observations of El Nino but the models couldn’t explain it. Then we got to the point where we could explain the historical record, did a great job on that, then a new one came and it didn’t do a good job. We realized there were different regimes in the ocean....So the most creative work has happened when those periods where everybody just doesn’t understand.
R: How about data analysis?
S: Uh...you can be a creative data analyst, searching for patterns in things. I don't think that is necessarily science...You are not scientific if you let beliefs interfere with the assessment of data. [eDF2, int]

It is very rare, if ever, that a scientist repeats the work of another scientist. Therefore, every step of research is an innovation. ...I do not work directly with data, so cannot give examples of that type. Others working in my field do take data, and the act of creativity is focused on the design of the device to take the data. [pTC4, vnos]

**Discipline-based Comparisons**

In comparison to the total, ESS scientists more frequently indicated the role of creativity in all stages (4 of the 5 ESS), and in finding patterns and building connections in particular (60% of the ESS, compared to 29.2% for the group). In general, the ESS affirmed a role of creativity in finding patterns and building connections slightly more so than the other discipline groups. However, the overall frequency of this subcode is quite low, making a context-based relationship very weak, at best.

**Approach-based Comparisons**

When comparing based on approach, there is slight tendency for those scientists involved in E/D or D methods to connect creativity with pattern recognition (4 of the 7). None of the theoreticians fell within the subgroup of “finding patterns/build connections.”

**Socio/cultural Influence**

Tables 12 and 13 present results of scientists’ views of social and cultural influences on science. There were three main clusters of subcodes. One focuses on the
reasoning processes, or the how science is conducted. One focuses on the questions that are asked, or what science is conducted. The third focuses on the products of science.

Scientists’ Views

Political, economic, societal pressures on what questions are probed. In general there were few occurrences within this main aspect other than the view that society and culture influence what science gets done because of political/economic/societal pressures that influence funding and other agencies that support the institutions of scientific research. Fifteen of the 24 scientists (62.5%) voiced this view. The following quotes are representative. They also demonstrate a blending of views that fell within other subcodes.

Certainly in the environmental you get sort of a scientific answer and you get a societal value and you get politics layered on top of that and you get uh...there are heaps...at least in the environmental area you are heaping all of these human factors on top of the scientific answer and then people try to support whatever it is they want to do for whatever reasons with the scientific argument. [wEL3, int]

That is an interesting question. I think here the science we do is influenced by politics. And in a rather devious way, through budgets. How they get you in the end is in the budget. They just don’t put the research money in. There is a political effort out now to stop certain areas of research, like gene cloning. Some areas are out to stop it and somewhat succeeding. But we don't have the types of things that happened in the 19th century, well 20th century Russia for example where you had to hold to certain theories to be allowed to be a scientist. There are some third world countries where you have to tow the political line...things that are clearly obvious to scientists...where if you said the obvious you'd be in trouble. Things like pollution, for example, have to be ignored. But here, society is very much a thing to be recanted with, I'd say. ....[cDFC2, int]

Science reflects social values...what science receives funding. However, once a problem is identified, proper use of the scientific method should lead to unbiased results. Nevertheless interpretation of these results may under some circumstances reflect personal biases e.g. gender and racial bias that colored interpretations of brain sizes. [NEFL1, vnos]
Table 12. Scientists’ Views of Socio/Cultural embeddedness: Grouped by Discipline

<table>
<thead>
<tr>
<th>Aspect</th>
<th>Subcode</th>
<th>Total</th>
<th>Grouped by Discipline</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total</td>
<td></td>
<td>24</td>
<td>10 5 5 4</td>
</tr>
<tr>
<td>Number of scientists</td>
<td></td>
<td></td>
<td>LS ESS Ph Ch</td>
</tr>
<tr>
<td>Percentage of group</td>
<td></td>
<td></td>
<td>% % %</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>LS ESS Ph Ch</td>
</tr>
<tr>
<td>Questions/problems studied (What)</td>
<td></td>
<td></td>
<td>% % %</td>
</tr>
<tr>
<td>human endeavor</td>
<td></td>
<td></td>
<td>% % %</td>
</tr>
<tr>
<td>s/c influenced: political, economic, societal pressures</td>
<td></td>
<td></td>
<td>% % %</td>
</tr>
<tr>
<td>not s/c influenced</td>
<td></td>
<td></td>
<td>% % %</td>
</tr>
<tr>
<td>products</td>
<td></td>
<td></td>
<td>% % %</td>
</tr>
<tr>
<td>universal application/generalizability</td>
<td></td>
<td></td>
<td>% % %</td>
</tr>
<tr>
<td>not universal use</td>
<td></td>
<td></td>
<td>% % %</td>
</tr>
<tr>
<td>SM: not s/c influenced</td>
<td></td>
<td></td>
<td>% % %</td>
</tr>
<tr>
<td>other</td>
<td></td>
<td></td>
<td>% % %</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Socio/cultural process/conclusions (how done, reasoning)</th>
<th>influence on reasoning through processes, assumptions</th>
<th>4 16.7</th>
<th>3 0 0 1</th>
<th>30 0 0 25</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>not s/c influenced: international community, technology</td>
<td>5 20.8</td>
<td>3 0 2 0</td>
<td>30 0 40 0</td>
</tr>
<tr>
<td></td>
<td>external: political and economic pressures</td>
<td>3 12.5</td>
<td>1 0 1 1</td>
<td>0 0 1 1</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1 4.2</td>
<td>0 0 0 2</td>
<td>1 0 0 2</td>
</tr>
<tr>
<td></td>
<td></td>
<td>3 12.5</td>
<td>0 1 3 0</td>
<td>0 20 60 0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>15 62.5</td>
<td>6 4 2 3</td>
<td>0 20 40 0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>4 16.7</td>
<td>0 1 3 0</td>
<td>0 20 40 0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>3 12.5</td>
<td>0 1 2 0</td>
<td>0 20 40 0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1 4.2</td>
<td>0 0 0 1</td>
<td>0 20 40 0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2 8.3</td>
<td>1 1 0 0</td>
<td>10 20 0 0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>4 16.7</td>
<td>1 1 1 1</td>
<td>10 20 20 25</td>
</tr>
<tr>
<td>Aspect</td>
<td>Subcode</td>
<td>Total</td>
<td>Grouped by Approach</td>
<td>Percentage of group</td>
</tr>
<tr>
<td>--------</td>
<td>---------</td>
<td>-------</td>
<td>---------------------</td>
<td>---------------------</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Number of scientists</td>
<td>E</td>
<td>E/D</td>
</tr>
<tr>
<td></td>
<td></td>
<td>24</td>
<td>total</td>
<td>10</td>
</tr>
<tr>
<td>Socio/cultural</td>
<td>process/conclusions (how done, reasoning)</td>
<td>24</td>
<td>total #</td>
<td>%</td>
</tr>
<tr>
<td></td>
<td>influence on reasoning through processes, assumptions</td>
<td>4</td>
<td>18.7</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>not s/c influenced: International community, technology</td>
<td>5</td>
<td>20.8</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>external: political and economic pressures</td>
<td>3</td>
<td>12.5</td>
<td>3</td>
</tr>
<tr>
<td>Questions/problems studied (What)</td>
<td>human endeavor</td>
<td>1</td>
<td>4.2</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>s/c influenced: political, economic, societal pressures</td>
<td>15</td>
<td>62.5</td>
<td>8</td>
</tr>
<tr>
<td></td>
<td>not s/c influenced</td>
<td>4</td>
<td>16.7</td>
<td>0</td>
</tr>
<tr>
<td>Products</td>
<td>universal application/generalizability</td>
<td>3</td>
<td>12.5</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>not universal use</td>
<td>1</td>
<td>4.2</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>SM: not s/c influenced</td>
<td>2</td>
<td>8.3</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>other</td>
<td>4</td>
<td>16.7</td>
<td>1</td>
</tr>
</tbody>
</table>

E: experimental; E/D: combination of experimental and descriptive; D: descriptive (non-experimental); T: theoretical
Several participants indicated needing to be attentive to the agendas of funding agencies in order to be successful as a researcher.

You try to match your research idea with the funding agency. Sometimes you put a spin on it, or rationale that will convince the funding agency you want to work on a problem here that has implications to these problems the agency has interest in. ...It is stakeholder-driven research. Just doing research for research sake is gone...Funding is shorter duration now. There is a desire for results faster now...You have to be flexible with what you are willing to research. Be more dynamic and responsive or else you are not going to make it. [IEF3, int]

Socio-cultural influences on the processes of science (how science is conducted and reasoning). Only four participants fell within this subcode. For example,

I think it [science] is absolutely performed within a cultural context. Scientists are human beings. Humans are part of their culture and scientists will certainly operate, we can pull examples up through history on how people’s ideas are embedded in the culture they live in. There is no way around that. We can try as individuals to pull ourselves out and be objective, but there is no way. European science is very different from US science. There is big cultural differences there, certainly in the environmental sciences. Chinese science is even more different than US science. You look at also cultural differences within disciplines. Ecological sciences are very very different from chemical science. Applied science is really different from basic science. [BELF1, int]

A molecular biologist discussed his views of the international science community and differences he has seen in how science is conducted.

R: Do you see culture as influencing science? Do you see science as being universal? Can it be both?
S: Absolutely. The whole idea of science is being able to convince someone else that the underlying principles such as general applicability and basis in reality, and so that is what everybody agrees on. There are rules and if they understand the concepts they will see where it takes. That is the beauty of dealing with different cultures. Uh...they will take what you say and interpret it slightly differently that lead to new insights and go in different directions. It will reveal different aspects of it. That is where the communication is really
important. I think [unclear] globalization. In fact I deal with people all over the world. That is one of the great benefits of working in this place.

R: Working with people from all over the world, do you see any differences based on culture in how science is done?

S: Sure. Very much so. Americans are extremely competitive and hardworking and driven, pushing at the latest technology. They follow trends and fashions. In fact if you don’t follow the latest trends you get penalized in your grant reviews because you are doing something that is old and everybody knows it. In the European system it takes forever to get established and the probability of succeeding is ...once you are there you’ve got great freedom. You can work on things that interest you to a great extent but you don’t have to be competitive and right at the cutting edge. I really like South American scientists because they can’t afford to buy a box of reagents. They think about their experiments a lot more. My friends there are a lot more philosophical. At first I am uncomfortable listening to what they are talking about because it isn’t real rigorous, but you have to sit and listen and suddenly the brilliance comes through, that they have sat back and gotten at the bigger picture. And the scientists from Japan are a lot more difficult to understand in a lot of ways. A lot of them do get very focused toward whatever their major professor or director of the institute is focused on. They tend to work very hard to get a very high level of skill on some particular problem. But often times they are looking at one tree in the forest and have no idea of what the forest looks like. But there are exceptions and they are very important. They also have a different philosophy in how they make their judgments in what science is.

R: What do you mean?

S: I don’t really know how to explain that, but the thought process is different in many ways. [OEL1, int]

Within this subcode, there were views suggesting influences differed depending on the discipline in reference. The following interview excerpt demonstrates one participant’s view that socio-cultural influences varies depending on the type of science.

S: Natural sciences tend to be universal.

R: Natural sciences?

S: This would include mathematics, physics, and chemistry. What we know about the world, or let's say the questions we pose and what we choose to explore about nature in the natural sciences, though, are usually determined by cultural and political forces....I am often quoted by my students, I believe there are only two things that enable science. That is either a desire to gain economic advantage or a desire to gain political advantage. If you look at the history of science, these are the forces that drive science. The third
component... is intellectual curiosity.... Biological sciences tend to be both [universal and culturally embedded]. Investigations in biology usually are very very cultural and driven by society.... just look at the issues faced right now. Cloning, the issues of stem cell research because we can only get them from embryos. These are tied into very very divisive issues.... Often times investigations in previous times were extraordinarily controversial around the time of Darwin. The whole idea that humans evolved from monkeys, uh-uh. Even today is still not universally accepted. So these are very much involved with culture and society. .... The natural sciences people don't get so excited about. .... The social sciences, just by the nature, it is reflected in the name 'social sciences' of course, are social and cultural. There is not question of it. The questions asked, and thus the approaches to scientific inquiry can definitely reflect social and cultural issues, but the outcomes must be universal if they are to qualify as science. This is why often natural scientists, like myself, will say the social sciences aren't really sciences. Sometimes we... just the process of investigation interferes with the thing you are observing in such a way that one has to question the validity of the observation. So you can often times get a room full of psychologists and they will disagree about how a person is behaving. They may not, I don't know. But it gets difficult for them to meet this standard of repeatability..... When you get into the social sciences, it is really hard for us to become totally objective. We can't escape our biases. That isn't meant to demean those sciences... So a scientist in this country would pursue one of these investigations very differently than say someone even in Europe, which is very close to our society. But somebody in India or Sri-Lanka or Africa or Asia, may look at this totally differently. And the kinds of questions that occur to you or me to ask may never occur to them to ask. 

R: Versus someone in, say, chemistry?
S: In chemistry, we can take this thing and this thing and put them together and say, "oh, it turned blue." Well, I don't have to be Asian or European or African to see that it turned blue, assuming I am not colorblind. So I can ask the question, why did that turn blue? That is not...a cultural issue is not going to come into play. [bEL3, int]

This exchange represents connections between views of subjectivity, reproducibility, and socio-cultural embeddedness. With regard to her own research, astronomy, one scientist felt there was no impact from society because the discipline does not directly effect society. However, she sees socio-cultural impacts within other cultures and sciences:

I think that science itself is universal, but because science is a human activity that has impact on the lives of people. It is affected by cultural values and
beliefs. Because astronomy has little direct impact on our lives, the work of astronomers isn’t much affected by cultural values (except in terms of access to technology to carry out our work...) ...Scientists whose work is reflective of cultural values (e.g. African scientists who maintain the AIDS is not a sexually transmitted disease, or paleontologists who maintain that the fossil record shows the world was created in 4000 BC because of their religious beliefs) are generally marginalized because their work ignores a large body of evidence. [pEDFC2, vnos]

No effect of society or culture on what science is done. Four scientists voiced a position that science is universal, or at least, is ideally universal. They stressed that regardless of current beliefs, with sufficient information, the products of science would be the same regardless of the investigator.

Science is done by humans, and humans live in a particular age and are heavily influenced. Perhaps scientists in the 19th century were heavily influenced by belief in ESP, ghosts, and spiritual powers, and perhaps in the current era we are too heavily influenced by cultural perceptions and ideas. Science should ultimately develop a universally applicable framework, which provides a broad and deep explanation of the world around us. ....Ultimately, the truth, rather than our opinions, will prevail. [jTC4, vnos]

Effect of an international community on minimizing socio-cultural influences.

An interesting category emerged in reference to the effect of international collaborations and technology on minimizing cultural differences.

In some sense, social values are important because funding from research is largely publicly supported. In another sense, science is universal because it seeks to understand enduring aspects of nature independently of current social and political customs. Moreover, scientists have a long history of international cooperation...[UEL1,int]

I work with so many different people from so many different places that I am lead to believe that science is universal. [jTC4, int]
Discipline-based Comparisons

Within the largest subcode of “political, economic, societal pressures on the questions asked (what science is done),” there was little apparent difference among the disciplines. Interestingly, 80% (4 of the 5) of the ESS group appeared within this subcode, with very little representation within other subcodes of this main category.

Approach-based Comparisons

The theoreticians were underrepresented within the largest cluster of “political, economic, societal pressures on the questions asked (what science is done)” (1 of the 4 in the T group; 1 of the 15 total). The remaining 3 theoretical physicists fell within the “not socio-culturally influenced” subcode with respect to what science is conducted. They represented 3 of the 4 total within this subcode. Even though this view, as well as the subcode of “not s/c influence” with respect to how science is conducted, was not widely apparent within the total group, those engaging in a theoretical approach were more likely than others to consider science void of socio-cultural influences.

Theory/law

Tables 14 and 15 present results for scientists’ views of scientific theory and law. The subcode of “different” has three nested subcodes to indicate respondents within those areas thought theories and laws were different for one or more of the reasons cited.
Scientists' Views

Hierarchical relationship: Theories turn into laws with sufficient time and/or repeated testing. Thirteen (54.2%) of the participants held a hierarchical view of the relationship between scientific theories and laws. Statements representing this view included ideas of certainty, whereby theories are less certain than laws.

A theory is when we're pretty sure about something but not so sure that we aren't leaving room for a little doubt; there is no doubt in a law. [mDF1, vnos]

S: Hypotheses die every day. Theories being higher than hypothesis. Theories are longer lived. Laws have been around for a while....I mean a theory tends to be very broad in scope. It is very very basic.
R: And few things make it to theory?
S: Yes, and even fewer make it to law. Like laws of thermodynamics.
R: Do you deal with laws in what you do?
S: We use laws of thermodynamics. They govern everything.
R: Can you give me an example of a theory that drives what you do?
S: We don't really work at that level. ...we have hypotheses related to our work. We work less with theories. I think of theory as something that has a lot of weight and has multiple emerging lines of evidence and are very basic in nature, like global warming is a theory. It is not a law. It is very much a theory though because it is in flux.
R: At what point would you get a law?
S: It has to stand the test of time. I think theories are constantly being tested. Things that make it into law are the test of time and the preponderance of evidence and debate are in that direction. .....Theories move into law over time. [fEF3, int]

Several participants suggested the conventions of the scientific community establish when a theory becomes a law.

There are several degrees of certainty in science. The first level is an hypothesis that can be tested as a way of understanding the cause of observed phenomena. If the hypothesis continues to be supported by various studies, it eventually reaches the status of a theory. Continued support of the theory eventually elevates it to the status of a law. Of course, at any time, a study producing data inconsistent with even a law can require that it be modified or replaced, but the likelihood of this decreases from hypothesis to theory to law. [SEDF1, vnos]
Table 14. Scientists' Views of Scientific theory and law: Grouped by Discipline

<table>
<thead>
<tr>
<th>Aspect</th>
<th>Subcode</th>
<th>Total</th>
<th>Grouped by Discipline</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>total</td>
<td>LS</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Theory/law</td>
<td>hierarchical (theories become laws)</td>
<td>13</td>
<td>5</td>
</tr>
<tr>
<td></td>
<td>different</td>
<td>8</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>levels of confidence differ: laws &gt; theories; foundations</td>
<td>5</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>theories more general &amp; complex/laws simple</td>
<td>3</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>theories more likely to change</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>no difference</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>differs with discipline</td>
<td>6</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>no laws in field of work</td>
<td>6</td>
<td>3</td>
</tr>
<tr>
<td>Aspect</td>
<td>Subcode</td>
<td>Total #</td>
<td>Total %</td>
</tr>
<tr>
<td>---------------------------------------------</td>
<td>----------------------------------------------</td>
<td>---------</td>
<td>---------</td>
</tr>
<tr>
<td>Theory/law</td>
<td>hierarchical (theories become laws)</td>
<td>13</td>
<td>54.2</td>
</tr>
<tr>
<td></td>
<td>different</td>
<td>8</td>
<td>33.3</td>
</tr>
<tr>
<td></td>
<td>levels of confidence differ: laws &gt; theories; foundations</td>
<td>5</td>
<td>20.8</td>
</tr>
<tr>
<td></td>
<td>theories more general &amp; complex/laws simple</td>
<td>3</td>
<td>12.5</td>
</tr>
<tr>
<td></td>
<td>theories more likely to change</td>
<td>2</td>
<td>8.3</td>
</tr>
<tr>
<td></td>
<td>no difference</td>
<td>1</td>
<td>4.2</td>
</tr>
<tr>
<td></td>
<td>differs with discipline</td>
<td>6</td>
<td>25.0</td>
</tr>
<tr>
<td></td>
<td>no laws in field of work</td>
<td>6</td>
<td>25.0</td>
</tr>
</tbody>
</table>

E: experimental; E/D: combination of experimental and descriptive; D: descriptive (non-experimental); T: theoretical
During the interview we discussed his views further:

S: E=mc² was initially a hypothesis. Now it is a theory of relativity. Sometime it will be the law of relativity. It wasn’t initially the theory of gravity. It was a hypothesis. That is a law now. Everything has continued to support that. At some point, I don’t know if there is a rule for when things become a law. I think it is sort of by convention maybe somebody starts calling it a law and people just agree.

R: Do you all just vote or what?
S: I’ve never gotten to [laughs]. That may be something. There are conferences of zoologists for instance who establish rules for scientific nomenclature and things like that. There may be some body of physicists that meet to decide that this theory has become...reached the status of a law, get some agreement on that. Biologists fight among each other too much. The physicists work together... [SEDF1, int]

Another participant was asked for an example of a change in status.

S: A law, I consider to be something that is at least at the time and space in which it is generated, it is an invariant statement about nature and how it works. A law is basically a statement about fundamental behavior of the world. A theory is an idea supported by both empirical observations and by reference to other accepted theories and ideas that represents a statement about nature. For a law is in a sense a very very established theory. That is what I would say. I recognize a law in one generation may turn out to be a theory...may lose that status in another. But that is a rare phenomenon.

R: Do you have an example of that happening?
S: Well, for example, I mean, the way that Newtonian mechanics was a sense found to be only a special case on Einstein’s theory of relativity.....that was sort of......well they didn’t lose their status of laws, but they somewhat, well they lost some of their status as the...the broadest overarching statement you could say about....they were subsumed by something that came after. ....[tries again to find an example...unsuccessfully]

R: Are there cases where theories have become laws?
S: Oh Yeah...I think....uh....uh....uh....[long pause]......

R: First of all, what would make that happen?
S: uh...I think if the theory is found to be substantiated by multiple lines of evidence, repeated by multiple people over a sufficient length of time.

R: What makes it sufficient?
S: Well, that it...it is all a level of agreement. It is all in the agreement...it is all in your view...what we agree on changes. We agreed that the Earth is the center of the universe and then at some point somebody came up with a theory that said it is not. There were enough other observations that could be fit in that new theory that were not explained by that other theory, or maybe it just fit ocam’s razor, that it was simpler way of describing something. Plate tectonics is another one, it is a funny one because it has been a theory for a
long time... There is something there, you know, with how the continents align, but there was no mechanism; no one could figure out how it happens. So is plate tectonics a theory? I think it is more than a theory. It is a principle. Now whether it is a law in the same way that force equals ma is a law, I am not sure. There is a certain complexity, but are the continents moving? Yes they are moving. You cannot explain the forms of the world without that. It doesn't make sense. Something which has that quality to it, I call a law. I move it out of the realm of theory...

His examples and attempts to find examples are consistent with his view of difference in power in explanation based on certainty. His views are consistent. As ambiguity decreases, certainty increases, as such, status increases. He views laws as having higher status compared to theories. However, he couldn't solidify an example of change in status in either direction. This discrepancy was dismissed as simply not making sense. The status change should happen. It is as if the rest of the community has not caught up the times yet. As indicated in the previous quotes, several participants expressed similar views. Their view is that there must be some point at which the "community" agrees to make the switch. Yet they themselves have not been involved in making any decisions about status changes. Even though they cannot provide examples, they maintain that the community decides.

Others stated they didn't necessarily see that laws were applicable, again due to their view of laws as infallible. They saw a hierarchical relationship, but laws as truly unlikely. Many of the statements were critical of conventions of labeling theories and laws and showed preference for calling everything a theory that had an evidentiary basis.

I guess in certain language, certain uses there is nothing wrong with it [law]. But a law really is a theory. It is just a theory that people have been using for so long that people have stopped calling it a theory. Again I don't believe in certainty at all in science. I don't believe certainty exists. I believe we approach certainty as a limit, but we don't get there. I guess you could use a
law for something known, like gravity. It exists. Gravitational theory, we refer to gravitational theory, but it is a law. It is as close to a law as one could get. You drop a ball and it falls to the ground, and we know why. We know those masses attract. In my own science, very very rarely actually deal with things on that level with certainty so it doesn't come up much. [BEF1, int]

Some scientists suggested calling certain theories “laws” in light of their centrality to the field and acceptance through repeated confirmations.

I would classify "atomic theory" as a law, in that it is as well confirmed and as central to the field. [gEL3, vnos]

In discussing whether laws change, one scientist responded, "Yes, obviously, because a law is nothing more than a theory dressed in fancier clothes."

*Different types of knowledge, not hierarchical.* Eight (3.3%) saw theories and laws as different and not necessarily hierarchical:

A law is a well established phenomena that can usually be described with a few words or single equation. An example would be the laws of conservation of energy or matter. Theories are more general and more complex statements of a set of principles and predictions about some natural phenomena. For example, a well established theory is evolution by natural selection. This can be described briefly but the theory is very wide and deep encompassing many topics and many diverse lines of evidence. [eDF2, vnos]

Theories tell us among all the possibilities in the world, what possibilities are more probable. Theories change as better approximations of nature (models) evolve. Theories encompass new complexities. Scientists develop theories to generalize so that science isn't merely a collection of experiments, observations and data with no connections. Theories also become tools for human applications. [PEDF1, vnos]

This position also included describing theories and laws as different in terms of their generality and complexity. In the following exchange, the physicist describes theories as having solid evidentiary basis. This was a typical view of theories among most of the participants. He continues then to describe theories and laws in terms of levels of complexity and role in explaining phenomena.
S: Theory we mean something exact, something complete, something we think is the final answer. So for particle physics there is a theory now known as quantum chromodynamics. It is a field theory. To solve that problem requires exchange of 16 different particles simultaneously. So it requires hundreds of equations to be solved simultaneously, and they are integral equations. That has taken years of computer time for most elementary, even models there, how to solve that. But in theory one has a complete mathematical description. In practice you say let's model it by limiting the number of particles. That makes it a model...

...The theory is the Bible; the law is the Ten Commandments. You know, you'd say there often are a few basic laws which express a lot of what is in the theory, basic principles you might call a law. And if you look at it more abstractly, many of what we call the laws of nature are conservation laws, at least in physics. They are a direct consequence of symmetry in the world. ....

R: Does every law have a theory attached to it?
S: That is a good question. You can find a law of nature which is an experimental law, and then it is often the job of the theory to explain why that law is true. So if you have conservation of charge, there has to be some theory telling you why charge is conserved. That case is very hard to explain what that is but it means the theory has to be a gauge theory and it has to be from thermodynamics...it builds up. One of the consequences is that law...Theories don't have to be simple. Laws tend to be simple. Like the Ten Commandments. Those are the basic laws...

R: Do you think laws can change?
S: A physicist always has an open mind. If there is a law, it means, this is what experiment has showed us it holds. Then what typically happens it only holds in the realm of validity where the measurements have been made. It might be that the measurements, if they are made at the atomic or subatomic level that the law no longer holds. So you don't say the law doesn't hold, or is no good, you say it doesn't apply in this realm...[ITC4, int]

_Differs by discipline._ A few of the scientists indicated a difference in applicability of theories and laws, based on scientific discipline (6 participants).

R: Do biological sciences have laws like the examples you have given for physics or chemistry?
S: Yeah, in fact I think there are more laws, at least from the way I look at it, than biologists believe, and they just haven't learned how to use the information that is available from chemistry and apply it to biology... [OEL1, int]
Most theories and laws involve physics and chemistry, both of which I know little about. In biology, when you involve the element of life (neurological function and complex gene sets) there are always surprises and no laws, therefore only theories. [mDF2, vnos]

_No laws in field of research._ Six participants specifically stated that laws did not apply to their particular field of research. Most of these related these views to the idea of laws representing “certainty” and that within complex systems, certainty was not attainable.

We have no laws on atmospheric sciences. Things are too nebulous to really speak of laws. Everything is almost a law, approximate. We certainly use the second law of thermodynamics. That is one of our backbones. The other one is conservation of energy. The other one is conservation of momentum. These are laws. So we use them [cDFC2]

Very interesting questions. In my field, there is no such thing as scientific law. Law somehow implies infallibility, which is impossible given the bizarre way that science works.....nothing is a certainty in science, making ‘laws’ unrealistic. [MEDF1, int]

**Discipline-based Comparisons**

With regard to “hierarchical relationship,” there were no apparent differences in discipline groups’ views in comparison with the total sample. The most interesting feature of the discipline groups is the physicists and chemists lack of response toward “differs with discipline” and “no laws in field of work.” The respondents within these subcodes were from the LS and ESS groups only. This finding may serve to reinforce the idea that many researchers within all disciplines give higher status to laws (as evidenced by the 54.2% of the total group holding a hierarchical view). Achievement of this status may be associated with the scientists’ view of attainment of certainty within the discipline. The results for tentativeness also add
support to this notion. Sixty percent of the ESS and the physicists indicated a non-hierarchical difference in theories and laws, but these groups did not cluster differently from the total sample with respect to the three subcodes related to “different.”

**Approach-based Comparisons**

Approach-based comparisons suggest the E/D group tended to cluster slightly more within the hierarchical subcode than the other approach groups (4 of the 5 E/D participants). The theoreticians did not report “no laws in field” or that theories and laws “differ with discipline.”

**Observation and Inference**

Tables 16 and 17 present results of scientists’ views of observation and inference.

**Scientists’ Views**

Only 58.3% (14) of the scientists had responses that affirmed a role of inference in science. This may be a function of the instrument, as the responses from the other 10 scientists did not provide a clear indication relative to this aspect. Statements indicative of “affirming inference” tended to be subtle and contextualized within remarks related to other topics.

The facts of science go beyond just experimental data. The facts of science involve interpretation of experimental data.....may call a theory. A fact of science is a logical structure, put together from experimental data with some interpretation of analysis. A fact of science is that energy is conserved. A lot of information goes into that. [kEL4, int]
Table 16. Scientists' Views of Observation and Inference: Grouped by Discipline

<table>
<thead>
<tr>
<th>Aspect</th>
<th>Subcode</th>
<th>Total</th>
<th>Grouped by Discipline</th>
<th>Number of scientists</th>
<th>Percentage of group</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>24</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>10</td>
<td>5</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>5</td>
<td>4</td>
</tr>
<tr>
<td>Obs/inference</td>
<td>affirmed role of inference</td>
<td>14</td>
<td>14</td>
<td>58.3</td>
<td>6</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>4</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>not clear from responses</td>
<td>10</td>
<td>10</td>
<td>41.7</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>4</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1</td>
<td>1</td>
</tr>
</tbody>
</table>

Table 17. Scientists' Views of Observation and Inference: Grouped by Approach

<table>
<thead>
<tr>
<th>Aspect</th>
<th>Subcode</th>
<th>Total</th>
<th>Grouped by Approach</th>
<th>Number of scientists</th>
<th>Percentage of group</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>24</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>10</td>
<td>5</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>5</td>
<td>4</td>
</tr>
<tr>
<td>Obs/inference</td>
<td>affirmed role of inference</td>
<td>14</td>
<td>14</td>
<td>58.3</td>
<td>6</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>3</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>not clear from responses</td>
<td>10</td>
<td>10</td>
<td>41.7</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1</td>
<td>1</td>
</tr>
</tbody>
</table>

E: experimental; E/D: combination of experimental and descriptive; D: descriptive (non-experimental); T: theoretical
The conclusion can often be very abstract or subtle. So you might be trying to find out if nature at its most elementary level is symmetric say, between right and left. ...reverse the direction of time. The way you determine that is looking at say how many electrons are given off right and left at some nuclear decay. And the decay happened to have some axis of symmetry. If the numbers are not equal then you’ve broken the symmetry, you know you’ve broken a fundamental law of nature. ...the problem is of course, there could be some other kind of nuclei present also, a contaminant. Even if the contaminant is only present in 1 part per million, that is enough to make the measurement unreliable...I think what is very hard for people not in the field [of particle physics] to appreciate is how indirect the measurements are. So they are not direct measurements like you make with a ruler. There is much inference. There are many steps. So you say you are measuring something that has nothing to do with a ruler or devices. It almost always has to do with counting events. [ITC4, int]

To this day, nobody has 'seen' an atom, or an 'electron', or any other of these submicroscopic particles. Nevertheless, we have a pretty consistent picture of the behavior of atoms....Our model of atoms is based on quantum mechanics. [ITC4, vnos]

A few participants connected their views of inference with the role of models. For example, in response to the question about the evidence for the atomic model, one scientist said,

It comes from the measuring behaviors of the larger entity and making models about what is inside that would be best predict those behaviors. Those models then allow you to make further predictions that provide further tests and evidence into the individual mechanisms going on inside. [MEDF1, int]

Discipline-based and Approach-based Comparisons

Except for lacking representatives from the ESS group and abundance of representatives from the Ph group in the “affirm inference” subcode, there are no other clusters or trends based on discipline or approach.

Models

The scientists described the involvement of models in their research in a variety of ways (Tables 18 and 19), from model development to model use through testing of
Table 18. Scientists' Views of Models: Grouped by Discipline

<table>
<thead>
<tr>
<th>Aspect</th>
<th>Subcode</th>
<th>Total</th>
<th>Grouped by Discipline</th>
<th></th>
<th></th>
<th>LS</th>
<th>ESS</th>
<th>Ph</th>
<th>Ch</th>
<th>LS</th>
<th>ESS</th>
<th>Ph</th>
<th>Ch</th>
</tr>
</thead>
<tbody>
<tr>
<td>Models</td>
<td>explain or organize obs/predict/test</td>
<td>17 70.8</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Models</td>
<td>understanding of system/complexity made</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Models</td>
<td>simple/abstract made visual</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Models</td>
<td>mathematics</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Models</td>
<td>physical system</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Models</td>
<td>analogy</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Models</td>
<td>mental construct</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Models</td>
<td>representation of reality</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Models</td>
<td>more specific than theory</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Models</td>
<td>Directing framework</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Models</td>
<td>other</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Models</td>
<td>N/A</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Total</th>
<th>Number of scientists</th>
<th>Percentage of group</th>
</tr>
</thead>
<tbody>
<tr>
<td>24</td>
<td>10 5 5 4</td>
<td>70 100 40 75</td>
</tr>
<tr>
<td>LS</td>
<td>3 1 4 1</td>
<td>30 20 80 25</td>
</tr>
<tr>
<td>ESS</td>
<td>2 2 4 1</td>
<td>20 40 80 25</td>
</tr>
<tr>
<td>Ph</td>
<td>0 2 1 0</td>
<td>0 40 20 0</td>
</tr>
<tr>
<td>Ch</td>
<td>1 0 0 0</td>
<td>10 0 0 0</td>
</tr>
<tr>
<td>LS</td>
<td>0 2 0 0</td>
<td>0 20 0 0</td>
</tr>
<tr>
<td>ESS</td>
<td>0 0 0 25</td>
<td>0 0 0 25</td>
</tr>
<tr>
<td>Ph</td>
<td>10 0 20 0</td>
<td>10 0 20 0</td>
</tr>
<tr>
<td>Ch</td>
<td>20 0 20 0</td>
<td>20 0 20 0</td>
</tr>
<tr>
<td>N/A</td>
<td>20 0 20 0</td>
<td>20 0 20 0</td>
</tr>
<tr>
<td>Aspect</td>
<td>Subcode</td>
<td>Total</td>
</tr>
<tr>
<td>--------------------------------</td>
<td>------------------------------------------------------</td>
<td>-------</td>
</tr>
<tr>
<td>Models</td>
<td>explain or organize obs/predict/test</td>
<td>17</td>
</tr>
<tr>
<td></td>
<td>understanding of system/complexity made</td>
<td>9</td>
</tr>
<tr>
<td></td>
<td>simple/abstract made visual</td>
<td>9</td>
</tr>
<tr>
<td></td>
<td>mathematics</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>physical system</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>analogy</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>mental construct</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>representation of reality</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>more specific than theory</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>Directing framework</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>other</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td>N/A</td>
<td>2</td>
</tr>
</tbody>
</table>

Total Number of scientists: 24

<table>
<thead>
<tr>
<th>Grouped by Approach</th>
<th>Number of scientists</th>
<th>E</th>
<th>E/D</th>
<th>D</th>
<th>T</th>
<th>Percentage of group</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>10</td>
<td>5</td>
<td>5</td>
<td>4</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>6</td>
<td>5</td>
<td>4</td>
<td></td>
</tr>
</tbody>
</table>

E: experimental; E/D: combination of experimental and descriptive; D: descriptive (non-experimental); T: theoretical
predictions. The following results represent how the scientists describe models within their field. Again, the subcodes are not mutually exclusive. Many of the representative quotes included here were coded in multiple subcodes. There was not an overall pattern that describes the multiple coding. For example, representatives of the largest subcodes were dispersed across the other subcodes.

Scientists’ Views

*Explain or organize observations/used for prediction/testing.* Seventeen of the 24 (70.7%) scientists indicated models were explanations or ways to organize observations that also involved testing predictions. Most responses were specifically related to the participant’s research.

As models become more complex, such as general circulation models of the atmosphere and ocean, the models are used as predictive tools. They're used to predict how climate will change as we change the composition of the atmosphere. [cDFC2, vnos]

All scientists use models and if they say they do not then they are failing to understand what they are doing. In my research I use the model of a trophic cascade that indicates how predator-prey interactions from the top of the food web propagate down the food web to affect lower trophic levels. This model explains some of the variability observed in food web dynamics and the relative abundance of predator and prey groups in ecosystems. [PEDF1, vnos]

A scientific model is a description of a physical system that provides an understanding of what the system is and how it works. A scientific model allows us to organize our information about a system and to predict how the system might evolve or react... We use mathematical models of stellar atmospheres to compute what the spectrum of a star ought to look like. We compare the predicted stellar spectrum with the observed stellar spectrum to determine the composition of the star. [pEDFC2, vnos]

A model is a quantitative mathematical/physical/biological model that explains observations in a verifiable manner...The purpose of a scientific model is to understand and predict observable phenomena in the universe, ie, verifiable truths...[Example from work] the numerical weather prediction
(NWP) model. With the use of a computer, an NWP model combines a mathematical and physical model of data and observations to obtain an initial state of the atmosphere. The model is scientific because its predictions can be verified quantitatively by peers.

This atmospheric scientist continued during the interview by discussing modeling of a system:

You have to model the statistical ensemble. You have to model what is happening on the average over the whole cloud. You are probably aware that the treatment of clouds in climate models is one of the weakest links in the chain of things that we need to put together to say something sensible about global warming. And we don't do it very well. The models are all over the map, depending on how they parameterize the cloud process.

Several responses within this subcode demonstrated a connection between the scientists' views of models and their views of certainty and hierarchy of scientific knowledge.

It is a mental or physical construct. The model is a way to test whether we got our ideas right. It is the assembling of those ideas into a model and watching them and comparing to what we really see that tells us if we have all our ideas right about the real system, or if something is missing or goofed up. If they are similar, then that tells me there is a good chance the ideas that went into making the model are actually pretty good at representing what is going on in reality. Then you can test it and try a different set of conditions. If they do, then it means the model is working, at least for these conditions, and it has some predictive function. One is to test the input to see if I have my ideas straight and the other is to make predictions.

So the model is useful where it can be predictive. We have a prediction on the structure of the atom and how it should function under new circumstances. And then we can test that. So that would be through experiment, to test the predictions made by the previous model. It is useful because it can help predict what new compounds might be, what kinds of molecules might be constructed under what conditions. There can be practical uses as well. Models are useful to guide experimentation and serve as a provisional understanding of a phenomenon. A useful model in my work is that ribosomal pauses lead to errors, like frameshifting. It is a model because it is based on fundamental laws about chemical reactions. This is only a model because we cannot directly measure the reaction rates. It is something to be refined, but it is getting better all the time.
The theory of natural selection is also a model that explains much about the origin and behavior of biological systems. It provides a basis for making predictions about species responses to environmental changes. A lot of these conclusions are drawn from tests with models that show that if you create this kind of structure it accounts for the behavior that you measure. Again, just because you can come up with a model that explains it doesn't necessarily mean that is the only model. Just maybe we haven't thought of the model that works better…Models work at all these levels [hypothesis, theory, law]. A hypothesis is a model. The model becomes more robust as it becomes elevated to theory and then law. But a model initially is a hypothesis. [SED1, vnos and int]

One participant began to question her views as she thought about the model of the atom. In her VNOS response, she described a model as, “a representation of a phenomenon.” She explored her ideas further during the interview.

It makes me pause and think really about it. Because to a certain extent, I guess you could say what these are is models. And, I mean, if you haven't really seen it, it is a model. Drawing a relationship to religion, What is God? …What does God look like? Different people draw different images. They haven't ever seen him, that I know. Or her. I guess...so what they base it on and why there are images of what an atom is or drawn the way they are…I can't tell you. .....And I think, I teach this stuff! [NEFL1, int]

In response to a prompt to discuss the development of the atomic model, one participant explained the explanatory and predictive power of this model across disciplines.

The prediction was that if the plum model existed, the scattering would be very weak, that is very small angles of scattering. But what they observed were huge, very huge angles. They only way you could explain these huge angles was with the planetary model. From the origin of that experiment became our idea of what the atom had to look like…the planetary structure. That held up. From that structure as well, they were ultimately able, on the basis of that model, refinements to that model, explain and predict the phenomenon of how the atoms interact with light…. Once the planetary model became acceptable, things that could be predicted from this model were consistent with what physicists were observing then it was quickly discovered that it was also consistent with the chemists, this whole body of knowledge that chemists were building. All of a sudden the world was falling in place. Chemists could see very neatly how their atoms stuck together and
begin to explain things. Linus Pauling came along and used the model, extended the model, to explain the chemical bond and all of modern chemistry ... Of course over the years the model continues to be used and refined in ways we hadn’t even imagined. We are comfortable with that until some day we bump up against something we can’t explain with the model. At that time we go back and try to adjust the model or come up with other explanations. It’s progressive. [bEL3, int]

*Complex made simple/ abstract made visual.* Nine participants (37.5%) describe models more specifically as a means to simplify a complex process or system or a means to visualize an abstract concept. Most representatives from within this subcode were distinct from the previous in that rather than considering models as explanations of observations that serve a predictive function; models are considered limited, but useful, explanations because they serve to simplify natural phenomena that would otherwise be too complicated to investigate further.

I use models all the time. In nature, systems are invariably far more complex than the idealized simplifications we rely upon to establish our theories and laws. [cDFC2, vnos]

Three of the scientists who demonstrated this view of models included a description of model use within the context of ecology. Typical descriptions included,

A scientific model helps to explain a natural situation. Often it is a small scale general version of a more complex phenomenon. Scientific models help us to grasp a complex situation as a more watered-down version. In the field of landscape ecology, scientists often cut fields into different patch sizes and patterns and study animal movements in them to model (simulate) how larger animals move about in larger more complex landscapes. Models can be increased in scope and complexity to further explain the variability we often encounter in nature. [mDF1, vnos]

Community ecology has often been compared to nuclear physics and astrophysics. ... [more complexity, the less predictive ability.] You have to use approximation methods. The same thing happens in looking at the structure in an atom where they don’t know actually where every electron is... probability theory... Community ecologists often end up doing the same
thing. For example, we might end up lumping all the bacteria together because it's just impossible to figure out the details of what is going on with individual species. Or we may not even sometimes care about individual species. There is actually a big division in community ecology against people in ecosystem ecology. They are really just interested in energy flow through trophic levels and nutrient flow through trophic levels and don't care at all about the individual species. They don't think they are important for the level of predictions they want to make... On the other hand, community ecologists, from my approach, tend to look at individual species and in doing so we often tend to neglect understanding how multi [unclear] are restricted by energy flow or nutrient flow.... A model is a simplified view of something complex used to analyze and solve problems or to make predictions. Simplification allows the user to focus on particular factors of interest while, of course, ignoring or holding other factors constant. [MEDF1, vnos]

Four of the physicists also appeared in this subcode. Their comments included:

So we have models, theoretical models, mathematical equations you need to solve. In the past they have had to be solved approximately with very simplified models. We use the term models as distinct from theory because theory you think is correct. Model is simplified. [ITC4, int]

I think it is impossible to make an accurate representation [of the atom]. I think you have to start somewhere to explain things at a certain level of complexity... A cloud isn’t really the right idea. Repeated measurements of the position of the atom over and over again you would see something like a cloud. But that is misleading because when you measure the electron around the atom you will find it at a particular point. The cloud represents the probability that you will find it at a particular point. If you look for it you will see it. You can’t see it without changing what is there. There is an interaction between the observer and the thing that you observe. It is very very complicated. If you are studying quantum mechanics and haven’t been thoroughly confused by the theory, then you haven’t really understood it... The models are okay as long as you understand the limitations of them. That isn’t really how it is but it’s the way we think about it... We are showing pictures here that relate to certain aspects of an atom. That is what you do when you see an elephant. It depends where you are looking on the elephant and what scale. [jTC4, int]

This physicist provided an additional example from his field:

It [a model] permits us to try to isolate and explain a few aspects of a mysterious system, without having to 'get it all right' in a consistent fashion. [For example] Quark model of Strongly Interacting Particles..... This model doesn't even pretend to tell us everything we know about protons, but it allows us to study situations in which the strong force is the most important interaction. This model may be an adequate description of protons in a
nucleus, but it will not explain how we can extract energy from the sun or a thermo-nuclear reaction. [jTC4, vnos]

**Mathematics.** Nine participants referred to models as mathematical representations. Within this subcode were statements to demonstrate the role of mathematics in dealing with complexity. As the complexity of the phenomenon increases, capabilities of mathematics becomes more important.

...the practice in environmental science has been it is far too complicated, we know it [the phenomenon in nature] is 1000 or 2000 [factors involved], but why do we represent it as 1 or 2. So that has been the practice in the field for 20 years. Just not to deal with the complexity. What we are working on are mathematical methods to say, ok folks lets stop kidding ourselves. This problem is far more complicated than that. So let's just accept the high degree of complexity and deal with the best way we can. [wEL3, int]

So for particle physics there is a theory now known as quantum chromodynamics, QCD. It is a field theory. To solve that problem requires exchange of 16 different particles simultaneously. So it requires hundreds of equations to be solved simultaneously, and they are integral equations. That has taken years of computer time for most elementary, even models there, how to solve that. But in theory one has a complete mathematical description. In practice you say lets model it by limiting the number of particles. That makes it a model. [ITC4, int]

**Directing framework.** A few scientists explicitly made reference to models as a theoretical framework. This view is a specific example of connecting a type of scientific knowledge to the theory-laden nature of science (subjectivity).

Without models, observation would amount to cataloging data ... There is a lot of data and its doesn't mean anything until you have a model. If you have all these data and lots of satellites taking all these data...it doesn't tell you what to look for. It just tells you whether a model you have is plausible or not. It is all indirect. [pTC4, vnos]

A gene network is a scientific model, postulating patterns of interacting among gene products following an analogy with a computer wiring diagram. It illustrates a mechanism, and helps develop hypotheses about other genes that must be involved to produce the observed phenotype. [KEDF1, vnos]
Discipline-based Comparisons

The highest category for the total group, the “explain or organize/predict/test”, was also the highest for the discipline groups except the physicists. The physicists had a greater tendency (4 of the 5 Ph; 4 of the 9 total) to fall within the category of “complexity made simple/abstract made visual.” Interestingly, 100% of the ESS discipline group fell within the former category (explain or organize/predict/test). The physicists clustered in the “complex made simple/abstract made visual” and “mathematics” subcodes (4 of the 5 Ph). Overall, there were scientists who clearly emphasized the use of models in their research, either through model development or model testing. The two atmospheric scientists and the aquatic ecologist especially stood out from the group as ones who strongly emphasized models throughout their questionnaire and interview responses. All three of these scientists reported using models for testing predictions in their work.

Approach-based Comparisons

100% of the E/D group and 80% of the D group described models fitting with the “explain or organize/predict/test” subcode. Three of the four theoreticians (75%) responded favoring the “complex made simple/abstract made visual” description. The T also had a higher tendency to explain models as mathematical entities (75% versus 37.5% for the whole group).

Experiment

Tables 20 and 21 present the results for the scientists’ views of scientific experiments.
Scientists' Views

Despite the majority agreement toward a traditional view of experiment (70.8%), the responses were diverse and lacked of consistency related to “hypothesis-driven” (41.7%) and “requirement for science” (12.5% yes versus 54.2% no) and “requires replicas” (33.3%).

Traditional. Seventeen (70.8%) of the participants held traditional views of scientific experiment; that being that experiments involve identification and manipulation of variables and controls in efforts to establish cause/effect relationships.

I think in the traditional way. There are conditions where only one thing is different. At least two different replicas of the same thing. You have your experiment and a control. [SEL1, int]

Manipulation of a variable in a defined setting in relation to an unmanipulated control or reference. [PEDF1, vosi]

A scientific experiment is a set of manipulations and observations designed to test (hopefully rigorously) a particular hypothesis. [KEDF1, vosi]

Hypothesis-driven. Ten of the 24 indicated experiments were necessarily hypothesis-driven. However, none of the participants volunteered the view that experiments did not require hypotheses. These results suggest that those who claimed the need for hypotheses held this criterion as one of importance, while the remaining may or may not have held this as a criterion at all. It was noted that the inclusion of a hypothesis was sometimes sufficient to designate an experiment; whereas, others explained experiments as hypothesis-driven investigations that also included manipulation and controls.
<table>
<thead>
<tr>
<th>Aspect</th>
<th>Subcode</th>
<th>Total</th>
<th>Percentage of group</th>
<th>Grouped by Discipline</th>
<th>Percentage of group</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Experiment</td>
<td>traditional: controls/variables/manipulation</td>
<td>17</td>
<td>70.8</td>
<td>8 3 4 2</td>
<td>80 60 80 50</td>
</tr>
<tr>
<td></td>
<td>hypothesis driven</td>
<td>10</td>
<td>41.7</td>
<td>5 0 1 4</td>
<td>50 0 20 100</td>
</tr>
<tr>
<td></td>
<td>hypothesis not required</td>
<td>0</td>
<td>0.0</td>
<td>0 0 0 0</td>
<td>0 0 0 0</td>
</tr>
<tr>
<td></td>
<td>test of models</td>
<td>2</td>
<td>8.3</td>
<td>0 1 1 0</td>
<td>0 20 20 0</td>
</tr>
<tr>
<td></td>
<td>inclusive: any test of idea, with or without manipulation</td>
<td>2</td>
<td>8.3</td>
<td>1 0 1 0</td>
<td>0 20 20 0</td>
</tr>
<tr>
<td></td>
<td>variable by context: lab versus field</td>
<td>6</td>
<td>25.0</td>
<td>3 3 0 0</td>
<td>30 60 0 0</td>
</tr>
<tr>
<td></td>
<td>requires replicas</td>
<td>8</td>
<td>33.3</td>
<td>4 1 1 2</td>
<td>40 20 20 50</td>
</tr>
<tr>
<td></td>
<td>required for science</td>
<td>3</td>
<td>12.5</td>
<td>1 0 0 2</td>
<td>10 0 0 50</td>
</tr>
<tr>
<td></td>
<td>not required for science</td>
<td>13</td>
<td>54.2</td>
<td>5 2 4 2</td>
<td>50 40 80 50</td>
</tr>
</tbody>
</table>
Table 21. Scientists' Views of Experiment: Grouped by Approach

<table>
<thead>
<tr>
<th>Aspect</th>
<th>Subcode</th>
<th>Total #</th>
<th>Total %</th>
<th>E</th>
<th>E/D</th>
<th>D</th>
<th>T</th>
<th>E %</th>
<th>E/D %</th>
<th>D %</th>
<th>T %</th>
</tr>
</thead>
<tbody>
<tr>
<td>Experiment</td>
<td>traditional: controls/variables/manipulation</td>
<td>17</td>
<td>70.8</td>
<td>7</td>
<td>3</td>
<td>4</td>
<td>3</td>
<td>70</td>
<td>60</td>
<td>80</td>
<td>75</td>
</tr>
<tr>
<td></td>
<td>hypothesis driven</td>
<td>10</td>
<td>41.7</td>
<td>6</td>
<td>3</td>
<td>0</td>
<td>1</td>
<td>60</td>
<td>0</td>
<td>20</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>hypothesis not required</td>
<td>0</td>
<td>0.0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>test of models</td>
<td>2</td>
<td>8.3</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>60</td>
<td>40</td>
</tr>
<tr>
<td></td>
<td>inclusive: any test of idea, with or without</td>
<td>2</td>
<td>8.3</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>20</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>manipulation</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>variable by context: lab versus field</td>
<td>6</td>
<td>25.0</td>
<td>1</td>
<td>3</td>
<td>2</td>
<td>0</td>
<td>10</td>
<td>60</td>
<td>40</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>requires replicas</td>
<td>8</td>
<td>33.3</td>
<td>6</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>60</td>
<td>20</td>
<td>20</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>required for science</td>
<td>3</td>
<td>12.5</td>
<td>2</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>20</td>
<td>20</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>not required for science</td>
<td>13</td>
<td>54.2</td>
<td>4</td>
<td>4</td>
<td>2</td>
<td>3</td>
<td>40</td>
<td>80</td>
<td>40</td>
<td>75</td>
</tr>
</tbody>
</table>

E: experimental; E/D: combination of experimental and descriptive; D: descriptive (non-experimental); T: theoretical
[A scientific experiment is] a test of an idea about how nature works. An observation or protocol is designed such that if the idea is correct, one expects one observation; and if the idea is wrong, the one expects to see something else. [UEL1, vosi]

To me it [a scientific experiment] is a hypothesis and a test. A strict definition of a test is something that will establish the probability of a hypothesis being correct. This usually means conducting experiments and conducting them with sufficient number of repetitions that you gain confidence that the outcomes are reproducible. A test can be a theoretical test, that is you can arrive at it by analysis through mathematics. ...I tend to think of it more often in experimental terms. That is conducting experiments on matter, instruments. [bEL3, int]

VOSI item #5 elicited ideas about experiments and hypothesis testing. Most respondents did not consider the scenario to be experimental. They reported the investigation was not experimental due to the lack of controls and manipulation of a variable.

[The investigation is] not experimental yet. Not at this point. However, the hypothesis that tooth structure and food source could be tested by experiment to see if each group could process different food or if new animal species, selected to represent meat eaters and plant eaters, continued to show the tooth difference. [SEDF1, vosi]

The few who reported the investigation to be an experiment either had an inclusive view of what constitutes experiment or determined there must be a hypothesis, and therefore it is an experiment.

If systematic, then it is good science. Systematic means the animals are chosen selectively, and he studied a variety of them....Yes experiment: careful and unbiased observations of a natural phenomenon. [kEL4, vosi]

With a retro-fit hypothesis to the investigation in VOSI #5, several scientists thought the situation was an experiment. In these cases, the scientists equated hypothesis-testing with experiment, regardless of manipulation or control.
Hypotheses may be implicit or explicit. For the teeth example, yes it is an experiment because there was an implicit hypothesis. [gEL3, vosi]

**Experiments are not required for the development of scientific knowledge.**

Thirteen (54.2%) participants indicated experiments are not required for the development of scientific knowledge. In reference to the VOSI item #5, scientists’ responses included,

Do I consider it to be an experiment? Not so much. It is an observation. He has not actually done a manipulation to see, generate a hypothesis and manipulate the system to see if this does vary. You can still publish this. You can still address the information. Uh...but it is not actually testing a mechanism. It happens all the time in medicine. [OEL3, int]

All experimental units are identical to the control with the exception of the thing being tested...At one time we came to the conclusion that salmon live in saltwater and spawn in fresh water through observation...If I felt that experiment was absolutely necessary, I couldn't be a field biologist or a wildlife habitat biologist. [mDF1, vosi]

[A scientific experiment is an] investigation of a physical system in which one or a few variables are changed while most fixed; cause/effect. [Experiments are] not required. In astronomy we make many observations of celestial objects to learn about them, but cannot conduct experiments on them. [pEDFC2, vosi]

A slightly different tone was demonstrated by one of the theoretical physicists. He stated a general definition of an experiment, “Test of the consequences of a scientific theory or model.” He added that experiments were not necessarily required because new knowledge could be developed without new data.

Some scientific progress can be made with data at hand. New models and scientific schemes can unify those data. The very mathematical consistency of the model or scientific scheme can be a strong argument for it. String theory is in such a state now; no one can even formulate a mathematically consistent theory. [pTC4, vosi]
Experiments are required for the development of scientific knowledge. Three (12.5%) participants said experiments are required. These three were also among the five who stated experiments require hypotheses, and as is discussed below, they also indicated all methods (or method) of science are hypothesis-driven.

R: [Referencing item #5 VOSI] Given this situation, he notices this relationship, do you think this is a reasonable conclusion?
S: Perhaps. It is based on a limited number of observations. There is no experiment here. It is a series of observations. It is not set up to test a hypothesis. These are what I would call casual observations. So is this investigation scientific? No. ...perhaps there is some scientific tendency here. ...Certainly there are people who are making correlations and testing them, probably by looking to see what else falls within these models. [fEF3, int]

The two others in this subcode did not voice the traditional view of scientific experiments. These results suggest some of the scientists hold different ideas about what constitutes scientific experiment (about 30% in this sample) than the traditional view of controls and variables to establish cause/effect relations.

Experiments require replicas. Eight participants (33.3%) indicated valid experiments needed to include replication.

[Experiment involves] manipulating, probing, testing the system. Not just getting a single number or analysis, but rather a series of analyses that are framed to answer a certain question. [fEF3, int]

There are conditions where only one thing is different. At least two different replicas of the same thing. You have your experiment and a control." [SEL1, int]

[A scientific experiment is] A controlled and repeatable procedure designed to answer a scientific question with a quantifiable degree of uncertainty. An example would be a laboratory experiment with two "cloud droplets" of different radii falling in a vertical wind tunnel to determine the efficiency with which they collide and coalesce. Collision/coalescence is an important precipitation mechanism in clouds that are everywhere warmer than the freezing point of water. Unlike observations of collision/coalescence in a real cloud, the laboratory procedure can be repeated. [eDF2, vosi]
In response to the question about the requirement of experiments for the development of scientific knowledge, this scientist responded,

No. If that were true, we would know very little about the atmosphere. You can observe nature and model the results without studying the phenomena in a laboratory. [eDF2, vosi]

*Experiments are variable by context.* Six participants (25%) suggested experiments differ with the context of investigation. For example, experiments conducted in the laboratory are different in form than experiments conducted in a field setting. The type of system under investigation was also mentioned as a determining factor in the type of experiment that is possible.

Experiments are simply part of the package. Experiments need context; they need backing concepts and theory. Not all scientific phenomena are readily addressed by the type of conventional experiments I described above. I think that ecology especially is evolving toward a view of comparisons of multiple models with data to ascertain which model fit the data best. This represents a change in perspective from traditional experiments and traditional statistical analysis of experiments. [PEDF1, vosi]

Experiments in physics are different from experiments in astronomy because we can’t go and scoop up the stars and make changes. We can’t investigate particular variables in isolation. Our stars are what they are and we have to deal with nature as it appears through our telescopes. And so that means that I need to have a well developed sense of experimental technique because I have to be able to select samples of stars for observation that will allow me to reduce the number of variables. For example if I want to study age affect, I don’t want to be selecting stars that have dramatically different rotation speeds or dramatically different compositions. I need to be able to select stars with similar properties except for the age. [pEDFC2, int]

S: Chemistry is largely laboratory driven discipline. Experimentation happens in the lab. Then the applied...also can take place in the lab is a different type of problem solving, not solely getting at the basic principles. So there is a bit of a divide starting. ...See there is a lot less control on the environment and I think your classical chemist who has everything under control, time, temp, concentrations, etc...You don't have control outside in the environment. In terms of approach, there are chemists out there, the classical traditional, who
have every parameter under control. And then there is environmental chemistry who has less control. People in the lab don't understand how the people in the field can function because we have even less control...

R: How about outside of chemistry? Say astronomy?
S: Is it representative? That is the approach. The general idea, I'll have an idea and I'll set up experiments to test the idea. That I think underlies everybody. It is how you do your daily business is where things differ. [fEF3, int]

The following quote from an atmospheric scientist represents a view of experiments as inclusive, that is, not necessarily involving direct manipulation. The statement also mentions a requirement for reproducibility and prediction. The statement indicates a difference between “observational experiments” and traditional laboratory experiments.

We don’t have a laboratory. We can’t get in a laboratory. The best thing we can do is an observational experiment. We try to design a strategy, the example I gave of looking straight down...if we do it enough times we should get the same result if everything is behaving alright. Is it consistent or inconsistent? The same thing is with the ship tracks. We design it in such a way that we can tell if the clouds were behaving, we would be able to tell you that. But the fact that they weren’t we could tell you that. We have all the data. What we do with the data is the difference. Doing it in a way that is difficult to argue with and also is repeatable. So if someone tried to do the same thing we did they would come up with the same answer. The answer is definite though, that going after the modeling community I can say, "Your model is not right." An experiment, an observational experiment, we try to analyze the data in such a way to come up with answers that are definitive. Ok? That is the essence of an experiment. Designing a strategy, whether it is building a piece of equipment or something, that tells you something new about the system. Something not known before in a way that uh..gains us a little on where we stand as far as our predictive capabilities or understanding. [cDFC2, int]

Discipline-based Comparisons

There were no apparent differences among scientists’ groups within the “traditional” subcode. One of the interesting discipline-based features is relative to views of experimentation being required for science or not. Four of the five
physicists indicated experiments are not required. Three of these physicists were theoretical-based. The chemists were split (2 each) within the “required for science” and “not required for science” groups. However, all of the chemists considered experiments to be hypothesis-driven investigations, in contrast to none of the ESS. The ESS group did not indicate one way or the other about ties of hypotheses to experimentation. This absence may suggest a limited conceptual link between hypothesis and experimentation in comparison to the scientists in the other groups. The ESS did tend to cluster within the “variable by context” subcode (3 of the 5 ESS; 3 of the 6 total), with none of the physicists or chemists represented. Overall, these results suggest the chemists had a greater tendency to view experiments as hypothesis-driven investigations; the physicists suggested experiments are not required for the development of scientific knowledge; and the Earth and space scientists did not give as much prominence to hypotheses within experiments as the other groups; nor did they consider experiments a requirement within science; but they did describe a difference in experimental procedures depending on the context.

Approach-based Comparisons

Comparisons based on investigative approach reveal more clusters within this main category (Table 21). Nine of the 10 scientists who reported experiments as “hypothesis-driven” were among the E or E/D groups; whereas none of the scientists in the D group explicitly connected hypotheses with experiments. In addition, 4 of the 5 scientists grouped within the E/D approach suggested experiments are not required, as did 3 of the 4 theoreticians. The three scientists who indicated “required for science” were within the E and E/D groups. Sixty percent of the E/D group
(representing 3 of the 6 in the total sample) also suggested experiments in the
laboratory were different than experiments that were field-based. These results are
not surprising given that the E/D participants conduct both experiments and
descriptive type research. When combined with the D approach group, 5 of the 6
total who indicated a difference in experimentation based on the context of discipline
were within the E/D or D groups. Another interesting feature is within the
experimental (E) group’s greater indication that experiments require replications
(60% of the E group, 6 of the 8 within this subcode).

Purpose

Tables 22 and 23 present results of scientists’ views of the purpose of
scientific inquiry. Purposes of science fell within both “basic” and “applied”
categories.

Scientists’ Views

More participants indicated a “basic” purpose (58.3%) than an “applied”
purpose (16.7%) Several indicated both ideas.

[The purpose of scientific is] To develop new knowledge, to deepen
understanding about in my case aquatic ecosystems, to provide information
useful for managing the environment (although this latter point is not a
purpose of scientific inquiry per se). [PEDF1, vosi]

The primary goal of scientific inquiry is to generate as full an understanding
as possible of the physical world around us...A secondary goal is to use that
understanding in ways that improve the life conditions of our companions.
[gEL3, vosi]

[The purpose of scientific inquiry is] To learn about the world. There may
also be useful applications. My study of the CGA codon was designed to
learn about the translational mechanisms, which are fundamental to life.
There may be practical implications for those who express foreign genes in E. coli. Those individuals may wish to change CGA codons to something else. [UEL1, vosi]

**Discipline-based and Approach-based Comparisons**

There was little apparent difference among discipline and approach-based groups with respect to views of the purpose of scientific inquiry. The LS showed a tendency over the Ch within the “basic: understanding” subcode (80% versus 25%). There was no apparent difference in these groups within the “applied” subcode.

**Methods**

The multiple subcodes for this aspect suggest the scientists voiced a diversity of views related to methods of scientific investigations. These results are displayed in Tables 24 and 25.

**Scientists’ Views**

*Single scientific method.* One participant commented during the interview, “If you say “the scientific method” to any scientist, you would get laughed at! It is absurd.” His prediction did not hold for this sample of scientists. Six participants (25%) said there is a single Scientific Method that all scientists use or should use. How they defined that method ranged from traditional experiments to anything that formally tested a hypothesis.
### Table 22. Scientists' Views of Purpose: Grouped by Discipline

<table>
<thead>
<tr>
<th>Aspect</th>
<th>Subcode</th>
<th>Total</th>
<th>Grouped by Discipline</th>
<th>Percentage of group</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>#</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Purpose</td>
<td>Basic</td>
<td>3</td>
<td>12.5</td>
<td></td>
</tr>
<tr>
<td></td>
<td>understanding</td>
<td>14</td>
<td>58.3</td>
<td></td>
</tr>
<tr>
<td></td>
<td>curiosity</td>
<td>2</td>
<td>8.3</td>
<td></td>
</tr>
<tr>
<td>Applied</td>
<td>improve quality of life</td>
<td>6</td>
<td>25.0</td>
<td></td>
</tr>
<tr>
<td></td>
<td>answering questions</td>
<td>4</td>
<td>16.7</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Predict</td>
<td>1</td>
<td>4.2</td>
<td></td>
</tr>
<tr>
<td>Discovery/Serendipity</td>
<td></td>
<td>3</td>
<td>12.5</td>
<td></td>
</tr>
<tr>
<td></td>
<td>differs among disciplines</td>
<td>3</td>
<td>12.5</td>
<td></td>
</tr>
<tr>
<td></td>
<td>no response</td>
<td>3</td>
<td>12.5</td>
<td></td>
</tr>
</tbody>
</table>
Table 23. Scientists' Views of Purpose: Grouped by Approach

| Aspect                        | Subcode              | Total | Total | E | E/D | D | T | Percentage of group |
|------------------------------|----------------------|-------|-------|---------------|-----|-----|---------------------|
| Purpose                      | Basic                | 3     | 12.5  | 2  | 1   | 0  | 0  | 20                  |
|                              | understanding        | 14    | 58.3  | 6  | 4   | 2  | 2  | 60                  |
|                              | curiosity            | 2     | 8.3   | 1  | 1   | 1  | 0  | 10                  |
| Applied                      | improve quality of life | 6     | 25.0  | 2  | 2   | 1  | 1  | 20                  |
|                              | answering questions  | 4     | 16.7  | 1  | 1   | 1  | 1  | 10                  |
|                              | Predict              | 1     | 4.2   | 1  | 0   | 0  | 0  | 10                  |
|                              | Discovery/Serendipity| 3     | 12.5  | 1  | 1   | 2  | 0  | 10                  |
|                              | differs among disciplines | 1     | 4.2   | 1  | 2   | 1  | 0  | 10                  |
|                              | no response          | 3     | 12.5  | 1  | 0   | 1  | 1  | 10                  |

E: experimental; E/D: combination of experimental and descriptive; D: descriptive (non-experimental); T: theoretical
<table>
<thead>
<tr>
<th>Aspect</th>
<th>Subcode</th>
<th>Total</th>
<th>Grouped by Discipline</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Methods</td>
<td></td>
<td>Total</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Single Scientific Method</td>
<td>6</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Multiple methods ...</td>
<td>16</td>
<td></td>
</tr>
<tr>
<td></td>
<td>all hypothesis driven</td>
<td>7</td>
<td></td>
</tr>
<tr>
<td></td>
<td>not all hypothesis driven</td>
<td>8</td>
<td></td>
</tr>
<tr>
<td></td>
<td>reductionist</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td></td>
<td>expt, descriptive</td>
<td>11</td>
<td></td>
</tr>
<tr>
<td></td>
<td>expt, observational, theoretical</td>
<td>7</td>
<td></td>
</tr>
<tr>
<td></td>
<td>methods influenced by complexity of systems</td>
<td>5</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Hierarchy:</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Validity/confidence of claims influenced by methods:</td>
<td>9</td>
<td></td>
</tr>
<tr>
<td></td>
<td>expt &gt; descriptive</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Maturity of the science: Expt more mature than desc</td>
<td>4</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Theoretical sciences: pushing the limits of</td>
<td>4</td>
<td></td>
</tr>
<tr>
<td></td>
<td>science/mathematics</td>
<td>9</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Discipline differences: complexity of systems, ability to repeat</td>
<td>4</td>
<td></td>
</tr>
</tbody>
</table>
Table 25. Scientists' Views of Methods: Grouped by Approach

<table>
<thead>
<tr>
<th>Aspect</th>
<th>Subcode</th>
<th>Methods</th>
<th>Total</th>
<th>Grouped by Approach</th>
<th>Number of scientists</th>
<th>Percentage of group</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>Total</td>
<td>E</td>
<td>E/D</td>
<td>D</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Methods</td>
<td>Single Scientific Method</td>
<td>Single Scientific Method</td>
<td>6</td>
<td>25.0</td>
<td>3</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>Multiple methods ...</td>
<td>Multiple methods ...</td>
<td>16</td>
<td>66.7</td>
<td>6</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>all hypothesis driven</td>
<td>all hypothesis driven</td>
<td>7</td>
<td>29.2</td>
<td>5</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>not all hypothesis driven</td>
<td>not all hypothesis driven</td>
<td>8</td>
<td>33.3</td>
<td>4</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>reductionist</td>
<td>reductionist</td>
<td>2</td>
<td>8.3</td>
<td>2</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>expt, descriptive</td>
<td>expt, descriptive</td>
<td>11</td>
<td>45.8</td>
<td>7</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>expt, observational, theoretical</td>
<td>expt, observational, theoretical</td>
<td>7</td>
<td>29.2</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>methods influenced by complexity of systems</td>
<td>methods influenced by complexity of systems</td>
<td>5</td>
<td>20.8</td>
<td>3</td>
<td>2</td>
</tr>
<tr>
<td>Hierarchy:</td>
<td>Validity/confidence of claims influenced by methods:</td>
<td>Validity/confidence of claims influenced by methods:</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>expt &gt; descriptive</td>
<td>expt &gt; descriptive</td>
<td>9</td>
<td>37.5</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>Maturity of the science: Expt more mature than desc</td>
<td>Maturity of the science: Expt more mature than desc</td>
<td>3</td>
<td>12.5</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>Theoretical sciences: pushing the limits of science/mathematics</td>
<td>Theoretical sciences: pushing the limits of science/mathematics</td>
<td>4</td>
<td>16.7</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>Discipline differences: complexity of systems, ability to repeat</td>
<td>Discipline differences: complexity of systems, ability to repeat</td>
<td>9</td>
<td>37.5</td>
<td>4</td>
<td>2</td>
</tr>
</tbody>
</table>

E: experimental; E/D: combination of experimental and descriptive; D: descriptive (non-experimental); T: theoretical
Well, I think the scientific method is propose a question, design some experiments, get some observations, test the hypothesis. That is still very valid. But creativity can be part of that process of thinking of the questions, designing the experiment, and then looking at the data. So I think the scientific method is very much in tact and it is not overly grandiose. It is actually quite simple. [fEF3, int]

Science inquiry is guided by the basic scientific method of observation, question, hypothesis, experiment, evaluation. I can observe patterns in how the abundances of two species are correlated. I can ask how such patterns could occur and devise a testable hypothesis, then I can conduct and evaluate the test. I believe that my approach is representative of all scientific research. [MEDF1,vosi]

The following response demonstrates a connection between view of scientific method, requirement for empirical basis, and accuracy of the knowledge produced.

Science is distinguished by the scientific method as the general approach to determining the answer to a question/hypothesis...As a scientist, I would say that we use the scientific method to explore and understand the natural world...from the data we collect we attempt to develop a picture of that world...a picture that is as accurate as possible. Art, on the other hand, is sometimes accurate, but need not be so. [NEFL1,vnos]

Four of the scientists who considered there to be a “scientific method” defined that method as hypothesis testing. Such a method could be experimental or descriptive.

S: [scientific method ] involves having a hypothesis and then choosing a system in which you can make observations that will address the hypothesis.
R: Does every investigation involve a hypothesis?
S: yes. I think every investigation should be driven by a hypothesis.
R: what makes a test scientific?
S: A test is scientific if it first of all addresses a hypothesis and if the assumptions that are made are well controlled, where possible, or consistent with other experimental data. [UEL1, int]

I have several different scientific methods. I tend to use whichever one is most convenient. In general though I think the scientific method involves observing something which ...and then asking questions about it. It is not just observation. Religion may be content with observation, and science involves questioning. Why it is, what made it that way? Those are questions, primary motivating questions. Starting with those questions. The scientific method involves a statement of one or more hypotheses that are essentially answers
to the questions they post. To me the scientific method is more that just saying I am interested in a river, I am interested in velocity and patterns in the river. I am going to go out and make measurements of the velocity of the river and then I am going to try to extract from that principles. To me the scientific method, is not always followed, but to me there is a step in between. You have to make some observations about velocity of rivers and go "ah hah, I think I know what makes this pattern happen. I think the flow is responding to different energy states in water..." And then you set up a test of that...You are collecting data that specifically helps you test whether the idea you had was correct. The same thing would apply in a geology setting. You walk up to a rock and you are trying...the rock represents a window in time...the sediment has different layers and so I have initial thoughts that this layer represents this or that...I have to make a set of observations or measurements that are consistent enough with what other people agree with that I can convince others. That is all part of the scientific method....[GDF2, int]

Despite his view that there is a single scientific method, one scientist reported problems and changes regarding scientists conforming to the scientific method:

Scientists follow a much more strict format [than artists]. The scientific method of observation, question, hypothesis, test, evaluation, new questions. It is also some recognition in some places that perhaps we have conformed too much. It is like funding at the National Science Foundation. A bunch of us write grants and we send them in. Their funding rate is only bout 17%. It is very hard to get a research grant. NSF reviewers do not typically score high on unique methods. They seek conformity...[The SM] is changing in odd ways, certainly more communication between more and more scientists through electronic access and things like that is probably making it...uh...making people conform to a single scientific method more so than in the past. [MEDFI, int]

**Multiple scientific methods.** Sixteen participants (66.7%) said scientific investigations could follow multiple methods. The methods included experimental and descriptive approaches (11 participants), and seven participants included "theoretical" methods in their responses.

This is not easy to answer because forms of inquiry are diverse. They can involve everything from mathematical theory to lab bench experiments to field studies. The norm in my field might be a mixture of observation and experiment that is derived from conceptual or theoretical models...My
approach follows my interest in aquatic ecosystems. I apply models that
develop predictions or at least alternatives and then conduct field experiments
and/or comparisons among ecosystems to test these models. I am also very
interested in testing ideas on whole ecosystems (e.g. an entire lake) and not
just small units within an ecosystem (e.g. a patch of enclosed water). This
reflects a bias that interesting and important answers lie at this scale of
experimentation... [Do you think your approach to scientific inquiry is
representative of all scientific research, some fields of research, or specific
only to your field? Explain with examples.] Yes and no. The differences lie
in the scale and complexity of the study object (ecosystems). The similarities
are that the way of asking questions, conducting experiments, thinking about
results is generally quite similar to other areas of science [PEDF1, vosi]

In regards to wildlife habitat research, you are opening a can of worms with
this question. Some have a very strict concept of research that requires that it
be performed following the rigid scientific method (observation, question,
hypothesis, experiment, conclusion). These folks feel that unless an
experiment is performed (all things held constant except for the thing being
tested), science has not occurred. Regrettably, this rigid conformity to the
scientific method is almost always lacking in wildlife field research. If I
hypothesize that small mammal populations will decrease when we thin the
forest a certain way, I am confronted with numerous problems if I want to
conduct an “experiment.” First of all, it is impossible to find multiple
identical forests in order to replicate (a must) the experiment. Even if I find
forests that are close in composition, etc., it is unlikely that I’ll be able to get
the landowners to thin the forests to my specifications at the same time, etc.
Consequently, most wildlife habitat work is observational not
experimental... In observational research, one finds forests (for instance)
already in the condition of interest... and looks at the wildlife habitat
relationships that exist. The rigid experimental folks say this is not research,
but often this is the best we can do. Certainly more inferences can be drawn
from experimental work, and we habitat folks jump at the chance to do
experimental field work when we can (usually with the cooperation of land
management agencies), but most of our work is observational by necessity.
[mDF1, vosi]

The following excerpts describe theoretical methods as valid science. Those
who expressed these views tended to have direct experiences with other methods,
either through colleagues or within their own research.

I am an experimentalist, so when you ask me what an experiment is I gave
that answer. There are colleagues within my department here who are
theorists. They do physics, but they don’t do experiments. They do
investigations by theory. They may use the theory to analyze experimental
data, or not. There may be no experimental data. String theory, for example, is on the fringes of particle physics... We can't visualize what it means. They are looking at an area where there is not experimental data and there isn't likely to be any. So their success is in a way based on the beauty of the symmetry or the aesthetics of a theory. How it holds together, how it looks. But yet that is valid research. [kEL4, int]

Most of my work is testing models. Model development is a whole other field. That might be the theoretical side. So I put myself in the observational side as opposed to the theoretical side. The models themselves are so complex. [cDFC2, int]

Developments in technology have advanced what science is possible, how science is done, and what is considered science.

Simulation you do two things. One you simulate nature. It is typically like rather than designing an experiment and then testing it out, you are just testing it on the computer first to make sure it is working. With simulated data coming in and simulated analysis. Then a simulation can be, you can actually solve the equations on the computer in some numerical way. So that has lead to the ability to solve equations or mathematical systems which weren't solvable otherwise. That in itself has lead to a whole new branch of science, known as computational science, which is not the same as computer science. But it is where the emphasis is on applications, solving problems, and you use the computer to do it. What is new to you is you have a whole new set of problems, ones which were either too complicated or too hard or which had no analytic solution, which you can now solve. That has very much changed science, what science has been able to look at and solve. [ITC4, int]

The lack of a single scientific method was connected to views of subjectivity in the following statement by the nuclear physicist:

As scientists we all have these kind of paradigms in mind of the way things are supposed to work. The scientific method I think implies a completely impartial investigation. We know that's definitely not true. In a lot of ways we test out their pet theories. "I have a model that I like best and I'm going to do experiments and interpret them within that model. I'm going to look for evidence that reinforces that model." When I find that, I stop. Then again, if I carry the analysis a bit deeper, I might find something that contradicted the model. [kEL4, int]
The forest ecologist had the following to say during the interview in response to a follow up question about “the scientific method.” Through her experiences, she had connected problems she sees with narrow perspectives of new science graduate students with the teaching of “the scientific method.”

R: what are your thoughts on this idea of a single scientific method?
S: Oh, I am revolted by it. I am absolutely revolted by it. Let me tell you a little story along those lines. I am the chair within our department on the curriculum development for our graduate program. We have developed an introductory course for our graduate students called “posing questions” or something like that. The whole idea is to take our incoming graduate students and introduce them to the art, and it is an art that you develop over a long period of time, in posing interesting, useful, wonderful testable scientific questions. This last year was the first year we ever taught the course. I didn’t teach it. Somebody else taught it. The students just hated it. They came in and they said, “We know the scientific method. We know this. You are treating us like babies. We don’t need to learn this any more. We want to apply it.” And I’m just thinking, “Oh my god, you have no idea. This is not something you learn in high school and from then on it is rote.” It is absolutely the opposite of that. What we are trying to teach you is as sophisticated as learning to paint like Picasso or something. You learn that by practicing, by exposure to lots of people who do it. You develop your own style. I am revolted by the concept that there is this single rigid approach. It should be beautiful. It should be artistic. It should be flexible, depending on your question. [BEFL1, int]

**Role of hypotheses:** All hypothesis-driven. The frequencies within the subcodes of “all hypothesis-driven” and “not all hypothesis driven” were similar for the total sample (29.2% and 33.3%, respectively). The following representative quotes are from those who indicated all methods are hypothesis-driven. Eliciting descriptions of scientific inquiry yielded responses such as:

I take the broad view that anything which poses a reasonable hypothesis and tests that hypothesis is a valid scientific inquiry. [gEL3, vosi]
The VOSI #5 question also elicited scientists' views of criteria for an investigation being "scientific." The requirement for hypothesis-testing to be scientific was evident in responses such as:

- He is justified because he poses a reasonable hypothesis and can further test; yes scientific because "It is based on observations from nature, and it leads to a testable hypothesis. [UEL3, vosi]

In this case, the participant retro-fit a hypothesis that was not included in the scenario. He equated "scientific" with "hypothesis test."

**Role of hypotheses: Not all hypothesis-driven.** In contrast, evidence of holding a view of inquiry that did not require hypothesis testing was seen in the following responses that provide examples from the scientists' research or experience. Eight participants (33.3%) voiced this type of view.

This [my own research] is all more or less hypothesis driven research.... There is other research, for example the initial stages of the selenium research, where you don't really have a hypothesis to begin with. You just have a problem and you may start with the data gathering activity and just gather data.... but it is not the same sort of controlled experiment. You are just gathering data and looking at it. My wife is an oceanographer, and she does a lot of what I call data gathering activities. ...You aren't going to do too many controlled experiments in the ocean itself... You're able to test precisely in the laboratory. [wEL3, int]

There is in the wildlife literature quite a few articles, most recently, saying we can all make hypothesis. But are we making up these hypotheses just to satisfy someone’s need to say we have hypotheses. These articles are now saying this is really stupid. We are making these hypotheses just to do that. The reality is that we don’t have a clue... It can be hypothesis driven. But a lot of times it is not, at least in the habitat research. ...If you have once species, you could okay, my hypothesis is that ‘spotted owls are going to decline if we get rid of the tree.’ But then you say, ‘what a dopey hypothesis. Doesn’t everyone assume that anyway?’ So what you really are doing is saying, ‘okay I know they are going to disappear. Is there a threshold they can tolerate? Do they just kind of peter out as you go through time? Do they like a scatter or particular pattern better than that?’ A lot of those things you could come up with a hypothesis either way. Once you got the results you ask
why do you think that was. Particularly if you see this kind of pattern and somebody else sees that kind of pattern, then maybe you can say ‘that is kind of a consistent response. So why would they respond this way?’ Given that I could hypothesize that they would do just the opposite...So yes, there are those who feel you have to have a hypothesis, and very often you look through the literature and there are those journals that if you don’t put up front, “here are my hypotheses,” they won’t publish it. Which is why people have come up with sort of hokey hypotheses to satisfy someone’s need to have one. When the reality is it is something that you created after the fact and you had no idea when you started. I’ll be very honest...I think some of the direction that we go and how we present things and how we would like people to think that we have thought about things, is how you get it in print.

One participant compared two types of investigations from her research area to support her claim that not all scientific investigations mandate hypothesis testing. She provided an example of a gene being expressed in the top flower of a plant, whereas in another plant the gene is on in all the flowers. She said in this case she would conduct hypothesis-driven research based on the pattern of the top flower having the gene expressed. She would seek this pattern in other plants. In contrast, she indicated she might just want to check something out to see if anything interesting is going on.

In that case, there would be no hypothesis. Also developmental morphology work is not hypothesis driven at all. ‘How does this thing get from a vegetative meristem up to producing a big open branching structure with flowers all over it? That is just look at it and see. ...In some ways the look at it and see stuff is the most fun because that is where you mostly discover the unexpected.’

Again, responses to the VOSI item #5 were very informative with respect to this subcode. Responses demonstrated some of the scientists were explicitly aware of different situations that would require a hypothesis and others that would not. A chemist said the following in reference to the structure/function scenario (VOSI #5).
In many cases in our work we don't have anything [hypotheses] either. We just think, "I wonder what would happen if we change this? I wonder what would happen if we make this hotter or colder or something like that." No hypothesis at all, just "I wonder what would happen if we did this?" ...Nothing guiding. Just strictly curiosity. [bEL3, int]

Hierarchy of methods. Within the subcode of “hierarchy,” 37.5% (9 participants) of the participants indicated that claims made through experimental methods were more valid than claims made through non-experimental methods. For example, in discussing the methods typically employed by astronomers, one participant said,

They still do hypothesis tests, but not experiments in the sense their observations are not repeatable. Like an experiment, but really only an observation of nature. In a different category because it is not repeatable....It is still scientific but it is uh...how shall I say...not likely to lead to the same level of certainty that you get with a repeatable experiment. [sTC4, int]

The following quote depicts one participant’s ideas of the problems and limitations with correlational studies as compared to experiments.

I could collect data in the field until I am blue in the face, but all that would be is a correlation. But again, I take that and I bring that into the lab and try to raise these things in the lab, manipulating photoperiod and see if I can change the type of eggs produced and relate it to the field and the sequence I saw... [in response to the dino queston] You can't really do an experiment. You are somewhat limited. [NEFL1, int]

The following excerpt from an interview with an atmospheric scientist further illustrates a view of validity differences. This excerpt demonstrates connections among several aspects that relate to this participant’s view of experimental methods being more valid than descriptive, such as prediction, socio/cultural issues in terms of funding, and justification.

R: Do you consider some types of science to be more valid than other types?
S: Yeah, I think so. I tell you what, in climate change we have a run into a difficult problem. Early on back in the 70s, it was perceived as a physics problem. Fill the atmosphere with these gases and we have these consequences. Which is a physical problem. No question about that. The research funding, and there's where it hurts, initially that is where the funding went. We have gone for many years now with essentially constant level funding. But now that funding is going into ecology this...ecology that. What happens to this worm in that environment? We have to save the ecosystem of this and the ecosystem of that. The trouble with this is I see these ecosystem studies as ill-posed and open-ended. Really open-ended. Largely simply descriptive with little real predictive capability. And because there are so many different parameters that affect the results, scientists are left to observations so they can't really constrain their study very well. So we didn't really learn anything. So what we are left with is a bunch of descriptive work. When...how the devil are we going to be able to say anything if we don't get the physics in the problem right? So you see the budgets diffuse out to a large number of researchers in various fields and part of that is simply political also. They outnumber use 10 to 1. Congressman A or Congressman B looks at his district and says we need more ecosystem research because he's got 10 people doing that type of work for every physical scientist doing that type of work. It is very clear why the budgets have gone that way. So the result is we still are left with a basic problem. We still haven't solved a physics problem. All this eco-research is going to be meaningless. They will be prepared for something...but we have no idea what for. So there is that type of frustrations...My feeling is that anytime you have a science that relies so heavily on the eloquent speakers and what not, which I think most of ecology does because people are very good at telling stories, I get suspicious because where is the hard evidence and how does the evidence come together to really make their story? People are really good at making arguments...when you have such open-ended systems how do you really explain what is going on and is that the only answer? What are you finding? That is the first question that comes to mind is whether this has really gotten us anywhere or not. [explains that one counter example kills the whole point] [cDFC2, int]

In contrast, not all the scientists presented such a view. Some gave equal importance to investigative methods. In response to the VOSI-Sci question #5, the structure/function correlation, one scientist said:

That is where being scientific is hugely important. These types of studies [teeth/food source example] are important. It might not get down to the cause of why the difference is there or why the homeobox that caused these teeth to have these shapes. There is a difference in shape. Can you explain how...have a reasonable hypothesis of why it might be that way...Do I consider it to be an experiment? Not so much. It is an observation. He has not
actually done a manipulation to see. generate a hypothesis and manipulate the system to see if this does vary. You can still publish this. You can still address the information. Uh... but it is not actually testing a mechanism. It happens all the time in medicine. [OEL1, int]

This response also provides an example of a scientific investigation that produces justifiable claims that does not have an up-front hypothesis. Another response representative of equal status among investigative methods is seen in the following:

R: if you cannot do manipulations, is it still scientific?
S: yes
R: and are your conclusions equally valid?
S: sure, you can't manipulate gravity, for instance. You can only measure it. But you can make predictions about what the strength of gravity should be in different places and then go measure them to see if the predictions are validated.
[SEDF1]

Finally, within this subcode were suggestions that differences in approach were related to the maturity of the science discipline. Three participants posited a view that descriptive sciences were immature sciences that have not yet produced enough understanding to become experimental. Geology and ecology were provided as examples of such immature or "early phase" sciences.

Geology is still in the descriptive phase. This type of mineral is found here; that type is found there...[cDFC2, int]

I think science varies in degrees of intellectual development and varying degrees of honoring of this scientific method. I think there are a number of disciplines, for example, that are very very much tied up with description. That tends to be early in the development of a scientific field where people just don't know a lot. You've got to assemble the basic stuff from which you make the more causal connections.
R: Can you give an example?
S: I think a lot of ecology. You are working with cryptic organisms .... We won't understand it. What is the role of predation? What is the relative importance of predation? A lot of what you have to do in that case...it is hard to keep track of all the population. There are big broad areas; they live in the ocean; they live in the wilderness. My job is easier. I work with rivers and water. The laws that I work with are primarily physical laws, laws of force
and gravity, density, temperature. Biological systems are much more complex. In order to understand biological systems, before you get to the point of saying, "I think that the salmon are declining because of x and y" you have to understand a lot about the basic building blocks of this fish. That involves a lot of observation. Whereas science ultimately wants to be headed in that direction, it is not there yet by nature of dealing with nature. ... Is it still science? Yes, they are all science. I think they are all to varying degrees scientific. I think sciences that are most mature are those that have integrated experimentation. Physics, chemists, biologists on the molecular level, use experimentation as the primary driver of how to understand things. [GFD2, int]

Discipline differences: Complexity of systems, ability to repeat. Nine participants indicated that methods differ according to discipline, and these differences are related to the complexity of the system under investigation. The examples within the “hierarchy” subcode that refer to immature and mature sciences are also included in this subcode. In general, the participants who expressed disciplinary-based differences identified a relationship between increasing complexity and decreasing ability to (1) identify and control variables and (2) reproduce the approach.

I tend to take a highly reductive approach to my science, and I recognize that other fields have far too many interrelating variables for that approach to be successful. I can isolate important variables (temperature, pressure, reaction solvent, compound structure) that influence how a particular chemical transformation occurs. Attempting such isolation of variables within a single cell, much less a larger organism, is much more difficult, and thus while biology can gain insights from the kind of work I do, biologists could not routinely use the same approach. [gEL3, vosi]

Research done in the field always has the problem of experimental vs observational research. Landscape research also has the added problem of replication. Some feel that without replication, there is no research. Regrettably, one can't replicate landscapes very easily and one is left with describing what is seen under a particular landscape configuration. If enough different landscapes are studied, one hopes that general trends will materialize. Lab work is very conducive to experimental research. [mDF1, vosi]
The following excerpt from the interview with the cell biologist demonstrates connections among his view of validity in methods and the complexity of the system under study. Although he states experimental and descriptive sciences are valid approaches, his following comment about ease of investigation and “finding the right answer” suggests he considers them to hold different status with regard to certainty.

S: Nearly as I can tell, it has all been looked at. So if you don’t have an experimental, if you don’t have a function/structure relationship, you don’t have a publication and you really haven’t learned much that isn’t already known.
R: Are you talking about specific to your field, or in science in general?
S: Specifically in biology, cell biology.
R: What about other types of science, like physics?
S: You mean in terms of being experimental or descriptive? I think if you are an astronomer you are limited to descriptive because it is hard to get stars into your lab. You have to hunt around to find ones that are doing what you want to know about. It clearly can only be descriptive.
R: Do you think it influences the validity of the findings?
S: Not necessarily. I mean again, it depends on what the hypothesis is you are trying to support. ....I have a lot of colleagues who are ecologists and we are always arguing about observations. They try to make experiments but when you go out to the field, you can't control what is out there. The only way you can control it is by describing what you encounter and pick out the ones that change and try to make sense out of it. They say that is experimental, and I say probably not. ....I think ecologists do good work, but in some ways it is a lot harder to do good ecology than it is to do good molecular biology.
R: why is that?
S: because the answers are easier to get. Experiments are easier to control. You can design yes/no experiments...You find your answer." [SEL1, int] 

**Discipline-based Comparisons**

In general more scientists fell within the multiple methods subcode than within the single scientific method subcode, regardless of discipline or approach group. The LS and Ch (50% from both groups) explicitly posited that all methods were hypothesis driven. In contrast, none of the ESS or Ph participants indicated that all methods were hypothesis driven. There seemed to be a fairly equal distribution of
all groups represented in the “hypothesis not required” subcode. Even though only six scientists in the total sample suggested there was a single scientific method, four were LS, while none were Ph. Furthermore, slightly more LS viewed experimental methods as more valid than non-experimental (50% LS vs 37.5% total). Relatively fewer of the Ph considered validity differences in methods or disciplines (1 out of the 9 total). However, the Ph did report theoretical methods as valid method of scientific investigation, whereas none of the Ch included such descriptions.

**Approach-based Comparisons**

The E and E/D scientists (50% and 40%, respectively) had a higher tendency to explicitly posit that all methods were hypothesis driven. In contrast, none of the D or T participants indicated that all methods were hypothesis driven. More E/D participants viewed experimental methods as more valid than non-experimental, with 60% being the highest frequency of any of the groups. Like the Ph, the T scientists were the only scientists to include theoretical methods in their description of multiple scientific methods.

**Anomaly**

Tables 26 and 27 present results for scientists’ views of recognition and handling of anomalous data. The scientists described an anomaly as an observation that is inconsistent with an expectation. They spanned several subcodes in describing what they do once an anomaly is recognized.
Scientists’ Views

*Initial identification of anomaly.* Identification of an anomaly requires the scientist to have expectations and background knowledge of the subject under investigation. In this way, several scientists’ responses revealed the importance of the theory-laden nature of science and creativity.

Anomalies are identified when data exceed expected bounds of predicted outcomes. This notion that we enter into experiments or observations without bias is not really accurate. Scientists usually have some idea of what is happening. [bEL3, int]

When we build these models and test them, we play games like this. We try to develop a test where we know what we should expect. We predict the results and see whether we get them or not. We see the failure of the prediction and start probing and say, “How come?” [cDFC2, int]

Expectations require a certain amount of knowledge. You recognize it based on your experience and training that you have certain expectations for a pattern of data or observations and you can only have that expectation after a lot of experience doing and reading and manipulating certain things. You recognize something doesn’t fall within that pattern... Without that experience [knowledge base] it is really hard to look at an outlier and anomalies in data, and to be able to have any intuitive sense, and a lot of it is intuitive... I think I wrote to you about the creativity in science... you can’t bring creativity to science without a lot of knowledge base to go with it. ...It is hard to get the real excitement of a discovery unless you have the knowledge base where that discovery means something. If it is just a piece of information that is stand alone, then it is just a piece of information. Unless you’ve got some mental construct and you get a piece of information that is outside of that and you’ve got some reason for believing this and this is outside, this is terribly exciting. I don’t know how you give somebody that experience if they don’t have the knowledge base. [BEFL1, int]
Table 26. Scientists' Views of Anomaly: Grouped by Discipline

<table>
<thead>
<tr>
<th>Aspect</th>
<th>Subcode</th>
<th>Total #</th>
<th>Total %</th>
<th>Grouped by Discipline</th>
<th>Number of scientists</th>
<th>Percentage of group</th>
</tr>
</thead>
<tbody>
<tr>
<td>Anomaly</td>
<td>inconsistency with expectation</td>
<td>17</td>
<td>70.8</td>
<td></td>
<td>10</td>
<td>LS 80 ESS 80 Ph 60 Ch 50</td>
</tr>
<tr>
<td></td>
<td>error in measurement: repeatable?</td>
<td>14</td>
<td>58.3</td>
<td></td>
<td>11</td>
<td>40 80 ESS 60 Ph 75</td>
</tr>
<tr>
<td></td>
<td>excitement/progress</td>
<td>10</td>
<td>41.7</td>
<td></td>
<td>8</td>
<td>60 20 ESS 40 Ph 25</td>
</tr>
<tr>
<td></td>
<td>natural variation?</td>
<td>3</td>
<td>12.5</td>
<td></td>
<td>5</td>
<td>10 20 ESS 20 Ph 0</td>
</tr>
<tr>
<td></td>
<td>Expand existing model</td>
<td>11</td>
<td>45.8</td>
<td></td>
<td>7</td>
<td>50 80 ESS 20 Ph 25</td>
</tr>
<tr>
<td></td>
<td>develop new model/theory to explain</td>
<td>7</td>
<td>29.2</td>
<td></td>
<td>1</td>
<td>10 40 ESS 20 Ph 75</td>
</tr>
<tr>
<td></td>
<td>discovery</td>
<td>2</td>
<td>8.3</td>
<td></td>
<td>1</td>
<td>10 0 ESS 20 Ph 0</td>
</tr>
<tr>
<td></td>
<td>ignore/discard</td>
<td>0</td>
<td>0.0</td>
<td></td>
<td>0</td>
<td>0 0 ESS 0 Ph 0</td>
</tr>
<tr>
<td></td>
<td>set aside for rainy day</td>
<td>3</td>
<td>12.5</td>
<td></td>
<td>0</td>
<td>0 20 ESS 0 Ph 50</td>
</tr>
<tr>
<td></td>
<td>depends on context</td>
<td>2</td>
<td>8.3</td>
<td></td>
<td>1</td>
<td>10 0 ESS 20 Ph 0</td>
</tr>
<tr>
<td></td>
<td>report finding/no explanation necessary</td>
<td>1</td>
<td>4.2</td>
<td></td>
<td>0</td>
<td>0 0 ESS 20 Ph 0</td>
</tr>
<tr>
<td></td>
<td>other</td>
<td>4</td>
<td>16.7</td>
<td></td>
<td>1</td>
<td>10 0 ESS 40 Ph 25</td>
</tr>
</tbody>
</table>
Table 27. Scientists' Views of Anomaly: Grouped by Approach

<table>
<thead>
<tr>
<th>Aspect</th>
<th>Subcode</th>
<th>Total</th>
<th>Total</th>
<th>Grouped by Approach</th>
<th>Percentage of group</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>E</td>
<td>E/D</td>
</tr>
<tr>
<td>Anomaly</td>
<td>inconsistency with expectation</td>
<td>17</td>
<td>70.8</td>
<td>7</td>
<td>5</td>
</tr>
<tr>
<td></td>
<td>error in measurement: repeatable?</td>
<td>14</td>
<td>58.3</td>
<td>5</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td>excitement/progress</td>
<td>10</td>
<td>41.7</td>
<td>4</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>natural variation?</td>
<td>3</td>
<td>12.5</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>Expand existing model</td>
<td>11</td>
<td>45.8</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td>develop new model/theory to explain</td>
<td>7</td>
<td>29.2</td>
<td>4</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>discovery</td>
<td>2</td>
<td>8.3</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>ignore/discard</td>
<td>0</td>
<td>0.0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>set aside for rainy day</td>
<td>3</td>
<td>12.5</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>depends on context</td>
<td>2</td>
<td>8.3</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>report finding/no explanation necessary</td>
<td>1</td>
<td>4.2</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>other</td>
<td>4</td>
<td>16.7</td>
<td>1</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td></td>
<td>24</td>
<td>100</td>
<td>10</td>
<td>60</td>
</tr>
</tbody>
</table>

E: experimental; E/D: combination of experimental and descriptive; D: descriptive (non-experimental); T: theoretical
Anomalies were also described in terms of intuition not fitting the existing model. In such circumstances, the scientist wants to alter the model, but there remains the need to find supporting data. This view also demonstrates a connection between model use, identification of an anomaly, acknowledgement of subjectivity, and the need for empirical basis. In the following example, the scientist discusses how his ideas do not fit with the existing model or the existing data. The excerpt begins after the scientist indicated “anomalies are cool.”

R: All the time? Do you ever go, “Ah! What’s this doing here?”
S: Oh, sure, but... not very much.
R: Ok, so it’s cool. So what do you do?
S: You try to figure out what it is.
R: It is intriguing?
S: Yeah. We had to deal with one today. It came up in the last week or so and we had to figure it out. There are these growth models that people use to show that plants are smart. They can actually forage like any animal could. So if you think of a clonal plant that makes another plant, sends out a runner, makes another plant, sends out a runner, makes another plant. People have actually figured out that those plants might be really smart, and the way they are really smart is that when the little runner hits a good environment, it branches a whole lot and it makes very short runners. So that makes it very likely that it will stay in that good area. And when it is in a bad area it sends out really long runners to get out of that area and it doesn’t branch. People have actually shown that plants do this. They fit this model to some degree. Well, a few years ago I had this idea that that simple model had to be wrong. People did all these computer simulations and all that. And I thought that it was wrong because it was all based on the idea that plants live in a dichotomous world: there are good areas and bad areas. Get out of a bad area and stay in good areas. But the problem is that the world isn’t dichotomous. It is a gradient. So you could put a plant at the bottom of a bad slope, ok, it has to climb to the top to get to a good areas. Using the strategy I just described, it might go a little ways up the hill and find it is better off than it was before and try to stay there. So, I didn’t think the simple growth idea of branding off in short internodes would work to get plants out of bad areas. It only works to get them out of the worst areas. So we built a big simulation model to do it, a big computer model that actually followed plants growing in all sorts of environments and patchy environments, dichotomous world or gradient where you have constantly changing resource levels across a landscape in some way. It seems to predict exactly what everyone always said using these simple growth strategies of changing the runners to these branching...
probabilities, do a pretty good job of running up hillsides. It doesn't make any sense to me. I keep thinking there is something wrong with the model...and so you work over the model and go okay, the model works pretty well. We actually discovered several things were wrong and fixed them. At some time you reach the point and go “Wow this is pretty cool. It does work. Why does it work?” It is not intuitive. It is kind of fun. It is also kind of scary because you make a model like that and the first time you make the model and you get the answer you want you have the tendency to stop there and there could still be an error in the model. That might be giving it to you but you got the right answer and you stopped.

_Rearable: True anomaly._ Once an anomaly is initially identified, 14 of the scientists (58.3%) reported a need to identify the anomaly as a true anomaly, as opposed to an error in the procedure. A true anomaly would be repeatable and predictable after the scientist has established that all equipment and materials are functioning properly.

[Once recognize an anomaly] this is where the repeating comes in. Usually one starts with the premise that you did something wrong. There is a dishrag in the works...a monkey wrench. So you look and verify that the equipment is running according to specifications...that you didn't make a blunder in preparing the sample or aligning the instrument, tuning it, something like that. You look for obvious mistakes....If you can't seem to account for this anomaly on that basis, then you have to begin to question the premise or hypothesis that is the premise behind the experiment or the hypothesis, that nature is not behaving in accordance with whatever your expectations were, assuming you had some to begin with. [bEL3, int]

There could be different ways of getting there. It's when you do an experiment that you think is pretty obvious and you know the outcome and it doesn't turn out that way. You repeat it to make sure you didn't reverse the test tubes and it still comes out wrong. You do that over and over again. It is frustrating, uh...and a lot of people think well...How do you have the patience to do it? That is actually where real progress comes from because when you understand that you realize you weren't thinking about it very well, and nature had a much better solution for the whole thing. [OEL1, int]

One of the physicists described several approaches to addressing anomalies. First, are the data wrong? The scientist needs to determine the reliability of the
anomaly. Second, the anomaly may not invalidate the existing framework because
the framework is a simplification of reality. As such, competing theories result.
Third, the anomaly may overturn the existing theory. He reported that scientists are
reluctant to accept this outcome because an established theory has a large body of
evidence. Within this situation, there is a need to blend a large body of evidence with
a small body of new evidence. In the event that the new finding is repeatable and
continues to disagree, there may be a need to abandon the current theory and develop
a new theory. “Scientists do that. It doesn’t happen that often.” He described science
is self-cleansing of false anomalies when they are not repeatable. This notion of
repeatability to identify a true anomaly is on a broad scale within the scientific
community and introduces the importance of peer review.

Anomalies are very important. The worst thing that could happen is to ignore
them. Generally scientists don't do that. If the data is respectable with the way
it was collected......One of the wonderful things about science is that it
cleanses itself of false anomalies. What I mean by that is the following. If
somebody pops up with a new observation that disagrees with existing ideas,
and the observation is important enough, this is the catch...meaning that it is
central enough to the subject. Then it will get repeated. The experiment will
get repeated or someone else will do it. It is a policing of this. It is quite
common in science for these anomalies to go away because....someone else
won't get the same thing or say the data are wrong. In that sense science is
self cleansing on important issues. [sTC4, int]

*Excitement/progress.* Once a true anomaly is identified, there were several
options and steps of options the scientists described. The first reaction expressed by
10 of the scientists was a feeling of excitement. These scientists reported an anomaly
means something new and interesting, and this is how science progresses.

We all secretly love that [finding an inconsistency] ...[laughs] because then
we have something to do ...that is the way science progresses. [eDF2, int]
Anomalies are cool. Anomalies are what make science fun...You try to figure out what it is. [MEDF1, int]

That is when things get fun. You've got to figure out did somebody screw up and write down the wrong number. Is this really an exception to the rule and if so what is going on here? You've got a lot of work to do. But that is the entry into the most wonderful part of science, opens doors to serendipity, to stumbling onto something totally fortuitously. That opens up whole new avenues. [BEFL1, int] [emphasis added]

That is when you come up with the new stuff, is when you absolutely can't rule out the possibility, that there is something new...some bit of biology there. That is the fun stuff. That is why you get up in the morning, for the things that don't fit. If all the data fit every existing theory, we'd be out of work. [KEDF1, int]

The wildlife biologist explained her position on finding anomalies as she described a situation at a conference when she presented results that were different from another presenter on the same topic.

People came up and said 'doesn't it bother you that she reported results exactly the opposite of you?' I said, not in the least. To me I personally feel the only way we are every going to get answers as far as habitat is by looking at those anomalies and asking why they occur. What are my other options? My option is to say, either you or I did something wrong...I know I didn't, and so therefore you must have. We both are thinking the same thing, so where does that get us? Or else you say, we both did it the best that we could and the reality is that is what we got. So why did we get that? Why under your situation did you get this and I got just the opposite finding? I think the reason at least in habitat is that you don’t have exactly the same stuff going on here as is going on there. [mDF1, int]

**Expand existing model/framework.** Eleven scientists explained that they examine the anomaly within their existing framework and attempted to explain within that framework or expand the existing model to accommodate the new information.

What I would probably do is I would consider first if there was a way of restating my own ideas that was consistent. If it were really different, then I would have to look harder. What I would really want to do is explain why I
found what I found and they found what they found and the two are not the same. That would raise a whole set of questions that would be important. So there is a sequence of steps that originally you look very hard at where that anomaly came from. Is it good data or bad data? Was it data taken the same way? Is the data consistent with the interpretation if you look at it in a certain way? [GDF2, int]

In the case of the inconsistency noted in 4, I've done my best to alert the cloud modeling community to my findings. For the most part, because they cannot explain my findings, they've chosen to ignore the findings— all but one cloud modeler, who thinks my findings are flawed. I've worked with him to ensure that his concepts have seen the light of day (i.e. published) even though I think there's something wrong with his model. It's not producing what I'm finding, and it's difficult to see how my results could be flawed. What's wrong with the model, or with my analysis of observations is a mystery that my take years to resolve. This give and take is simply an ongoing part of learning. [cDFC2, vosi]

In discussing competing models for the same anomaly, an atmospheric scientist described the need for better analysis and refinement of his model to explain the data. His statements indicate a critical role of creativity.

We are going to get better at our analysis of our data and when we do that it gets harder for people to say, “Ah.” or how do you say, it motivates people to start looking at the model and ask what is really going on here. How do we understand this? Obviously there is something strange going on here. By pursuing this and keeping the pressure up, I am hoping that people like John [colleague] will come along and start thinking again. “Well maybe if I did something else in my model... maybe we could pull this off.” [cDFC2, int]

The following is an excerpt from the interview with a molecular biologist who explores mechanisms of translation in prokaryotes. He discusses a situation where his expectations were not upheld by the data he collected. The exchange details the process he went through in dealing with the anomalous data. Rather than abandoning his theoretical framework, he alters his hypothesis and offers possible explanations based on the model that drives his work.

R: This one you talked about creativity happens in all parts of investigations. You gave the example of you designing an experiment and you got initial
results that did not meet your expectations. And this was the CGA you though was a stop codon. It didn’t work, and this lead you to conclude that it was poorly translated, but not a terminator. Can you relate this example to the question on anomalies? How was it recognized that this did not meet your expectations and what was the process you went through, both the thought process and the physical investigative process.

S: Ok, I’ll have to think back to how I figured out it wasn’t a terminator. Ok., I put in a large cluster of CGA codon into the gene and measured how much protein was made. It didn’t seem to matter how much CGAs, I got the same amount of protein. I also over-expressed the release factor. UGA is a stop codon. that didn’t have any effect on expression. So I was thinking that if the codon is read as a terminator and you over-express the release factor, it should be read more as a terminator, but it didn’t. So I figured something else had to be going on. It did take some creativity to come up with the experiment, with the hypothesis and the experiment. The experiment didn’t give the results I was looking for. CGA is still a rarely used codon. There is something wrong with it, but what could it be? And so it turns out it is slowly translated. This is the speedometer aspect that you are familiar with. We can show that the CGA codon is slow in selecting aminoacil tRNA, first of all. Second, if you put in ....CGA is a lousy context for reading the next codon. It is the last base pair, the AI base pair which is an unusual base pair, a wobble base pair, and that interferes with reading the next codon. That is why CGA is not useful. It interferes with two translational cycles. It is slow to be translated and it makes the next codon slow as well. That was totally different from the initial hypothesis. The initial hypothesis was based on the rare CGA in genes and the fact that it looks like UGA. Chemically UGA and CGA are almost the same things.

R: you recognized it did not come out as you expected and so you changed the expectations?

S: New hypothesis. Maybe it has some other attributes. Maybe it is slow to select tRNA. Maybe it is slow to select the next tRNA, next codon. So anyway, you read that and look at the literature, and put it all together. You make a conclusion about CGA codon. Hopefully we are going to find out the UUPy is a fancier start site in RNA polymerase gene. It is really rare, but occurs in RNA polymerase genes. It is somehow doing something there. Probably frameshifting, but we don’t know. You come back and work on that one. [UEL1, int]

Develop a new model. Seven scientists stated they may develop a new model to explain the anomaly. Five of these scientists also appeared in the previous subcode of “expand existing model.” Ideas within the current subcode related to strict notions
of falsifiability and, moreover, methods of paradigm shift within the scientific community.

You can only disprove. With one exception, if it is repeatable, the theory is gone. [bEL3, int]

S: Now, where you see paradigm shifts is when anomalies tend to add up and add up, often times in contradictory ways. So one of the ways to deal with this is to make the ad hoc correction to the hypothesis. That should lead to further predictions of where you should find what would be anomalies for the original hypothesis. And so you go and test that and either the original hypothesis is correct or you will really generate more anomalies. In an extreme case you generate so many contradictory anomalies that you need something completely new to accommodate everything you’ve got.

R: how often do you think that happens in your field?
S: not on a day to day basis, but certainly uh...I’m aware of year to year types of things that force people to rethink. [gEL3, int]

One of the astronomers gave details of a specific anomaly that lead to profound progress within her field.

S: There are other cases where our basic underlying model fundamentally changes. A recent example of that has to do with field of star formation where now understand, since 1995, that the formation of planets around stars, when stars are forming, is very common. We had no idea before the discovery of the first planet that was true. And we had no idea that we would find large planets like Jupiter very close to their central stars. We’ve got Jupiter out at 5 astronomical units in our system. But many of the planetary systems that we have found so far, there are planets like Jupiter closer than Mercury to their central stars. That is a huge change in how we look at the origin and evolution of planetary systems. They were unexpected.

R: How is that changing how you are looking at these systems?
S: it is changing the whole set of questions we ask. Instead of asking “do stars have planets” we are asking “how did these planets get here?” “what sort of process would lead to a planetary system that looks like this one and what impact does that have on the existing of planets like Earth?” our fundamental question is “are we alone?” what we expected to find are planetary systems like ours. What we have found are planetary systems that are very different than ours. And if in fact the majority of stars form Jupiter like planets that migrate in toward the central stars, and actually are destroyed because they impact the central star, that would say the existence of planets like Earth around other stars are very uncommon because they would be destroyed in this migration process. They would form, but be destroyed. That would make the existence of other Earths around other stars extremely rare.
That is an important question for us to know the answer to. And so because of this discovery we really have reframed the whole set of questions we are asking about life in a different set of models. Models of how we understand the origin and evolution of planetary systems has changed. We are exploring a whole different set of questions than we would have predicted 10 years ago. That is a bigger sort of model. It is a fundamental paradigm change in how we view planets. [pEDFC2, int]

*May set aside but do not disregard.* None of the scientists suggested ignoring the anomaly completely. Several suggested that in the event the anomaly could not be explained, it might be set aside for further investigation at a later time or included as part of the published report, acknowledging the outlier without explanation for its appearance.

And it would also be true that okay sometimes you will get something and you get maybe some copepod which is giant, and all the other ones were small. You go, Hmm, what is this weirdo here? And so, sure, there is the temptation to exclude that data point, but unless you can come up with a valid reason such as you gave it Wheaties and the other didn’t get Wheaties, then perhaps you can exclude that. There is always that temptation to make your data look better by excluding the oddballs, but if you adhere to the scientific method and are a good scientist you shouldn’t do that. [NEFL1, int]

My graduate mentor once had a series of many experiments where in two temperature regimes, data for a rate constant fell closely along a linear model. His student did one last experiment that was at an intermediate temperature, but the result fell nowhere close to the linear model. He published the full set of data, yet noted the lack of reproduced data for the anomalous point... It’s human nature to gloss over inconsistencies when one is trying to make an argument. A classic example lies in a reinvestigation of Gregor Mendel's publication of his original work in genetics. Modern analysis shows that the data he reported, while fitting his proposed model, has got to be incorrect according to current knowledge of genetic behavior. It is clear from his original records that he ignored “anomalous” data on the presumption that it was caused by experimental uncertainty/variation or other error. Had he included this data, his model may not have been accepted, yet by ignoring this data he committed what by today's standards would be recognized as scientific fraud. [gEL3, vosi]
A couple of participants discussed the reality of not being able to progress if one focuses on all the anomalies. They are set aside and maybe addressed at a later time, maybe not. In any event, science must progress:

I often see things in data that look interesting; interesting in that I didn’t expect it. I jot it down. The trouble is that I don’t get back to all these lists. There is always something to do. With satellite data, you are talking about billions and billions of bits of data. Evidence we might talk about one thing, where all these billions of bits are pointing. Making sense of these billions of bits…that is where the challenge is. And so maybe I can make sense of 999 million, but there are probably always a few that I can’t make sense of. Life is too short to worry about it. I am just not going to do it. The thing is, if I really did worry about it, I’d never get anything done. Period. [cDFC2, int]

You can’t be a perfectionist in this field. People who are either go into theoretical physics or become pathologists or they fail. Because there is no perfect experiment. There is no perfect set of data. [OEL1, int]

Discipline-based Comparisons

Most of the scientists who expressed excitement in relation to identifying an anomaly were LS (6 of the 10 in the LS group; 6 of the 10 total in this subcode). Interestingly, the LS and ESS indicated “expand existing model” more often than the “develop new model/theory to explain” option. Also, the LS and ESS groups showed preference to the former option more than the Ph and the Ch (9 out of the 11 total). The Ch reported “develop new model/theory to explain” more often than the other groups (75%). These results suggest slight differences among the disciplines in whether they prefer to work within their existing model (LS, ESS) or consider developing something new (Ch) in dealing with anomalous data.

In general, the physicists did not comment about altering or developing models when addressing anomalies. Their tendency was to (1) discuss anomalies in terms of mathematical inconsistencies that were resolvable through mathematical
adjustments; or (2) state the inconsistent data should be reported but not necessarily explained. The existence of true anomalies was not clearly articulated by the theoreticians. Again, they spoke in terms of mathematical inconsistencies rather than empirical inconsistencies:

In my work, two or more scientists can find different answers to the same question. Since we are working with mathematical models, not sketchy data sets, this cannot be just an example of different interpretation. Ideally when an inconsistency is found, the scientists get together and try to find the true differences in what appears to be the same question they are both asking. It may be that what seemed to be an innocuous detail is a distinction in the two computations that explains the different conclusions. It may also be that one of the researchers is wrong. THIS HAPPENS ALL THE TIME IN MY FIELD! In most instances, cooperation results in a resolution of the problem. [pTC4, vosi]

**Approach-based Comparisons**

Most notable in this group of comparisons is the difference in E/D preference for the “expand existing model” subcode (4 of the 5 E/D participants) versus none of E/D scientists appearing in the “develop new model/theory to explain.”

**Justification**

The VOSI-Sci items #5 and #6 and the VOSI-Sci item #8 elicited scientists’ views of justification of scientific claims. During the interview, participants were prompted to describe what they consider to be requirements within their field, and science in general, to justify claims. Results are presented in Tables 28 and 29. Most prominent features brought forth from the discussions on justification included views of reproducibility (internal and external), validity of different investigative approaches and disciplines, peer review, and a need to address alternatives.
Table 28. Scientists' Views of Justification: Grouped by *Discipline*  

<table>
<thead>
<tr>
<th>Aspect</th>
<th>Subcode</th>
<th>Total</th>
<th>Number of scientists</th>
<th>Percentage of group</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>#</td>
<td>LS</td>
<td>ESS</td>
</tr>
<tr>
<td>Justification</td>
<td></td>
<td>%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>reproducibility</td>
<td>internal: statistics</td>
<td>11</td>
<td>7</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>external: multiple</td>
<td>3</td>
<td>0</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>researchers/same results</td>
<td>3</td>
<td>0</td>
<td>2</td>
</tr>
<tr>
<td>consistency with others</td>
<td></td>
<td>10</td>
<td>7</td>
<td>1</td>
</tr>
<tr>
<td>experiments over description</td>
<td></td>
<td>9</td>
<td>3</td>
<td>1</td>
</tr>
<tr>
<td>peer review</td>
<td></td>
<td>5</td>
<td>3</td>
<td>0</td>
</tr>
<tr>
<td>address alternatives</td>
<td></td>
<td>8</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>predictions/tests</td>
<td></td>
<td>2</td>
<td>0</td>
<td>2</td>
</tr>
<tr>
<td>model/ predict/test</td>
<td></td>
<td>3</td>
<td>2</td>
<td>0</td>
</tr>
<tr>
<td>scientific method/hypothesis test</td>
<td></td>
<td>14</td>
<td>6</td>
<td>2</td>
</tr>
<tr>
<td>differs with discipline/context</td>
<td></td>
<td>2</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>not clear from response</td>
<td></td>
<td>4</td>
<td>2</td>
<td>0</td>
</tr>
<tr>
<td>other</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Table 29. Scientists' Views of Justification: Grouped by Approach

<table>
<thead>
<tr>
<th>Aspect Subcode</th>
<th>Total #</th>
<th>Total %</th>
<th>Grouped by Approach</th>
<th>Number of scientists</th>
<th>Percentage of group</th>
</tr>
</thead>
<tbody>
<tr>
<td>Justification</td>
<td></td>
<td></td>
<td>E</td>
<td>E/D</td>
<td>D</td>
</tr>
<tr>
<td>reproducibility</td>
<td>24</td>
<td></td>
<td>11</td>
<td>45.8</td>
<td>6</td>
</tr>
<tr>
<td>internal: statistics</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>external: multiple researchers/same results</td>
<td>3</td>
<td>12.5</td>
<td>1</td>
<td>0</td>
<td>2</td>
</tr>
<tr>
<td>consistency with others</td>
<td>3</td>
<td>12.5</td>
<td>1</td>
<td>0</td>
<td>2</td>
</tr>
<tr>
<td>experiments over description</td>
<td>10</td>
<td>41.7</td>
<td>4</td>
<td>4</td>
<td>1</td>
</tr>
<tr>
<td>peer review</td>
<td>9</td>
<td>37.5</td>
<td>4</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>address alternatives</td>
<td>5</td>
<td>20.8</td>
<td>4</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>predictions/tests</td>
<td>8</td>
<td>33.3</td>
<td>3</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>model/predict/test</td>
<td>2</td>
<td>8.3</td>
<td>0</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>scientific method/hypothesis test</td>
<td>3</td>
<td>12.5</td>
<td>1</td>
<td>2</td>
<td>0</td>
</tr>
<tr>
<td>differs with discipline/context</td>
<td>14</td>
<td>58.3</td>
<td>4</td>
<td>4</td>
<td>2</td>
</tr>
<tr>
<td>not clear from response</td>
<td>2</td>
<td>8.3</td>
<td>1</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>other</td>
<td>4</td>
<td>16.7</td>
<td>1</td>
<td>1</td>
<td>0</td>
</tr>
</tbody>
</table>

E: experimental; E/D: combination of experimental and descriptive; D: descriptive (non-experimental); T: theoretical
Scientists' Views

Reproducibility. Internal reproducibility (11 participants) deals with replicas within an investigation and measures of statistical analysis. In response to VOSI #5, a participant responded,

Do I consider this persons investigation [VOSI #5] to be scientific? The answer really comes down to how does he present his data and how does he present his conclusions. Are the results there that he can reproduce? [OEL3, int]

R: Ok. Let’s switch to your work now. Let’s say you’ve done investigations and compiled your research and say you are ready to publish. What is critical in your work to justify and have your claims be accepted?
S: Well, I guess for me, and I am talking just about my experimental work. When you run an experiment, first of all, you want replication. Replication is important. Sometimes I will get a paper to review where the person has replication but the person has only done the experiment once. That is suspicious to me. I like to see experiments repeated. The question is how often do you have to repeat them.

Well, I guess it depends on how much the data resemble one another. I don’t run out and publish experiments we have only run once. I want to do them two or three times at least. [NEFL1, int]

There is statistics in there, a certain number of replications, treatment of the data. Then when you say I am confident in this type of analysis, you can go out measure something with time or in a different system and measure something else. In order to move into a study where you are interpreting behavior of chemicals, as analytical chemists, we take great pains to make sure our data are solid; the actual numbers are solid. [IEF3, int]

The three most recent ‘must do’ items for wildlifers in the last several years have been: power analysis (statistics), AIC analysis (statistics), variable density sampling (a small mammal capture design). Scientists, if anyone, should know that "one size doesn't fit all," but they appear to be a "fashion conscious" group (to my personal dismay). [mDF1, vnos]

Often times I am sitting here reviewing a manuscript that an editor has sent me to review. You will look and there will be some experiment the author has conducted and they will draw conclusions from the outcome of this single experiment. Many times I or another reviewer will raise the question of “this
is only one observation." There is no evidence in your work that you even repeated the experiment yourself to give some statistical measure of its reliability. That you can do internally within your own laboratory. Always we like to see multiple observations of the same thing. That gives an indication of the variation within our own experiment. That enables us to establish reliability. If that variation is small compared to the magnitude of the observation, then we are confident that this is probably a good result. We say it is statistically significant. But if the variation is large, compared to the magnitude of the result, we will have very little confidence in the result. [bEL3, int]

External reproducibility (3 participants) deals with the notion of multiple researchers conducting similar investigations and getting the same results. This was considered separate from “consistency” (3 participants) which was more in reference to theoretical framework and/or current paradigm within the community. Again it should be noted that participants were not restricted to only one subcode, but single quotes were often applicable within multiple subcodes as well as main categories.

A new scientific claim in astronomy is the existence of a new form of energy called ‘dark energy’ which in fact comprises most of the total mass-energy of the universe. The acceptance of this claim has required investigation by several different teams of astronomers using different techniques. The claim must also be consistent with all existing data. [pEDFC2, vosi, 6]

[In reference to VOSI #6, justification of natural selection] Based on just these observations no. That conclusion would need a broader context of data on other physical characteristics and environmental variables. After all, even polar bears like to eat carrots. [pEDFC2, vosi]

**Experiment over description.** Consistent with the findings for “Methods,” is the subcode of “experiments over description (10 participants, 41.7%).” VOSI item #6 sought respondents’ ideas about justifying a causal claim (natural selection) from a correlation investigation. The following response demonstrates this scientist considers controlled experiments to be more valid approach than a correlational study. He rationalizes his position with claims of predictive function.
I mean, you would have to set up real experiments, I think...or, yeah, controlled situations in which you have controlled source of variability in the population and you have relatively limited sources of food and then... Otherwise you just have a correlation...You have to be able to do something with your conclusion, otherwise it is not justified. [eDF2, int]

The following responses suggest the added complexity of relying on qualitative, rather than direct experimental and quantitative evidence, makes causal claims problematic.

Likelihood of being wrong is much higher than in something like physics where you are colliding elementary particles together. The reason is that you are looking at something that is really rather complicated and the scientist in this instance is looking at one aspect of a full range of measurements or observations that might be relevant to the question. A chemist might come along, we actually have people who do that type of thing, and do isotopic measurements. You realize for example, it is possible by looking at your teeth to figure out what you eat, not just teeth in general, but there are particular molecules that arise...it is possible to go back to the bones, for example of Native Americans and figure out what they ate....We are dealing with an area of science which is complicated and some are more quantitative than others. And this particular example you are talking about, from my perspective, is someone less quantitative and therefore, as a nonexpert looking at this...I am not an expert looking at this...I would find it interesting and worthy of publication, but I would retain substantial skepticism. In the end it would come down to how the results were presented, not whether the results were presented. [sTC4, int]

[In reference to VNOS #8, the dinosaur extinction controversy] The data are incomplete. Also paleontology is an observational science, which makes experiments impossible. The scientists have to work with what they see in the rocks, and make educated guesses about what caused the patterns that they see. If they could do an experiment (revive the dinosaurs and try dropping a meteorite on them) it might remove some of the ambiguity, although maybe not. [KEDF1, vnos]

Some scientists drew examples from other areas of research to suggest experimental results were more justifiable than non-experimental. Interestingly, the following quotes are from participants who engage in descriptive research or a combination of experimental and descriptive research.
Solid, defendable conclusion in science come from repeatable cause/effect phenomena. Because we don’t have dinosaurs nor can we experiment with meteorites and volcanoes, all we have to examine is the effect. In the real world, multiple things can cause the same effect. ...It is always difficult to infer cause/effect after the fact (which is why experimental research is far superior to observational research, if it can be done). [mDF1, vnos]

If you can do a clear cut experiment and test a hypothesis, you should go ahead and do so... You set it up first by presenting a hypothesis either from a model or from previous literature or whatever. So once you set up a hypothesis correctly, that is all the justification you need for the experiment. Ok, because that model should kind of state 'here's what I am testing” and if that is what you are testing when the experiment must be designed that is what you are manipulating and it is going to make it very clear that everything else is going to be controlled or explained. [MEDF1, int]

The entomologist studies the effects of clearcuts and forest conditions on invertebrate populations. He saw differences in confidence of claims between investigations he is able to do involving direct manipulations of forest plots versus natural plots.

If you establish plots where the storm went then you are accepting the fact that certain trees blow down because of where they are, so that whatever differences you see could possibly be attributable to local soil conditions or where in the draining it is, whether it is on a ridge or in a valley is going to affect whether it blows down...you know certain things like that. Where if you go into an area and randomly pick plots that are going to be subject to different treatments, then you have a lot more confidence that it is the treatment that is responsible for whatever the change you see. [SEDF1, int]

*Prediction/tests.* Eight scientists explicitly stated claims needed to be the result of tests of predictions or hypotheses.

Not justified about natural selection [VOSI #6]. Because this is now a hypothesis that has to be tested. [bEL3, int]

A similar example is the disappearance of many species of large mammals in North America in the Pleistocene. A long running debate has occurred as to whether this was caused by disease, climate change, or human hunting. What happens over time is that the debate leads to investigations that uncover new evidence and the weight of evidence tends to favor one hypothesis or perhaps a combination of explanations or something entirely new. [PEDF1, vnos]
Even when different disciplines employ different approaches, one scientist insisted the claims needed to be predicated by a hypothesis to be justified.

S: Things can change as more information comes in. What is constantly in front of you is a model for how nature works. Certain parts of that model are probably solid and are not going to change. Other parts may change. My guess is that in the case of astronomy, just because they can't go out and manipulate things, it doesn't mean there can't be strong components of their model. But some of the components will change, and be subject to refinement. It is a limitation experimentally and technically to carry out observations and experiments, but I don't think it affects the strength of conclusions, necessarily.

R: So, to be clear with what you are saying, conclusions can be equally valid regardless of if it is gained through experimentation, or descriptively, observation?

S: I think that is true. Again, even a descriptive one should have some sort of experimental or hypothetical...something driving it. A retrospective type of experiment. [UEL1, int]

**Differs with discipline/context.** The largest subcode included 14 (58.3%) of the scientists who indicated that means of justification varies by context, both within and across disciplines. Some of the scientists had difficulty answering questions about justification in other fields. As one physicist said, “It is not simple to answer. [Justification] is laden with context. It varies enormously from field to field. Context and comparison is important.” Others were better able to articulate differences based on context:

If I make a claim about a river today, there has to be physics. If I make a claim about a river yesterday or last century or last millennium, it will have a different kind of evidence. But is has to be consistent with what other people know. [GFD2, int]

Generally, we want a true manipulative experiment. However, consider scientists that use genetic similarity to determine evolutionary trees. They cannot go back in time to make observations or conduct experiments. They can test their methods, but not their conclusions – in many cases, this is all that is possible to make a scientific claim. [MEDF1, vosi]
Well, to a physicist, what can be done by experimentation is as close as you can get to the word of God. I mean, you literally can say "the word of God," meaning this is what nature is. This is the real nature that we like to explore. So you don't try to replace experiment...as a physicist. As a computational scientist, people in like drug research, they often will use the computer to develop a drug rather than nature because it is faster. And they develop the codes to the point where it agrees with nature perfectly. The same is true for bomb builders, the simulation of nuclear reactors, nuclear piles...it is all done computationally. But as far as a physicist...you don't beat experiment. That is the ultimate test. What you do have, though, is, anytime you do something on the computer, it is an approximation for the mathematics because you are approximating numbers, which can be intimately precise, which are only stored to a certain length. And so then you ask the question, are your results believable? Are they truly a solution or are they just approximate solution that has errors? So this now comes close, similar to experiment. Yes, you've done a measurement, but unless you know how big your uncertainties are, you don't how reliable that measurement is. [ITC4]
conditions constant. If you do, the results are meaningless...Ecology is sometimes viewed as a loose science by people in physics and chemistry. They are just not used to dealing with variation. Where ecologists, I mean, variation...part of our job is trying to figure out what the source of all the different variation are. But is hard sometimes talking back and forth and understanding that, yeah terrestrial forests probably are a source of missing carbon but we probably aren’t going to find 1.2 gtons and be that precise. That is not very satisfying for them...But I think it is important for scientists to talk to each other for just that reason. So you can start sharing these perspectives and understand that, yeah how you do research and what you get out of it isn’t the same in all the fields. [SEDF1, int]

Peer review. The role of peer review in establishing and maintaining justification standards was expressed by nine participants (37.5%).

The ultimate test is whether the conclusions serve to change the way others think about the world. I suspect that many conclusions are never check[ed], because it may not make much difference to others whether they are right or wrong. If the results are significant, they will be checked by others, and then the reputation of the scientist hangs in the balance, depending upon whether the project has been carefully done. [jTC4, vosi]

It is hard to fit into one sentence. What makes it valid is enough of your peers saying it is valid. They are basing their opinion on their experiences of what is good and what is not. [SEL1, int]

Well, it means that some other peer could repeat what you did and reach the same conclusion. I think the notion of peer review and...uh...in a very general sense, is a key part of science. Like if you are the only person in the world to get this result and no other peer can seem to figure out what is happening, that is probably not science. There are also truths probably, religious and other truths that are not...you know...they are real for people and important part of human inquiry but not necessarily verifiable...There is no way of measuring, no way of quantifying whatever it is...You can quantify it. You can't verify it. You can't have a peer look at it and come to the same conclusions. It is not science. So the fact that it is verifiable is a key part of it for me. It doesn't mean there aren't other truths. It is just that these are truths were you can go out and put error bars on and see how you are doing and someone else can too. [eDF2, int]

One of the scientists distinguished between being a scientist and being in the profession of science. Peer review is part of the profession.
So that basically is what science really is. You get deluged with a bunch of facts, you have to sit back, you don’t worry about laws, theories, or principles, or anything. You day dream. You say, well...as I go through I sometimes find two different papers in the messy office and say, “wait a minute. Here is an interesting connection.” That is largely what you doing as a scientist. So what you are talking about with theories and principles and hypotheses, is how I can convince my colleagues that I have thought through this well and it might be a physically meaningful principle. So basically what it comes down to is that science is a way of simplifying and expressing patterns and a way of testing whether or not those patterns make predictions in a way that other people can apply it and understand it. So that is what the real divergence is between how you do science versus what being in the profession of science is. Doing science is playing. It is a lot of fun. The profession of it is convincing other people that you’ve really done something, that it is not an artifact. That is important and they ought to pay attention to what you’ve done even though you are not paying attention to what they have done. [OEL1, int]

Address alternatives. Five participants included a need to address alternatives in the process of justification. Without acknowledging and discrediting possible alternative explanations for the data, the conclusions remain suspect.

I think it is incumbent upon scientists when they come up with something ....to be serious about distinguishing it from prevailing theories. And to talk about the critical test that will ultimately distinguish the validity of the two theories. I think that is what distinguishes a lot of good science from bad science. Maybe I’m looking at it from too much of the perspective of an experimentalist, but ultimately I think it has to come down to things that can be measured in the laboratory. That’s what science is about. [kEL4, int]

[Justification requires] a conclusive experiment. An experiment in which every other possibility has been eliminated by controls. That would be a conclusive study. If the hypothesis was reasonable, didn’t violate any physical laws. The tests were reasonable, all well controlled, all consistent without any anomalies, then that would be acceptable as strong evidence for a particular model. [UEL1, int]

Discipline-based Comparisons

There are several interesting features that emerge from the comparisons by discipline groups. First, the LS and Ch groups were the only participants who
indicated justification through statistical measures (70% and 75%, versus none for the ESS and Ph). Two of the 3 within the “external reproducibility” subcode were ESS participants, whereas there are no LS or Ph. The life scientists (LS) comprise 7 of the 10 participants with the view of “experimentation over description” for justification of claims. Thus, 70% of the LS hold this view, in comparison to 41% of the total sample.

**Approach-based Comparisons**

When comparing the “internal:statistics” subcode by approach, there is a greater tendency for the E and E/D to hold this view of justification (9 of the 11 total; 60% of both groups). The D group only had one participant express this view, and there were no theoreticians who expressed statistical measures were a means of justifying their work. The E and E/D groups fell higher within the “experiments over description” subcodes (8 out of 10 total). Six of these eight are LS. Thus, there may be a tendency for the LS who are experimentalists or engage in a combination of experimental and descriptive research to hold the view that conclusions drawn through experimentation are more justifiable than those based on descriptive investigations. Regarding peer review, there was no apparent trend based on discipline groups or approach groups. However, only one of the D scientists mentioned the important role of peer review in determining justified knowledge. In comparison to the E and D groups or the total sample, more responses of the T and the E/D scientists indicated views that justification “differs with discipline/context,” with 100% of the T and 80% of the E/D scientists represented in this subgroup.
Table 30. Scientists' Views of Data and Evidence: Grouped by Discipline

<table>
<thead>
<tr>
<th>Aspect</th>
<th>Subcode</th>
<th>Total</th>
<th>Grouped by Discipline</th>
<th>Number of scientists</th>
<th>Percentage of group</th>
</tr>
</thead>
<tbody>
<tr>
<td>Data/evidence</td>
<td></td>
<td>24</td>
<td></td>
<td>10 5 5 4</td>
<td></td>
</tr>
<tr>
<td>Different</td>
<td></td>
<td>1</td>
<td>4.2</td>
<td>0 0 1 0</td>
<td>0 0 20 0</td>
</tr>
<tr>
<td>Data: measurements/observations</td>
<td></td>
<td>15</td>
<td>62.5</td>
<td>6 4 3 2</td>
<td>60 80 60 50</td>
</tr>
<tr>
<td>Evidence: interpreted data in relation to a question</td>
<td></td>
<td>10</td>
<td>41.7</td>
<td>4 3 1 2</td>
<td>40 60 20 50</td>
</tr>
<tr>
<td>Same</td>
<td></td>
<td>2</td>
<td>8.3</td>
<td>2 0 0 0</td>
<td>20 0 0 0</td>
</tr>
<tr>
<td>No evidence in field</td>
<td></td>
<td>2</td>
<td>8.3</td>
<td>1 0 1 0</td>
<td>10 0 20 0</td>
</tr>
<tr>
<td>No response</td>
<td></td>
<td>5</td>
<td>20.8</td>
<td>2 1 1 1</td>
<td>20 20 20 25</td>
</tr>
<tr>
<td>Other</td>
<td></td>
<td>3</td>
<td>12.5</td>
<td>1 1 1 0</td>
<td>10 20 20 0</td>
</tr>
</tbody>
</table>
Table 31. Scientists' Views of Data and Evidence: Grouped by Approach

<table>
<thead>
<tr>
<th>Aspect</th>
<th>Subcode</th>
<th>Total #</th>
<th>Total %</th>
<th>Grouped by Approach</th>
<th>Number of scientists</th>
<th>Percentage of group</th>
</tr>
</thead>
<tbody>
<tr>
<td>Data/evidence</td>
<td>different</td>
<td>24</td>
<td></td>
<td></td>
<td>10 5 5 4</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>E  E/D D T</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>E</td>
<td>0 0 0 1</td>
<td>0 0 0 25</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>E/D</td>
<td>5 4 4 2</td>
<td>50 80 80 50</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>D</td>
<td>10 41.7</td>
<td>30 60 60 25</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>T</td>
<td>2 8.3</td>
<td>10 20 0 0</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>2 8.3</td>
<td>20 0 0 0</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>5 20.8</td>
<td>30 0 20 25</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>3 12.5</td>
<td>0 20 20 25</td>
</tr>
</tbody>
</table>

E: experimental; E/D: combination of experimental and descriptive; D: descriptive (non-experimental); T: theoretical
Data and Evidence

Participants' views of the distinction between data and evidence were elicited through a direct question on the VOSI-Sci. Tables 30 and 31 present the results for this category.

Scientists' Views

Participants were asked what data is and if there was a difference between data and evidence. Fifteen participants (62.5%) described data as measurements or observations. The other nine either did not respond to that question (5), indicated no difference between data and evidence (2), or only discussed evidence (2).

*Data as observation/Evidence as interpretation.* Only 41.7% (10 of the 24 scientists) described evidence as the product of interpretation of data in relation to a question or theory. The message was common, but the articulation of views varied from simple to detailed.

Data are raw observations ("facts"). Size measurements, geographic distributions, gene sequences. Evidence usually requires an inference from the data. It is data interpreted in a particular context...[KEDF1, vosi]

Data is different from evidence, in that data generally is regarded as raw information; purely factual; the core thing that an independent investigator should be able to generate to document the reproducibility of an experiment. “Evidence” implies use of the data to support or falsify an hypothesis, i.e., data does not become evidence until it is used for a specific purpose. The same data may be used as evidence for any of a number of competing hypotheses. [gEL3, vosi]

Data is simply anything you measure. Numbers, you can measure or simulate, output, might be data. They are just numbers. Let me see if I can remember the line ...you can say there is data, from data you can get information, from information you can get understanding, from understanding you can get knowledge, from knowledge you can get wisdom...When you use the word evidence, you are implying that already something is being judged. It is like a
theory and you have set up a framework in order to test the theory and you are using the data in that particular context as a test. So it is evidence of a theory. But not all data is the evidence and it is only...you’ve already gone up the chain. If you are at that, you are looking for more than knowledge. Data is always knowledge but it is not always wisdom. It is not always going to teach you anything or understanding unless it is within the theory. So in order to convert the data to understanding you have to test it against a theory and then it can be evidence. [ITC4, int]

The following responses indicate evidence is linked to the success of a model.

[Data] Quantifiable measures of phenomena and their properties. Some examples would be measurements of temperature, wind speed, energy, or work. Evidence has a different connotation to me. Evidence includes data. However, I would regard a successful numerical simulation of a cloud with a simplified physical model as evidence for the importance of the physical processes included in the simplified model, and the relative unimportance of the neglected physical processes. [eDF2, vosi]

Data is the results of tests or observations. Data can be the outcome of laboratory experiments, the observations of geological or astronomical structures, or the outcome of computer modeling. In my own work the outcome of computer modeling lays the role of data. Data is a form of evidence; there are other forms of evidence, such as the mathematical consistency mentioned above. In many/most branches of science "real" data (laboratory experiments and observations) are used, not the outcome of computer modeling. The use of scientific evidence varies enormously from field to field. [pTC4, vosi]

*No difference between data and evidence.* Two scientists explicitly stated that there was no difference between data and evidence.

S: [Data are] Facts or figures used to make calculations or testing hypothesis/theories

[Is "data" the same or different from "evidence"? Explain and provide an example from your own work.]

S: Same. I don’t see what is to be explained by example here. [MEDF1, vosi]

*No use of evidence in field of research.* Two scientists reported not using evidence in their field of research. They suggested their work does not involve interpretation.

I take data to mean the raw or processed results from experiments, such as the counting rate from radioactive sample. I think evidence involves conclusions
drawn from analyzing data. I don’t recall ever applying the term “evidence” in my work. I don’t think that my field of research uses “evidence.” We take data and process the data to draw conclusions. [kEL4, vosi]

S: I guess some folks might use them interchangeably.
R: How do you use them in your work?
S: I would usually talk about the data that I collect. I don’t think that I would ever refer to something as evidence. [NEFL1, int]

**Discipline-based and Approach-based Comparisons**

Compared to the other discipline and approach groups, the theoretical physicists were underrepresented in the “evidence: interpreted data” subcode. Only one of these scientists discussed evidence in terms of data interpreted within a specific context of inquiry. There are no other suggested trends among the groups.

**Reproducibility**

Even though the concept of reproducibility emerged within the “experiment” and “justification” categories, the theme of reproducibility was emphasized consistently throughout the data to the extent that it warranted separation into a main category. This separation allows the collective examination of statements from several other categories as well as additional ideas that emerged related to reproducibility. Results of participants views are presented in Tables 32 and 33.

**Scientists’ Views**

*Statistics.* As found in the justification category, there is a connection of statistics with reproducibility (33.3% of the sample). Similar statements as those included in the “internal: statistics” subcode for justification were also coded within this subcode.
Table 32. Scientists' Views of Reproducibility: Grouped by **Discipline**

<table>
<thead>
<tr>
<th>Aspect</th>
<th>Subcode</th>
<th>Total</th>
<th>Total</th>
<th>Number of scientists</th>
<th>Percentage of group</th>
</tr>
</thead>
<tbody>
<tr>
<td>Reproducibility</td>
<td>statistics</td>
<td>8</td>
<td>33.3</td>
<td>10</td>
<td>50</td>
</tr>
<tr>
<td></td>
<td>requirement to be scientific/acceptable</td>
<td>9</td>
<td>37.5</td>
<td>5</td>
<td>50</td>
</tr>
<tr>
<td></td>
<td>depends on method/discipline</td>
<td>9</td>
<td>33.3</td>
<td>2</td>
<td>30</td>
</tr>
<tr>
<td></td>
<td>no mention</td>
<td>5</td>
<td>20.8</td>
<td>1</td>
<td>30</td>
</tr>
<tr>
<td></td>
<td>other</td>
<td>2</td>
<td>8.3</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

Table 33. Scientists' Views of Reproducibility: Grouped by **Approach**

<table>
<thead>
<tr>
<th>Aspect</th>
<th>Subcode</th>
<th>Total</th>
<th>Total</th>
<th>Number of scientists</th>
<th>Percentage of group</th>
</tr>
</thead>
<tbody>
<tr>
<td>Reproducibility</td>
<td>statistics</td>
<td>8</td>
<td>33.3</td>
<td>10</td>
<td>50</td>
</tr>
<tr>
<td></td>
<td>requirement to be scientific/acceptable</td>
<td>9</td>
<td>37.5</td>
<td>5</td>
<td>50</td>
</tr>
<tr>
<td></td>
<td>depends on method/discipline</td>
<td>9</td>
<td>33.3</td>
<td>2</td>
<td>30</td>
</tr>
<tr>
<td></td>
<td>no mention</td>
<td>5</td>
<td>20.8</td>
<td>1</td>
<td>30</td>
</tr>
<tr>
<td></td>
<td>other</td>
<td>2</td>
<td>8.3</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

E: experimental; E/D: combination of experimental and descriptive; D: descriptive (non-experimental); T: theoretical
Requirement to be scientific. The scientists are not in agreement with regard to the requirement of reproducibility to be scientific. Nine participants (37.5%) affirm a requirement and nine said the requirement for reproducibility depends on the context of the investigation. The remaining did not comment either way regarding requirements.

The VNOS item #2, asking for a comparison between science and art, intended to elicit participants’ views of the empirical basis and creativity in science. Rather, many responses distinguished science and art on the basis of reproducibility.

Statements aligned with the “required” subcode included:

The community of scientists is very important for evaluating work and weeding out incorrect ideas or flawed studies. Presumably criticism in art plays a similar role although the evaluations of critics are opinions. The evaluations of scientists while also partly opinion have the criteria that a work or result must be repeatable by others. Thus, science differs from art in the criterion of repeatability. [PEDF1, vnos]

The motivation for science and art are pretty much the same. They reflect a desire to understand and interpret the world, and are driven by the same internal dynamic. The difference is that scientific observations strive to be repeatable, and consistent with the observations of everyone, whereas art strives not to be repeatable, to be unique. If a scientist observes a phenomenon that no one else can see, it is generally dismissed, whereas an artist with a unique vision is considered a great success. [KEDF1, vnos,2]

Depends on method/discipline. Nine participants stated that reproducibility differed by the discipline or context of the investigation. In general, with increasing variability of the system under study, requirements for reproducibility decrease. These statements directly relate to scientists’ ideas of experimentation and justification.

A lot of sciences like chemistry or physics require repeatable experiments. We [Earth scientists] don’t get to do a lot of that. We do some, but not a lot. Being able to manipulate...you can’t manipulate a big river. You can do it at
a dam. We talk about things as experiments; we take a dam off a river and see what happens. But it is not really an experiment. We don’t have a control. We have a control in what the river looks like beforehand and afterwards, but we don’t have two rivers, one which you take the dam out and one you don’t. One of the conventions in the experimental sciences is there is a lot more emphasis on doing reproducible experiments as the basis by which claims are made. We can do some of it...I think experiments are limited, microcosms of nature. They are not nature themselves. But they allow you to make important... in part because they in some ways get away from some of the subjectivity which I think is inherent if you don’t do experiments. [GFD2, int]

Reproducibility of observations in astronomy is generally neither controlled nor reproducible (at least by the experimenter). Yet these results are certainly good science. [kEL4, vosi]

An experiment to me has to be repeatable. It has to be controlled conditions. We almost never, in my area, we almost never make an experiment. We can’t. The atmosphere is too large. You can’t simulate the effects of the rotation of the Earth in a laboratory. It is just not possible to get realistic results. That doesn’t mean we don’t do science. We just don’t do experiments. We have field programs. We have observational programs. We go to the field. We collect data. We try to simulate using physical laws and the computer and we try to predict what will happen in the future based on our understanding of the present. That is how we make progress. You can’t repeat... I mean... we went to the field to get this data I am working on a couple of years ago. If we went back to the field today at the same time of year and we wouldn’t find the same thing. It’s not repeatable. It’s not an experiment. [eDF2, int]

The one respect where science does not cleanse itself, and this is a problem, is that the scientific community is now so big that people can now get tenure on the basis of what looks like fairly good science, but it is science that nobody else repeats because it is not quite important enough. I am afraid that this is probably more true in biology than other areas of science, but it happens in all areas. There is the possibility of lower quality science slipping through because the community is so large that people are more interested in making new observations than checking whether someone got the right answer. [sTC4, int]

Differences in reproducibility because of the context were also related to level of certainty of the resultant knowledge.

[VOSI #5] Like an experiment, but really only an observation of nature. In a different category because it is not repeatable.... It is still scientific but it is
uh...how shall I say...not likely to lead to the same level of certainty that you get with a repeatable experiment.

By the way, the same thing happens in biology when people say, "wow, I went away and looked at something and this agrees with a theory." Like when Darwin went to Galapagos and made observations and so forth which helped him develop his views of natural selection. That is not really...that is a way of checking a hypothesis to see if it is compatible with the facts. It is less compelling because there is always the possibility because somebody else could come along and say they have this other theory that fits these observations. Now, of course, you could always do that in science. You could do that in areas of high certainty, but when you have repeatable experiments, an experiment where you can design all aspects of it, there is a greater likelihood of deciding among competing hypotheses than when you are looking at nature the way biologists, Earth scientists, and astronomers do. They just have to rely on what is there. You can't play God. You don't have control. [sTC4, int]

**Discipline-based Comparisons**

Most notable here is that only the LS and the Ch affirmed a role for statistics in association with reproducibility (50% and 75%, respectively), with the ESS and Ph scientists totally absent. Moreover, the Ph participants were the only group not to say reproducibility was a required criterion of science. About half of the other groups said reproducibility was a criterion of science. The ESS and Ph had greater tendency to state there were differences in the criterion of reproducibility based on the context (60% for both groups versus 37.5% total sample), and none of the Ch voiced this view.

**Approach-based Comparisons**

Among the eight who stated reproducibility involved statistical measures, seven were either experimentalists or E/D combination. None of the T scientists voiced this view or considered reproducibility a requirement within science.
Regarding "depends on method/discipline," the E had less tendency to express this view (20% of the group, 2 of the 9 total), and the E/D scientists had a greater tendency to express this view (60% of the group, 3 of the 9 total). The E/D group also showed greater tendency toward "requirement to be scientific" than the total group (60% of the group versus 37.5% total sample).

Prediction

Prediction is an aspect that connects with several other aspects including justification and reproducibility. In this regard, prediction was most frequently associated with measures of validity. That is, validity of claims as well as validity of the science discipline itself were related to predictive ability. VNOS items #1, #2 (science versus art), #6 (models), and VOSI items #5, #6 and #8 were most successful in yielding responses related to views of prediction. Some representative statements have already been used in support of views of models, methods, anomaly, justification, and reproducibility. Results of the scientists' views of prediction are presented in Tables 34 and 35.

Scientists' Views

*Affirmed important.* The notion of prediction as an important component within science and scientific inquiry emerged through responses of 87.5% (21) of the scientists. Like models and reproducibility, there was some variability in how much emphasis scientists placed on prediction. Only three participants made no mention of prediction. Typical responses indicating the importance of prediction include:
Table 34. Scientists' Views of Prediction: Grouped by Discipline

<table>
<thead>
<tr>
<th>Aspect</th>
<th>Subcode</th>
<th>Total</th>
<th>Total</th>
<th>LS</th>
<th>ESS</th>
<th>Ph</th>
<th>Ch</th>
<th>Percentage of group</th>
</tr>
</thead>
<tbody>
<tr>
<td>Prediction</td>
<td>affirmed important</td>
<td>21</td>
<td>87.5</td>
<td>5</td>
<td>2</td>
<td>4</td>
<td>2</td>
<td>50</td>
</tr>
<tr>
<td></td>
<td>requirement to be scientific</td>
<td>6</td>
<td>25.0</td>
<td>4</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>40</td>
</tr>
<tr>
<td></td>
<td>through models</td>
<td>10</td>
<td>41.7</td>
<td>3</td>
<td>5</td>
<td>0</td>
<td>2</td>
<td>30</td>
</tr>
<tr>
<td></td>
<td>through conclusions</td>
<td>2</td>
<td>8.3</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>Testable</td>
<td>2</td>
<td>8.3</td>
<td>0</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>20</td>
</tr>
<tr>
<td></td>
<td>no mention</td>
<td>3</td>
<td>12.5</td>
<td>3</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>10</td>
</tr>
<tr>
<td></td>
<td>differs with discipline</td>
<td>3</td>
<td>12.5</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>10</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Number of scientists</th>
<th>LS</th>
<th>ESS</th>
<th>Ph</th>
<th>Ch</th>
</tr>
</thead>
<tbody>
<tr>
<td>10</td>
<td>5</td>
<td>2</td>
<td>4</td>
<td>2</td>
</tr>
<tr>
<td>5</td>
<td>4</td>
<td>40</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>5</td>
<td>3</td>
<td>3</td>
<td>0</td>
<td>2</td>
</tr>
<tr>
<td>4</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>3</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>2</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Percentage of group</th>
<th>LS</th>
<th>ESS</th>
<th>Ph</th>
<th>Ch</th>
</tr>
</thead>
<tbody>
<tr>
<td>10</td>
<td>50</td>
<td>40</td>
<td>80</td>
<td>50</td>
</tr>
<tr>
<td>5</td>
<td>40</td>
<td>40</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>3</td>
<td>30</td>
<td>100</td>
<td>0</td>
<td>50</td>
</tr>
<tr>
<td>2</td>
<td>0</td>
<td>20</td>
<td>20</td>
<td>0</td>
</tr>
<tr>
<td>3</td>
<td>20</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>3</td>
<td>30</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>2</td>
<td>10</td>
<td>20</td>
<td>20</td>
<td>0</td>
</tr>
</tbody>
</table>
Table 35. Scientists' Views of Prediction: Grouped by Approach

<table>
<thead>
<tr>
<th>Aspect</th>
<th>Subcode</th>
<th>Total #</th>
<th>Total %</th>
<th>Grouped by Approach</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>E</td>
</tr>
<tr>
<td>Prediction</td>
<td>affirmed important</td>
<td>2</td>
<td>87.5</td>
<td>6</td>
</tr>
<tr>
<td></td>
<td>requirement to be scientific</td>
<td>6</td>
<td>25.0</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>through models</td>
<td>10</td>
<td>41.7</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>through conclusions</td>
<td>2</td>
<td>8.3</td>
<td>0</td>
</tr>
<tr>
<td>testable</td>
<td></td>
<td>2</td>
<td>8.3</td>
<td>1</td>
</tr>
<tr>
<td>no mention</td>
<td></td>
<td>3</td>
<td>12.5</td>
<td>2</td>
</tr>
<tr>
<td>differs with discipline</td>
<td></td>
<td>3</td>
<td>12.5</td>
<td>0</td>
</tr>
</tbody>
</table>

E: experimental; E/D: combination of experimental and descriptive; D: descriptive (non-experimental); T: theoretical
Science is the pursuit of understanding that allows one to predict how things work. It differentiates itself from religion and philosophy in that it focuses on repeatable phenomena. One can design tests to learn what works and what doesn’t. Science, as I pursue it, also incorporates an experience-based library of knowledge and mathematical rigor which are wholly lacking in disciplines such as religion and philosophy. The understanding also leads to predictive capability, as in orbital mechanics. One can predict trajectories with good accuracy under conditions for which there’s been no previous experience, as in orbiting a spacecraft around a plantlet, and have a high degree of confidence that the predicted outcome will be realized. [cDFC2, vsos]

R: How do you make the choices of which models to use?
S: It is the usual thing of how you do science. You do science by taking a body of data and analyzing it and making a model or theory that fits the data. Now this model or theory has some sort of predictive value. It can predict some particular outcome. We fit all the data we have for particular parameters. If there is something else that these particular parameters can predict, we can then go into the laboratory and measure. That is the way science works in different fields: analyze and predict; analyze and predict. [KEL4, int]

**Requirement to be scientific.** Six of the scientists strongly voiced their views of prediction as a requirement of science.

[In reference to VOSI #5] The conclusion is speculation until it can be used to predict. Gathering data, determining correlations, and characterizing beak shapes and nut types are all tools of scientific inquiry. It is not science if the person is not attempting to determine verifiable scientific truths. Stamp collectors gather samples and data, and organize them in systematic ways. But their activity is not considered scientific research. [eDF2, vsos]

Science is the pursuit of understanding that allows one to predict how things work. It differentiates itself from religion and philosophy in that it focuses on repeatable phenomena. One can design tests to learn what works and what doesn’t. Science, as I pursue it, also incorporates an experience-based library of knowledge and mathematical rigor which are wholly lacking in disciplines such as religion and philosophy. The understanding also leads to predictive capability, as in orbital mechanics. One can predict trajectories with good accuracy under conditions for which there’s been no previous experience, as in orbiting a spacecraft around a plantlet, and have a high degree of confidence that the predicted outcome will be realized. [cDFC2, vsos]
The following excerpt from the interview with a molecular biologist demonstrates his views and connections of his views of prediction. He considers validity of claims and differences in disciplines as they relate to predictability. The entire excerpt is included here to reflect the interview prompts and thought process the scientist underwent as he made connections and reinforced his view of the importance of prediction. This portion of the interview follows his description of experiments and the types of investigations he uses in his biomedical research.

S: Most of what I do is try to understand cause and effect relationships. What causes the nitric tyrosine in animals...can I give them a drug to block it and does it now have an effect on the disease process? Eventually I hope to test it in humans, but that is no longer an experiment. It is a clinical trial and I can’t control the variables. I can see if there is an effect on the outcome. In a way it is, but right now the evidence that this process relates to humans is based on observations in humans. I can reproduce it in animals. I can reproduce it in cells.

R: Ok, does the different kind of evidence then that you get in these types of procedures influence the validity of the findings?

S: I don’t know how to answer that. It is such a broad question...but sure. An observation just gives you something to look into, this might be the effect.

R: You could think of another example. Let’s go back to comparing fields of science...field ecology or astronomy or something else where you can’t do manipulative experiments, or something else in molecular biology. Can you compare the validity of the claims made in those fields?

B: You know I don’t know how to answer that because it is so broad in so many different ways. The answer is in astronomy for example, yes you make observations, but can you predict them from underlying theory. And so what it really comes down to is that you can do experiments in astronomy where you make predictions and you search for a particular type of phenomenon and that gives you some more certainty that what you predict will happen. It is useful, you’ve got a set of rules...things that will happen. You ask, does it happen and can you find examples. Then you have to use statistics to decide whether or not it happened by chance or is this more consistent with there being a relationship. Are relationships driving it or is it pure chance. That’s fine. So the validity, again, you never prove validity. You invalidate things. I guess that is where I am having trouble, struggling with this. How would you invalidate certain studies? Is there an experiment you can do to disprove something? If there is not an experiment you can do, then what you have is tautology, and it is not terribly useful. There has to be some defining
experiment that says this is wrong, this is a lower bound of how this would work.

R: Can you think of an example of where there is not?
S: Where you can't disprove something?
R: Yeah.
S: Well, for example, string theory. Bottom line is that it potentially has the ability to explain, be the theory of everything. But so far no one can come up with a single experiment to test whether it is real or if the generalized appeal is real. So you can explain the facts that are here, but can you explain something...a new observation? Do you make a prediction from this theory that no one has thought of before that you can test? Or can you do an experiment where you can actually isolate a string and the calculations indicate the accelerator would have to be larger than the known universe, which is beyond the realm of testing. So is it worth doing string theory? Sure, it is a very hot field, but the bottom line is that if it is going to hold up someone has to figure out a way to test it and make predictions you can't get from other theories.

R: So is it science?
S: It is science in that it is a developmental tool I guess. You have an underlying theory and framework that you are expanding, and you are explaining facts. The problem with string theory is that it hasn't predicted anything in a way that you couldn't explain by some other phenomenon. So it is not a theory of everything at this point. It is a theory that explains known facts. It is just it hasn’t explained, made a prediction you could test and other theories would not have worked. So, is it science? That’s probably putting too harsh a judgment on it because certainly you need to develop theories and give them time to evolve and see where they are going to go.

R: How does that differ from mathematics?
B: It is going to be really hard to distinguish the two. But basically mathematics is the underpinnings for the other sciences because basically it allows you to make predictions of physical phenomena and model it. The difference with mathematics is that you shouldn’t try to constrain yourself to try to explain physical phenomena. Trying to work through a set of postulates, things you can derive from it. Largely you are making conjectures. Conjectures in a way are hypothesis. You play with something for a while and it seems to make sense, and then you have to figure out from an accepted set of postulates how to prove that conjecture being true. That is how it has taken up to 400 years to prove some conjectures. So I think the difference between mathematics and science is that mathematics does not have to have a physical reality or prediction that you would test. [OEL1, int]

**Prediction through model use.** Ten (41.7%) of the scientists suggested prediction is attained through model use. With regard to models, the two atmospheric
scientists and the aquatic ecologist strongly emphasized models in science as well as
described model use in their own research. Interestingly, differences in views of
other scientific disciplines were voiced in relation to a discipline's ability to predict
based on models. The atmospheric scientist [cDFC2] quoted above as well as earlier
in reference to the "Methods" category held views against ecology as a valid science
because he saw a lack of predictive ability. Others expressed views of attaining
prediction through models.

[Discussing function of models in her work] Then you can test it and try a
different set of conditions. If they do, then it means the model is working, at
least for these conditions, and it has some predictive function. One is to test
the input to see if I have my ideas straight and the other is to make
predictions. [fEF3, int]

They compute the basic underlying models. I take those models and they
usually compute them for specific temperatures and specific gravities, and I
take those models and I interpolate within grids of models to get models that
apply to my specific stars. Then I use those models to predict what the
spectrum would look like given a certain composition. [pEDFC2, int]

**Diffs with discipline/context.** Three participants stated predictive ability
differed with discipline or context. These views are consistent with differences of
reproducibility and justification.

I often use the example in my class. We talk about particle physics when we
teach about modern physics. It is a little bit outside my area. It is kind of fun
for students to think about. For many years particle physics was a lot like
butterfly collecting. You catch a lot of butterflies and start organizing them in
sequences of size and color, by different body parts, whatever. Early days of
biology. The hairless things over here. The crawly things over here. The
flying feathery things over here. We began to notice some common
properties. That lead us to some next step in analyzing structures. In particle
physics, it was a lot like that. There was no prevailing theory. Then in 1960s
and 1970s there was a theoretical framework that evolved that explained
these different classifications. It was similar to butterfly collecting. You
would go into the lab and smash things together and look at the particles that
come out, classify them into different categories by mass or electric charge or
whatever. Looking at this huge collection of stuff and try to make some sense out of it. There were themes emerging....So now we have a theory that explains all the structure that now has some predictive value in the lab. We test some things and jiggle the grounders of the theory. The theory has evolved quite a lot in the last 40 years...That is an example in physics of something that started out with no predictive theory and gradually moved into an area of predictive science. Astronomy works in a similar way. You look at galaxies .....Now we have a cosmological theory that explains the origin of galaxies. [kEL4, int]

**Discipline-based Comparisons**

An interesting feature based on a cross-discipline comparison of views about prediction is related to models. All of the ESS participants emphasized models as a means through which predictions are tested.

**Approach-based Comparisons**

Within this “through models” subcode, seven of the 10 in the total sample utilized either a combination of E/D approaches or D approach (that is 70% of the E/D and D groups combined, as compared to 40% of the E and E/D groups combined). None of the Ph or T participants appeared in the subcodes of “requirement to be scientific” or “prediction through models,” although most of them indicated prediction plays an important role in science. These results suggest the Earth and space scientists, who also conduct at least some not-experimental, empirically-based research (E/D or D), showed stronger tendencies than the other groups to emphasize a connection between prediction and models.

**A final note about prediction**

Limitations in predictive ability has provided as a source of controversy among scientists. This controversy is related to the complexities of systems under
study and perceived images of what constitutes valid science. In addition to the
previously described views toward ecology, one of the scientists discussed the
progress his field has made in simply acquiring status as a science:

Meteorology has had a long struggle to be recognized as a serious scientific
discipline. The fact that scientific meteorological predictions can only be
expressed in terms of probabilities, and that the ranges of uncertainty are
large, has created the incorrect impression among the general public that
meteorology is not a quantitative scientific discipline with problems that are
as challenging as those found in nuclear physics or molecular biology. On the
other hand, the real needs of agriculture, commerce and defense have
provided a modest but continuous flow of research dollars to research is a
necessity. [eDF2, vnos]

He elaborated during the interview:

From the scientific perspective, I think early on it [meteorology] wasn't a
very prestigious science, like nuclear physics or some areas of chemistry, just
those basic disciplines. The scientists in those areas did not respect
meteorology.
R: and why so?
S: Well, partly because it is applied. The laws are pretty well known. There
are Newton's laws of motion. There is the ideal gas law. There is the first law
of thermodynamics...We can write these equations down. Everybody can
write them down. The problem is how do you write them up and make a
forecast [laughs]. I am sure that some nuclear physicists who were
contemplating the origins of the universe felt that problem was like an
engineering problem. Here is point A and here is point B. Here are the
equations. Why can't you do it? Sort of a grunt work assignment. You are not
really a scientist, more like engineering. I think that was a problem to get
respect in other scientific disciplines. I think over time some of the best work
in turbulence and fluid dynamics has been done by atmospheric scientists and
oceanographers. It has gradually gotten acceptance. The equations are just as
difficult...it is not like it is chopped liver. There are real scientific problems
that have interesting characteristics. I think gradually we have become
quantitative. In the beginning you were just agricultural weather forecasters.
They got lousy pay. No smart young person, unless they were a total weather
nut would ever aspire to such a position. They wouldn't go to the best schools
and get the level of mathematics and physics they needed to ...[eDF2, int]
This study examined scientists' epistemological views of science and sought relationships between views and authentic research contexts. This chapter discusses the results with respect to both research questions. Results of the main categories are discussed to suggest explanations and, when possible, to compare the findings with the literature. When not discussed, comparable research is lacking. Following this discussion, overall conclusions are presented, along with limitations of the study, implications for science education, and recommendations for further research.

Scientists' Epistemological Views of Science

The results show that these participants' epistemological views of science are complex, yet suggest overarching consistencies. There are consistencies at general levels of description and applicability with respect to particular features of NOS and NOSI. That is, the main categories of NOS and NOSI are applicable across all the contexts of these scientists, and many of their broad descriptions were similar. For example, the general description of "experiment" was quite similar for most of these scientists. The more subtle complexities emerge within descriptions, those being represented by the multiple subcodes, as the aspects relate to individual experiences and specific contexts.

Through written and verbal responses, the scientists showed conviction in their views, supported with examples from their research. The responses demonstrate
connections between individual authentic scientific inquiry experiences and these scientists’ views of NOS and NOSI. However, their views are not necessarily consistent with any particular philosophical position. It is important to note that the present study is not attempting to align these scientists with any particular philosophical stance (e.g. relativist, absolutist, etc.). On occasion these perspectives are useful to describe how perspectives in this study may be described and comparable.

Given the complexity of this study and results, each of the main categories is presented here for discussion. Results of the whole sample and group comparisons are discussed in order to address possible explanatory relationships. Furthermore, noticeable connections among categories and groupings are highlighted.

**Tentativeness**

Only 46% of the sample demonstrated views that scientific knowledge is inherently tentative. These, combined with the six additional participants who reported quite sophisticated examples of different levels of certainty, comprise 17 of the 24 scientists in the sample. The notion of differences in certainty due to complexity of the system involved had not been depicted in previous studies with scientists, teachers, or students. The remaining seven scientists indicated the knowledge either progressively approaches certainty or reaches certainty. That some of the participants had absolutist views of scientific knowledge is consistent with other reports (e.g. Behnke, 1961; Glasson & Bentley, 2000) while inconsistent with predictions that scientists hold views of open-minded realism (Harding & Hare, 2000). Many of the scientists in this study reported using “what works” with the
understanding that "what works" might change or might not be an exact
representation of the real phenomenon. In this way, they demonstrated what Harding
and Hare (2000) considered open-mindedness. However, except for the nine
participants who explicitly mentioned science as knowledge of reality or
approaching knowledge of reality, these scientists were seemingly comfortable with
empirically-supported relativist positions, seeing knowledge as "what works" as
opposed to "small scale truths" (Harding & Hare, 2000). They were consistent and
did not mix meanings in their responses. Overall, the participants in this study
primarily affirmed that scientific knowledge is subject to change, recognized there
are areas of science that are more certain than others, yet some viewed science as
progressing toward knowledge of external reality. Of the five who stated scientists
reach that knowledge, four were life scientists; none were physicists or chemists
(Table 4).

Why might these life scientists, as opposed to the physicists or chemists have
the greater tendency toward absolutist views? These results are contrary to what one
might expect given the "hard science - soft science" continuum (or demarcation in
some cases) that is often used to describe these science disciplines (e.g. Spieker,
1972; Van Bemmelen, 1961). In fact, some of the scientists in this study used these
descriptors, referring to the physical sciences as "hard" and the natural sciences as
"soft." These descriptors indicate the type of data relevant to the field, as well as the
number and controllability of variables (Knorr-Cetina, 2000; Mayr, 1997; Spieker,
1962). With seemingly more variables and less control of the system, one might
expect those in the life sciences and Earth and space sciences to hold more relativist
views of scientific knowledge as compared to the physicists and chemists. Likewise, one might expect those engaging in nonexperimental research to hold relativist views in comparison to the experimentalists. The results of this study do not support these conjectures. The fact that one of the atmospheric scientists demonstrated more absolutist views clearly goes against expectations based on the “hard science/soft science” rationale.

Combining the results of the chemists and physicists for review of this category, the life scientists still have a greater tendency toward absolutist views. These results might be explained by virtue of a difference based on discipline; these results might be a feature of the small sample size; moreover, these results might be a feature of this particular sample. Four of the physicists in this sample were theoreticians (Table 1) and may not hold views typical of physicists in general. Indeed, the theoretical research approach may affect a more tentative view because of the lack of direct empirical basis. In contrast, the life scientists explore living systems through empirical means. In any event, the views voiced by this group, and all the groups, should be reviewed in light of the sample characteristics.

Empirical Basis

The categories of “empirical,” “observation/inference,” “models,” “experiment,” “anomaly,” “data/evidence,” “justification,” “reproducibility,” and “prediction” relate in some way to the role empirical data in the development of scientific knowledge. With respect to the whole sample, the participants were fairly consistent in their views of all these categories except “justification” and “reproducibility.” The differences for these two categories are discussed below.
There was overwhelming agreement among these participants in the importance of empirical data in the development and justification of scientific knowledge. These results are consistent with reports of others (Bell, 2000; Glasson & Bentley, 2000; Osborne et al., 2003).

The theoreticians had a greater tendency than any other group or the whole sample to state that valid knowledge could be developed through mathematical and theoretical means, without extraneous empirical support (Table 7). This result is consistent with their research approach and reported conventions of their research community. They reported that developments in technology and mathematical theory has extended the boundaries of what is considered science, citing String theory as revolutionary in pushing those boundaries. These ideas, although valid from these scientists' perspectives, are beyond the level of practical consideration for the typical K-12 classroom, at least for the near future. These are extremely sophisticated ideas that have not been reported elsewhere in studies involving K-12 science teachers or students (e.g. Khishfe & Abd-El-Khalick, 2002; Lederman, Abd-El-Khalick et al., 2002), and not included among scientists' suggestions for K-12 science topics (Osborne et al., 2003).

Subjectivity

A majority of the scientists acknowledged the influence of current scientific theory and paradigm in directing scientific research. They recognized theory-laden observations and investigations from within their research contexts and were able to provide examples. This view is considered a more sophisticated view in comparison to only recognizing personal subjectivity, such as variances in taking measurements.
Few beginning teachers who participated in a science research internship were able to achieve that deeper level of understanding (Schwartz et al., in press). Perhaps the extended authentic experience, and moreover, true membership within the community (Lave & Wenger, 1991) provides opportunity to not only (1) develop expertise in the theoretical framework driving the research, but also to (2) contribute to that framework through original research, and (3) develop practical utility of that framework for acquiring funding necessary to maintain that research. The 15 scientists who demonstrated this view provided evidence of an almost commonsense notion of theory-ladenness. One of the responses relative to use of scientific models was, “Every scientist uses models, and if they say they don’t, they fail to understand what they are doing.” In other words, use of theory to drive one’s work may be implicit or explicit. The framework is there whether the scientist recognizes it or not. As these results suggest, not all of the scientists recognize it. The fact that 15 of the 24 scientists in this study did explicitly recognize this feature of subjectivity is interesting because of the contrast to other reports (Glasson & Bentley, 2000), and as mentioned, perhaps a function of their prominence and longevity within the scientific community.

The notable feature of the group comparisons for this category is in the subcode of “differs with approach.” The chemists consider there to be more subjectivity in qualitative approaches as compared to quantitative, where they describe quantitative as involving numbers. As discussed below, these results are consistent with their view of the importance of statistical analysis. Again, the
physicists cluster differently, with none of them agreeing with the chemists for this subcode (Table 8).

For the present study, four stated their work and that of other scientists is objective, and two of these mentioned the use of the Scientific Method as a mechanism for eliminating subjectivity. These types of views are more consistent with naïve notions of subjectivity, commonly voiced by learners before explicit NOS instruction (e.g. Abd-El-Khalick & Lederman, 2000; Lederman, Schwartz et al., 2001; Schwartz & Lederman, 2002).

Creativity and Observation/Inference

A majority of the participants affirmed a role of creativity and inference. According to Ziman (1995), pattern recognition is linked to subjectivity and is a mainstay of all scientific practice.

…the bodily senses are the only link between the human mind and the world he or she inhabits. Visual perception, by its intersubjective consensibility, is an essential element in the creation and validation of scientific knowledge, and pattern matching provides a standard of consensuality which is never completely superseded by more ‘objective’ devices such as mechanical instrumentation. (Ziman, 1995, p. 55-56)

For the given data set, there were no consistent responses beyond affirmation of inference, and just a few that expanded on the role of creativity. Thus, even though the results suggest most of these scientists consider creativity and inference to be important to their work, few could actually explicate the use of creativity in making meaning of data, as evidenced by the low frequency of responses within the “identifying patterns or building connections among data” subcode.
The main category of "data/evidence" relates somewhat to the category of "creativity" and "observation and inference." Yet, as Tables 30 and 31 show, the scientists did not elaborate on their views of evidence either. These scientists tended to emphasize the importance of empirical data and supporting conclusions with that data to justify claims to themselves and their peers. The nature of the discussions about data and evidence and means of justifying one's claims was very different from the nature of the discussion about science and art, even though both topics could introduce ideas of creativity in making meaning of data. Creativity was seen as a part of the process of science in general, but not necessarily as a part of developing solid claims specifically. There was a division between doing and justifying. The extent to which these scientists all saw this division is not determined from the current data. However, the comment from one of the molecular biologists on the distinction between doing science and being in the profession of science, in terms of having fun and being creative versus writing up the work to the acceptance of your peers, demonstrates this division. Perhaps some of the participants did not recognize the role of creativity in making meaning of data due to the muddy lines between doing science and being in the profession of science.

Sociocultural Influence

A majority of these scientists emphasized the importance of funding and the influence of political and societal pressures on the direction and continuation of scientific research. The influence was directed toward what questions get asked more so than on the reasoning processes involved with how the science is conducted. In this way, the sociocultural influences were seen as primarily external. The scientists
reported having to tailor their research programs toward the agendas of the funding agencies. Political and societal institutions establish standards and direct research through funding decisions, and are a recognized feature of scientific-social dynamics (Knorr-Cetina, 1999; Ziman, 1995). However, this feature is typically overlooked in the context of science education. Through contacts with scientists in practice, students and teachers have learned about the pressures of acquiring funding, and recognition of such requirements may also lead to recognizing the theory-laden NOS (Ryder et al., 1999; Schwartz et al., in press). As several of these scientists stated, the grant writing process itself mandates work be framed within current theory and directed toward worthy goals that fit within the visions of current scientific progress. However, this vision may or may not fit exactly with the preferred direction of the scientist, as in the case of the astronomer who reported having to change her research focus from older planetary systems to younger systems to satisfy the interests of NASA:

I look at what is going on in the field and I put the emphasis in the grant proposal in a direction that I think will be more fundable than necessarily in a direction that might be easier for me to carry out a research program. I put in a proposal for NASA funding this last summer to get funding for a graduate student to work on this program to detect planets around young stars and look at the evolution of planetary systems. It is a whole lot easier to do that project if we look at clusters that are 100 million years old or older because those stars are more stable and it is easier to make the measurements we need to make. The question that is more likely to get funded is what is happening in stars that are 20 and 30 million years old. So I had to revise my thinking in order to see how I would carry out this observational program to really target that age group. The older clusters are still interesting, but they aren’t from the point of view of NASA which is really focusing on the young systems...I actually changed how I wanted to do the program to address questions for younger systems. [pEDFC2, int]
The Earth and space scientists and the experimentalists held this view slightly more the other groups, with the theoretical physicists not stressing sociocultural pressures (Tables 12 and 13). Whether trends are related to discipline, research approach, individual situations, or combinations of factors is undeterminable.

An interesting feature of a few of the scientists’ views of socio/cultural influences, or lack thereof, is the role of the international community in minimizing differences across cultures. They specifically reported their involvement with international groups through conference attendance, collaborations, and advances in technology lead to a global scientific society. The extent to which this global perspective is shared by less established scientists or by scientists from other cultures would be interesting to pursue. Despite the low frequency of statements indicating a global community, there were still few scientists who identified an internal feature of culturally-based differences in how science is conducted. The more sophisticated philosophical view that scientific reasoning and the processes of developing and accepting scientific knowledge is influenced by the culture and society within which the science is practiced was only voiced by a few of these scientists. This result is not surprising, given again the international status of the scientific community and the high occurrence of external influences. Moreover, this more sophisticated perspective of sociocultural embeddedness seems to be generally difficult for individuals to develop, even through explicit instruction (e.g. Lederman, Schwartz et al., 2002).
Scientific Theories and Laws

These participants demonstrated well articulated and fairly consistent views of scientific theories, but some wavered in describing scientific laws. Over half reported hierarchical views that theories develop into laws with repeated testing and or after sufficient time, but it should be noted that, unlike typical naïve responses, most of these hierarchical views maintained laws as tentative. A typical naïve response states laws are theories that have been proven true through repeated testing (Lederman, Ebd-El-Khalick et al., 2002). None of these scientists used “proven true” to describe the transition from theories to laws. They tended to use the idea of a consistently established theory, historical use of terminology, and even the suggestion of a community vote to mark the use of the word “law” in a hierarchical transition. In this sense their hierarchical view was not typical of teachers’ and students’ naïve views (Abd-El-Khalick & Lederman, 2000; Lederman, Abd-El-Khalick et al., 2002).

Overall, these scientists’ views of theory and law were consistent with their views of tentativeness. Recognizing differences in the application of laws based on discipline is consistent with the view of different levels of certainty. Some of the scientists reported laws to be more certain than theories, and this certainty depends on the discipline under study. Those disciplines that have fewer variables, are more controllable and predictable, are more likely to have established laws, and thus, have more certainty attached to them.

Mayr (1997) described laws in biological sciences as distinct from laws in the physical sciences. Some of these scientists held similar views that laws differ with
scientific disciplines, and even that there are no laws in their field. Despite only six scientists falling into these two subcodes, there are suggested discipline-based distinctions. The life scientists and Earth and space scientists not only saw differences based on discipline, but they also claim there are no laws in their field of research (Table 14). In contrast, the physicists and chemists saw laws similarly across all the sciences. When comparing by approach, only the theoreticians fit with the physicists and the chemists (Table 15). What is intriguing about these results is that there are no notable patterns in how these groups described theories and laws that might help to explain the disparity in applicability. Their descriptions as hierarchical or different do not relate in any way to how they see laws applying across disciplines. Five of the six who stated no laws in their field directly related these views to their views of tentativeness. Again these scientists raised the notion of influences of complexity of the system under study on the ability to attain certainty. This theme is repeated throughout the results.

Models

Creating and using scientific models is central to scientific inquiry (Gilbert, 1991). There was overwhelming sentiment that models are used to explain or organize observations, then predict and test through further observations. The emphasis here is on empirical observation in the development and in the testing of models. In comparison, half as many scientists saw models as a means to visualize something abstract or simplify a complex process. This latter view seems to place less emphasis on direct observation and incorporates theoretical entities, although these are not necessarily mutually exclusive. These results show that these scientists'
perceptions and use of models fit broadly with published descriptions of functional roles of models in science, including descriptive, explanatory, and predictive characterizations (Justi & Gilbert, 2003; Van Driel & Verloop, 1999). The multiple descriptors that the scientists used for models, such as mathematical, physical, and analogical, are also consistent with prior characterizations. In contrast to the range and multiple categories of meaning for the seven aspects of models identified in the Justi and Gilbert (2003) study of teachers’ views of models, the present study suggests these scientists hold more consistent views of scientific models. Justi and Gilbert (2003) suggested a unique definition of model used by scientists, but they did not have empirical support. The present study provides a definition, with supporting data. According to practicing scientists from a variety of specialty areas, a scientific model may be a mathematical, physical, analogical, or mental construct that (1) explains or organizes observations, that then enable prediction and testing through further observations, (2) simplifies a complex phenomenon or renders an abstract concept visible, and (3) provides a framework for guiding further investigation.

The Earth and space scientists and the physicists were clearly disparate in their descriptions of models (Table 18). The approach groups were also distinct, with 100% of the E/D group describing models as “explain or organize observations/predict/test” (Table 19). When looking at the results for the “prediction” category (Table 34), again the Earth and space scientists and physicists are disparate with respect to the requirement of predictive ability and the role models play in achieving prediction. The E/D group falls into a similar pattern as the Earth and space science group (Table 35). These results suggest the Earth and space
scientists and/or those who engage in combination of experimental and descriptive research hold more similar views of scientific models than they do to theoretical physicists or even to the whole sample of scientists. That is, these former groups tended to emphasize models and their explanatory and predictive functions more frequently than the other scientists in this sample.

Experiments

The scientists were asked to provide a definition of scientific experiment on the VOSI-Sci questionnaire because prior studies with teachers and students have demonstrated the concept of “experiment” is understood and used in different ways (Lederman, Schwartz, et al. 2002; Schwartz et al., 2001). The results of the present study suggest that most of the scientists hold the view of scientific experiment as a traditional method of identifying controls and manipulating variables to establish cause/effect relationships (Simon, 2001). Where the groups demonstrate differences in their views, however, is in the perceived importance of (1) hypotheses, and (2) replicas in this experimental process, and (3) the requirement of experiments for the development of scientific knowledge. The results suggest this category has several context-based relationships, but the patterns may not be distinguishable between discipline or approach (Tables 20 and 21). For example, all of the chemists claimed experiments are hypothesis driven, and likewise a large portion of the experimentalists and the E/D scientists. None of the Earth and space scientists put priority on hypothesis testing, nor did the descriptive scientists. There may be interaction effects or overshadowing effects of one group character over the other.
Purpose

Educators are constantly emphasizing the importance of making science relevant to the learner. By connecting science to real-world situations, learning becomes more meaningful (e.g. Minstrell & van Zee, 2000). The results of this study suggest that what is meaningful depends on the scientist. From their perspective, the purpose of scientific inquiry is to generate understanding, with fewer of them focusing on applicability of that knowledge to the betterment of human or natural existence. Sometimes what is relevant to a scientist may be to satisfy curiosity about a phenomenon that may not have an applied function, or even relevance to anything tangible in the real world. Moreover, what is relevant to most of these scientists is influenced by governing agencies that provide research funds. Thus, relevance is relative, and in many instances the applicability of the knowledge produced is irrelevant to the scientist or even unknown at the time. Generating basic understanding predicates applicability. Some of the participants compared knowledge development and knowledge application as separate processes performed by scientists and engineers, respectively.

Methods

The fact that some of the scientists stated views of a single scientific method was alarming at first. However, when they clarified their view, only a couple held the traditional view of scientific method as an experimental approach. Thus, the importance of clarifying terminology through the interview process is highlighted with this example. These scientists were using the term of "scientific method" in a broad way to describe what scientists do, in general, involving asking questions and
collecting information and hypothesis testing. Nonetheless, this view is still narrow with respect to all methods of scientific inquiry because of the criterion of a priori hypotheses. What is disturbing is that any of the scientists hold convictions toward a single Scientific Method in the traditional sense that scientists only engage in a strict step-wise practice of hypothesis-testing and experimentation. According to these narrow views, several participants in this study do not do credible science.

The scientists here who said there are multiple methods, and that not all scientific investigations require hypothesis testing provided examples from their own contexts. Given that these are successful scientists, their multiple methods are examples of scientific research, regardless of the presence of formal hypothesis testing or controlled experimentation. These results further portray authentic scientific inquiry as comprising dynamic, creative, and multiple processes involved in the development of scientific knowledge (e.g. AAAS, 1996; Lehrer, Schauble, & Petrosino, 2001; Knorr-Cetina, 1999; NRC, 2000; Osborne et al., 2003; Simon, 2001; Ziman, 2000). A single method or even a general method of hypothesis testing is inconsistent with most of these scientists’ views and reports of their authentic practice.

About the same number of scientists considered hypotheses to be a requirement as considered hypotheses to be dependent upon the approach (29% versus 33%) (Table 3). As with the findings for “experiment,” the role of hypothesis testing was given opposite priority by the Earth and space scientists and the chemists (Table 24). In this category, the chemists and the life scientists favored hypothesis
testing, and the physicists and the Earth and space scientists were against a requirement of hypothesis testing.

The results based on group comparisons are difficult to interpret in any strict sense because they are supported by the individual’s research context. That suggests the research context for these experienced scientists relates to their perceptions of valid scientific methods. Some are more open than others to methods of investigation and justification outside of their area of expertise. Thirty-seven percent recognized differences in disciplines and considered there to be differences in validity of methods. However, 60% of the E/D combination approach group thought experimental methods are more valid than descriptive approaches. Perhaps their experiences with multiple investigative methods help to clarify distinctions between those methods.

Anomaly

Recognizing anomalies is connected with subjectivity, although only a few participants identified this positive element of subjectivity. Reproducibility is also related to identification of anomaly. According to several of these scientists, if the anomaly is repeatable, it is a true anomaly and, as such, something interesting to explore further. The two mechanisms of addressing a true anomaly identified by these scientists were to “develop a new model” or “expand the existing model.” As opposed to suggesting both or several options, these scientists tended to suggest only one mechanism of handling a true anomaly with respect to their work. The chemists were the most clustered, showing preference for the “develop a new model/theory to explain” subcode. The life scientists and the E/D scientists clustered in the opposite
direction from the chemists, with more preference toward “expand the existing model” (Tables 26 and 27). There is no clear reason for these clusters.

Investigations of how people respond to anomalous data have characterized several reasoning processes distinct from how they deal with expected data (Chinn & Brewer, 1993, 1998; Trickett, Trafton, Schunn, & Harrison, 2001). There were four common characteristics identified in the present study that are similar to what Chinn and Brewer previously identified (1993, 1998). These were: rejection (error in measurement: not repeatable), exclusion (report findings/no explanation necessary; anomaly the result of natural variation), abeyance (set aside for later), and theory change (develop new model/theory to explain). The subcode of “expand existing model” incorporates both the “reinterpretation” and “peripheral theory change” responses of Chinn and Brewer (1993, 1998). The rationale to combine these two responses is because the original model or theory is modified once the new data are explained. The model has broadened to include the anomalous data. There is a change in interpretation and subsequent applicability, what Chinn and Brewer call peripheral theory change. The response of “uncertainty” in Chinn and Brewer’s 1998 study describes the situation when the scientist cannot determine if the anomaly is valid or not. This response is consistent with the “set aside for later” response in the present study as well. The present study did not probe for details of the situations in which the scientists would distinguish between setting aside because they believe the data are valid and worth exploring further at some point, and those times when they would set aside data to determine the validity later.
Importantly, in opposition to the findings of Chinn and Brewer, is the finding that none of the scientists in the present study reported that they ignore unexpected outcomes. Reportedly not ignoring anomalous data is a result more consistent with Trickett et al. (2001). They explored how scientists respond to anomalous data during *in vivo* situations and found that astronomers and physicists noticed and attended to anomalies in precise ways, rather than ignoring them. An important distinction between the two studies is that the present study is based on self reports and not actual practice, as in Trickett et al. (2001). However, none of these studies compare scientists' perceptions of how they recognize and handle anomalous data to their actual practice in authentic contexts. Such a comparison was beyond the scope of the present study, but is recommended for future research.

**Justification**

These scientists recognized that the justification and acceptance of scientific knowledge depends on the context in which the knowledge is developed and the community that determines acceptability of the claims. These results suggest that negotiation of meaning is indeed a social activity (e.g. Lave & Wenger, 1991); and a particular scientific/social community has standards that are established *from within* and may not apply across the board of the scientific domain (Knorr-Cetina, 1999; Ziman, 2000). Recognizing that there are differences in standards for justifying claims, even though specific differences were not necessarily articulated by all of the participants, is evidence of awareness of science as a domain beyond the individual research context.
The group comparisons for both justification and reproducibility revealed considerably more differences than what was typically found within other categories. An interesting finding from this category was the different focus on use of statistics from the discipline and approach groups. The life scientists, chemists, experimentalists, and E/D groups emphasize statistics for justification, whereas the other groups were well below the total sample percentage (Tables 28 and 29). In fact none of the Earth and space scientists and the physicists claimed statistical analysis was important for justifying knowledge claims in their field. The focus of the Earth and space scientists was more on external reproducibility, that being other researchers conducting similar investigations and finding similar results. The physicists discussed the role of predictions and testing in knowledge justification, but with the caveat that this approach depended on the context of the investigation.

The majority of the life scientists and E/D groups’ description of use of experiment over descriptive methods to justify claims may be indicative of the standards of their community. The fact that the E/D scientists voiced this view is not surprising given they engage in both practices. This is similar to their claim of validity of methods (Table 25). Nonetheless, they still emphasized the use of statistics, in contrast to those strictly in the descriptive sciences.

Distinguishing Data and Evidence

Distinguishing data from evidence requires acknowledgement of interpretation in light of a question or problem-base. In this way, determining what data constitute evidence within a particular investigation is necessarily theory-laden. Even though 62.5% of the scientists recognized the theory-laden NOS, only 42.7%
of them explicated differences in data and evidence in terms of informed interpretation. Eight of the ten participants who were informed regarding evidence as products of interpretation, were also informed regarding theory-laden NOS. Eleven of the 15 who reported a difference between data and evidence were also informed of the theory-laden NOS. These results suggest a connection between views of theory-laden subjectivity and acknowledging a difference between data and evidence, for this sample. The slightly lower representation of the theoretical physicists who distinguished data and evidence suggest experiences with empirical data may be helpful in developing desired conceptions. Nonetheless, it is somewhat disturbing that nearly 40% of the participants did not distinguish the two concepts. Although these scientists, with the exception of the theoreticians, produce conclusions supported by evidence generated from the data, they do not necessarily hold formalized conceptions of the distinction between data and evidence.

Reproducibility

As discussed with reference to “anomaly,” reproducibility is an important element of scientific inquiry that involves reasoning processes related to identification of anomalies as well as justification of claims in the form of evidence. As several of these participants voiced, reproducibility of data becomes the evidence for a claim. Responses comparing science and art supported the importance of reproducibility in validating scientific claims, whereas unique findings were anomalous and/or fraudulent. The highest occurring subcode for this category was that reproducibility is a requirement for scientific knowledge, but this subcode was not agreed upon by a majority of the whole sample, with only 37.5% occurrence.
There was substantial diversity in how the scientists described the applicability of reproducibility, and like "justification" this diversity was somewhat discipline and approach-dependent (Tables 32 and 33). Again, the life scientists and chemists clustered together to emphasize the use of statistics to establish internal reproducibility, in contrast to the lower frequency of the Earth and space scientists and physicists in this subcode. In comparison to the Earth and space scientists and the physicists, the chemists and life scientists seem to give higher priority to the criterion of reproducibility. Clearly, the theoreticians did not give high priority to this feature within their own work, and only mentioned reproducibility within the context of other science areas. These results suggest that views of reproducibility are related to the authentic context of the participant. Ziman (2000) describes similar context-based parameters to, and expectations of, reproducibility, such as those described by the participants in this study.

Prediction

Prediction is not a typical criterion explicitly emphasized in K-12 classrooms, but the emphasis on prediction across the contexts of this study and others (Osborne et al., 2003) suggests educators may want to give attention to how prediction relates to justification, models, distinguishing data and evidence, and possible variances in applicability within scientific research. These areas should be emphasized within inquiry classrooms (NRC, 2000), and based on the present results, those who engage in authentic scientific inquiry connect prediction to all these elements.

Some of the results may be explained by context and the connection of views between prediction and the function of scientific models (Tables 34 and 35). The
Earth and space scientists, E/D scientists, and the descriptive scientists all clustered to the favor of prediction attained through models. In contrast, none of the theoretical physicists voiced this view. These results are consistent with their reported use of scientific models in their research. One would expect that since the chemists report such high claims for hypothesis testing and reproducibility, that prediction would be viewed as a requirement. However, the results here indicate only two of the chemists consider prediction important, and not necessarily a requirement. Only the life scientists and Earth and space scientists appeared in the "requirement to be scientific" subcode. Nonetheless, the physicists recognized prediction as an important component of scientific research.

Ernst Mayr (1988, 1997) states that physicists and biologists would be more likely to hold opposing views of prediction, with the biologists not placing as much credence in the role of prediction in the biological sciences as physicists do to physical science. He relates his view to the implication of scientific laws having absolute predictive power.

The ability to predict was therefore the classical test of the goodness of an explanation in physics...Prediction in the vernacular sense, that is, the foretelling of future events, is as precarious in biology as it is in meteorology and other physical sciences dealing with complex systems...the ability to predict is not a requirement for the validity of a biological theory. (Mayr, 1988, p. 19-20)

Mayr (1997) also stated, "For the biologist, it is not so important that his theory survive the test of prediction; it is more important that this theory is useful in solving problems." (p. 54) With the exception of the physicists who stated prediction and testing were important in justification of claims (Table 28), the results contradict Mayr's ideas about scientists' relative considerations of prediction (Table 34). Even
though only 25% of the participants claimed prediction was a requirement of science, those who made this statement were the life scientists and Earth and space scientists.

**Views of a Common Subcode: Variability by Context**

There was a fairly common subcode suggesting the applicability or meaning of a main category depended on the science discipline or method of investigation. The subcode relating to “Variable by context” appeared in 10 of the 16 NOS and NOSI categories. They include: justification (58% representation of the total sample), reproducibility (38%), methods (37.5%), Experiment (25%), theory/law (25%), tentativeness (16.7%), subjectivity (16.7%), purpose (12.5%), prediction (12.5%), and anomaly (8%). Despite the high occurrence of this subcode among the categories, the overall representation of participants is very low, with only “justification” having more than 50% of the participants recognizing justification of claims may depend on the context in which the claim was developed. For this subcode, views are disparate between experimentalists and the E/D combination researchers with respect to categories of experiment (Table 21) and reproducibility (Table 33). Compared to the experimentalists, the E/D scientists are more likely to claim that the experimental approach as well as the requirements and methods of achieving reproducibility differ based on the context of the investigation. The other approach groups were not noticeably different in their responses within the “variable by context” subcode. The discipline groups are more consistent. These results are intriguing, yet equivocal. At best, they suggest “variability by context” is a feature that applies to these scientists’ view of NOS and NOSI. Those who engage in
multiple types of investigations may have an advantage in recognizing and describing that variability as it applies to experimentation and reproducibility.

**Views of a Common Subcode: Statistics**

The role of statistics emerged within the categories of reproducibility (Tables 32 and 33) and justification (Tables 28 and 29). Additionally, although not a distinct subcode, statistics was mentioned in relation to subjectivity as a rationale for quantitative methods having less bias than qualitative methods. Across these main categories, the chemists, followed by the life scientists, had a greater tendency to emphasize statistics than did the Earth and space scientists or the physicists. Likewise, the experimentalists and the E/D scientists placed higher emphasis on use of statistics than the descriptive or theoretical scientists. Given the overlap of disciplines and approaches, it is impossible to distinguish among these relationships or determine interaction effects.

**Views of a Common Subcode: Hypothesis Driven**

The categories of experiment (Tables 20 and 21) and methods (Tables 24 and 25) contain a subcode related to requirements of hypotheses. Distinctions among groups emerge for this subcode. Compared to descriptive and theoretical scientists, the experimentalists and the E/D scientists have a greater tendency to view experiments and scientific methods in general as hypothesis driven processes. The chemists and the life scientists also demonstrate this tendency over the Earth and space scientists and physicists in the present sample. Again, given the overlap of disciplines and approaches, it is impossible to distinguish among these relationships
or determine interaction effects. The chemists, the life scientists, the experimentalists, and the E/D scientists are more likely to hold views in favor of hypothesis testing and statistical analysis for methods, reproducibility, and justification than are the physicists, Earth and space scientists, descriptive scientists, and theoreticians in this sample.

Consistency and Interdependence among Aspects of NOS and NOSI

The results demonstrate that, in general, these scientists held consistent views across the 16 categories of NOS and NOSI. For example those who viewed science as inherently tentative did not report that scientific laws are definitive unchanging facts. They either reported laws were also susceptible to change or laws were not applicable to science because science is inherently tentative. In general, there were few inconsistencies in responses or inappropriate examples. Moreover, the quotations and discussions show that many of the participants were able to demonstrate interconnections among aspects of NOS and NOSI. These scientists as a group were considerably more definitive and sophisticated by means of demonstrating connections, providing examples, and showing conviction in their perspectives, than teachers or preservice teachers from prior studies (Abd-El-Khalick et al., 1998; Bell et al., 2000; Lederman et al., 2001; Schwartz & Lederman, 2002; Schwartz et al., in press; Westerlund et al., 2002).
Are Variations in Scientists' Epistemological Views of Science Related to Authentic Science Context?

This investigation identified 16 unified core categories of these scientists' epistemological views of science that are applicable across the science disciplines and contexts involved in this study. On the level of broad generality, epistemological views do not seem to differ across disciplines or approaches. However, there are variations in how these scientists described the categories, and these variations are represented within the more minor subcodes. Some of the variation can be related to contextual issues of discipline and/or research approach, yet no overarching pattern emerges that serves to explain all the tendencies observed. With a few exceptions, variations in these scientists' views are more idiosyncratic, emerging at levels of specificity that are tied to individual contexts and experiences.

Distinctions cannot be made between trends based on discipline and approach. For example, the chemists and the life scientists clustered similarly with the experimentalists and the E/D combination scientists for some of the categories. All the chemists and half of the life scientists were also experimentalists, and four of the other five life scientists were E/D scientists. As such, when these groups clustered together, the common factors are both disciplinary and investigative approach.

Most notably distinct were the theoretical physicists. They frequently clustered differently from other groups, and they would not consistently cluster with the same group. That these four scientists routinely stood apart from the rest suggests
something about their experience, be it knowledge, research experience, scientific community, or something else, has lead to their different perspectives of NOS and NOSI. This group having consistently different perspectives in comparison to the other discipline and approach groups is evidence for contextually-based epistemological views of science, within this sample of scientists.

Furthermore, accounts of context-based differences have been identified for most of the main categories. Some of them overt (use of statistics for reproducibility); most of them merely suggestive (anomaly: expand existing model). Some of them make intuitive sense, such as the finding that the scientists who engage in both experimental and descriptive research had a greater tendency to state experiments are not required for the development of scientific knowledge and also that experimental methods can vary by context. All of these examples support the conclusion that there are relationships between these scientists’ epistemological views of science and their authentic science context. Clearly, the views expressed by these participants are contextualized within their authentic practice. However, the extent to which the relationships may be predictable based on discipline or an investigative context is equivocal from these results.

Why isn’t there a consistent pattern? The fact that there are typically 50% or fewer participants represented within the subcodes suggests the level of description may be too specific to identify a pattern. That is, beyond the general affirmation and broad descriptions of categories as they apply within the scientist’s research, the participants provided details very specific to their contexts. In doing so, detailed profiles and descriptions of participants’ contexts were generated. Many subcodes
emerged, and there is low overall frequency within many of the subcodes. The emergent subcodes enlighten the finer detail of these participants’ views of NOS and NOSI. It is at this level of specificity that most context-based variances in views were identified. However, given the low overall frequency, these trends are idiosyncratic in nature. It is likely that broader representation of the subcodes would dissolve these subtle differences. Broad representations of inquiry, for example, were found to apply to a range of scientific contexts in the investigation by Reiff et al. (2002). In that study, deeper contextual details were not sought. In the present study, the connections among categories suggest consolidation of categories would be possible. However, the detail of views would be lost. The emergent multiple subcodes serve to describe views more specifically such that subtleties and contextually-based connections are not likely overlooked.

Thus, for the present study, with the exception of “Justification” and “Reproducibility” and a few categories with more minor trends, there are generally no differences between disciplines or investigative approaches with respect to broad perspectives of NOS and NOSI. The more minor exceptions include “Prediction: through models”, “Models”, and “Methods: hypothesis-driven.” The results suggest the broad perspectives of NOS and NOSI are consistent with the guiding framework for what is appropriate for K-12 science curricula.

Scientists’ Epistemological Views of Science: Informed or Naïve?

A primary aim of this study was to describe practicing scientists’ epistemological views of science. Those views have been described, compared and
contrasted. Prior studies put results in terms of "adequacy" of views as they compare to the desired, contemporary perspectives. This position is considered to be the "informed" view, and the current perspective is that advocated for K-12 and post-secondary science. As is evident in the subcode percentages for the total sample and as discussed above, the results show that all of these scientists did not necessarily hold informed conceptions of all NOS and NOSI categories. Overall, however, the results show these scientists, as a group, demonstrated somewhat more informed conceptions than has previously been reported (Behnke, 1961; Bell, 2000; Glasson & Bentley, 2000; Kimball, 1967-68; Pomerory, 1993). Yet, the methodology of this study generated a great deal more description of epistemological views than any of these prior studies, and as such, it is difficult to directly compare outcomes. An additional note on the methodology is to come later in this chapter. Nonetheless, engaging in authentic scientific inquiry, as a successful member of the scientific community, is not necessarily sufficient in and of itself to ensure informed conceptions of NOS and NOSI, or conceptions the same as others within the scientific community. Those who engage in authentic scientific inquiry may or may not develop epistemological views of science aligned with positions for scientific literacy as described in current reform documents relevant for K-12 science. Scientists' views may be tightly bound to the context of the individual scientist, and, as has been shown in this study, individual contexts vary considerably across and within disciplines. An individual scientist may not consider alternate contexts any differently than his or her own. These results demonstrate success of the scientist does not depend on epistemological views of science (Elby & Hammer, 2001). The
scientists reported that they do not consider epistemological issues in their everyday work. Considering the nature of the practice and having abilities to participate successfully in the practice are two different perspectives. Scientific literacy, as described for the general citizenry, does not demand that everyone be a scientist. Rather, the intent is that people understand where scientific knowledge comes from, including the confidence and limitations of that knowledge. The level of commitment or vested interest one has within the scientific community may inhibit one's abilities to reflect and accept some of these tenets of NOS and NOSI, such as tentativeness. One can be a very successful scientist and not consider the nature of the discipline. On the other hand, as demonstrated by some of these scientists, experience within the scientific community may enable advanced views and specific insights not easily attainable without such experiences. The key may be in the ability of the scientist to reflect on his or her practice. How willing and able is the scientist, whose career depends on solid claims, to transition perspectives to consider some topics as tentativeness, subjectivity, and reliance on inference? Scientists may be literate of how they conduct and reason through their own practice, in their own community. This perspective may or may not be representative of NOS and NOSI as these concepts apply on a more general scale.

The results of the present study do show that many of these scientists were able to explicate sophisticated views of some of the identified NOS and NOSI categories. It is the detail and sophisticated explanation that is informative for understanding general and subtle connections between these categories and authentic
science examples. Many of these subtle, albeit primarily idiosyncratic, connections have not been identified in prior studies, such as those reviewed in Chapter two.

**Implications of the Study**

Elby & Hammer (2001) hold the position that "sophisticated epistemology does not consist of blanket generalizations that apply to all knowledge in all disciplines and contexts; it incorporates contextual dependencies and judgments" (p. 565). If a goal of science education is to aid the development of sophisticated epistemological views of science (whereby "sophisticated" indicates alignment with postmodern perspectives relevant to scientific literacy); and if teachers are to teach about NOS and scientific inquiry within a variety of contexts toward this goal, such as within the contexts of life science, chemistry, Earth and space science, and physics; then it seems imperative to explore possible correlations and variances in views of NOS and scientific inquiry that might stem from learning experiences within different scientific disciplines. The results of this study support Elby and Hammer’s claim of contextualized sophisticated epistemology. For the most part, it was at finer levels of specificity that the relationships between context and views emerged. The question is, then, how applicable is this level of sophistication and specificity to the K-12 learner?

**The Adequacy of the Generalized Treatment of NOS and NOSI**

The relationships that emerged in this study were primarily idiosyncratic and specific to the few scientists within subgroups. The level of specificity, the examples
provided, and the extent of the participants' reflection on their years of practice
yielded descriptions of sophisticated epistemological views, with a few relationships
among context groups. The practical application of the expressed detail to the
classroom is questionable. At the broad levels of generality, the subcodes with the
highest frequency, the scientists demonstrate overall consistency. This, and the
impracticality of introducing all the finer perspectives of authentic science practice
into school-based science, lead to the conclusion that the generalized treatment of
NOS and NOSI across science disciplines is appropriate for the K-12 context. With
the numerous distinctions and nuances associated with authentic science practices,
there is a danger of losing the “forest through the trees” if those nuances are the
focus of science instruction rather than the broader, overarching commonalities
among the contexts. That is not to say there aren’t important context-based issues
that should be addressed to promote an inclusive image of scientific practices.
Instructional objectives for NOS and NOSI are likely more practically attainable and
relevant to the goals of scientific literacy when kept at levels of generality shown
here to apply across science disciplines and approaches. It is important to be mindful
that the target of scientific literacy is the general citizenry. All individuals are
consumers of science, all types of science.

Redefining “Generalized”

The recommendation here is that K-12 science should target the development
of generalized epistemological views of science, where the “generalized”
classification includes: (1) understanding of the 16 core categories (14 common NOS
and NOSI aspects, 2 newly introduced aspects) as described in reform documents
and broadly in this study, (2) understanding connections among NOS and NOSI aspects, and (3) recognizing that aspects exhibit “variability by context.” This redefined level of generality presents a comprehensive picture of NOS and NOSI that is inclusive of authentic science practices. It recognizes that the identified and agreed upon general aspects are connected and interdependent to each other and the context of scientific inquiry. Finally, it introduces the notion of awareness that not all science is the same, and what is standard within one area may not be appropriate within another. Such awareness may promote a view of science that is more inclusive of the multiple practices, conventions, and communities. Based on the results of this study, awareness of “variability by context” is common across many aspects of NOS and NOSI.

It is also important to consider the developmental appropriateness of addressing all 16 core categories and the other two elements of this generalized classification of NOS and NOSI with respect to all K-12 learners. There is the possibility that various features of this scheme are inappropriate for certain levels of learners. This is an area of suggested future research.

Informing Explicit Instruction: Pedagogical Content Knowledge beyond Generalizations to Promote Inclusive Image of Authentic Science Practice

As has been shown, some contexts may be more conducive than others to promote desired conceptions of NOS and NOSI, as well as undesired misconceptions. Explicit teaching is necessary, regardless of context and regardless of level of generality. Indeed, explicit teaching is necessary if the concept is valued as a cognitive learning outcome, no matter what the topic. For teaching of NOS and
NOSI, teachers may need to develop conceptions beyond generalized levels of affirmation or applicability. Describing elements of pedagogical content knowledge (PCK) for NOS and NOSI is an emerging area of research (Bartholomew et al., 2002; Schwartz & Lederman, 2002). Teachers' abilities to identify and utilize science inquiry episodes, in different disciplines and through a variety of investigative approaches, to teach about NOS and NOSI very likely require more sophisticated conceptions of these aspects. That is, the more subtle and even idiosyncratic level descriptions may be important for teachers to identify and understand in order to explicitly teach NOS and NOSI within different contexts.

Furthermore, knowing that different contexts may promote certain images of science over other images may be a critical step in teachers' PCK for NOS and NOSI. For example, if (1) instruction focuses primarily on reproducible experimental methods, and (2) engaging in this approach tends to correlate with a view of justification as requiring replicas and statistics (as seen in this study), and (3) teachers explicitly teach about justification of scientific claims utilizing this experimental approach, then the students would in all likelihood develop a view of justification as requiring replicas and statistics. Part of the explicit instruction would need to raise an awareness of other examples of justification through nonexperimental valid investigations. The ability to raise such awareness would depend on the teacher’s PCK (Shulman, 1986). Within the context of a science lesson, unit, or curriculum, a teacher needs to know what elements of NOS and NOSI are applicable and, additionally, what other areas need to be addressed to fill in gaps to promote a more complete, even if generalized, view of science.
“Generalized” must present a representative picture of scientific inquiry and NOS. As such, a teacher must be aware of that representative picture and be able to select learning tasks accordingly (Gesss-Newsome, 1999; Shulman, 1986).

Why is awareness and additional attention important? Referring back to the above example, the view that all of scientific inquiry requires internal replicas and statistical measures dismisses scientific claims developed through nonrepeatable and/or descriptive approaches, such as some of the studies conducted by the atmospheric scientists and the ecologists in the present study. They still conduct scientific research, the conventions of which are established and maintained by a community of their peers. The standards for justification of claims may differ with context, but they are nonetheless scientific. As described in this study, reproducibility can be achieved in different ways. The use of internal replicas is one way. Repeating investigations several times is another way. Different researchers repeating the investigations is still another way. Sometimes results get published without being reproduced. As one of the scientists claimed, however, if the results are important, somebody will repeat them. This is what he called the “self cleansing nature of science.” Justification of scientific claims involves reproducibility, statistical analysis, empirical data, consistency, predictive ability, acceptance by peers, addressing of alternative explanations, and “changing the way people think about the world.” The extent to which all of these criteria apply to any one scientific investigation is dependent on the context of that investigation. Explicit instruction, thus, necessarily depends on the context of the investigation.
Authentic Scientific Inquiry as a Context for Developing Conceptions of NOS and NOSI

The situated cognition perspective of Brown et al. (1989) supports the integration of NOS, scientific inquiry, and science subject matter. To promote meaningful conceptions, NOS and NOSI should be explicitly taught within contextualized experiences that can include authentic science experiences (Ryder et al., 1999; Schwartz et al., in press). Furthermore, interconnections should be emphasized through explicit instruction to further contextualize the aspects.

Participants in the Osborne et al. study (2003) also indicated ideas related to the scientific endeavor were not easily separated, nor should they be. When learning about NOS, preservice teachers have reported difficulties in distinguishing individual aspects in authentic scientific contexts (Schwartz et al., in press). Similar to the scientists in this study, some of the participants in the Schwartz et al. study provided anecdotes from research settings that exemplified multiple aspects of NOS.

Awareness of an artificial separation of NOS into distinct categories was a turning point for several of the preservice teachers in developing more meaningful conceptions of NOS. The results of the present study demonstrate that a variety of authentic contexts can provide opportunities for establishing connections among NOS and NOSI aspects. Participants from all the discipline areas demonstrate connected views.
Consensus on NOS and NOSI

The results lend support to the growing consensus on aspects of NOS that scientists consider important (Osborne, et al., 2003) and commonalities of NOS and NOSI identified through review of reform documents (AAAS, 1993; NRC, 1996) and recommendations by science educators (Chinn & Malhotra 2002; Driver et al., 1996; Hodson, 1998; Lederman & Abd-El-Khalick, 2000; Millar, 1989; Millar & Osborne, 1998; Minstrell & van Zee, 2000; Reiff et al., 2002; Ryan & Aikenhead, 1992; Smith et al., 1997; Smith & Scharmann, 1999). A distinction between the Osborne et al. (2003) study and this one is that whereas Osborne et al. consulted scientists about what they thought was relevant for science teaching, this study consulted scientists about their views of science as they related to authentic research practice. As such, the present investigation presents a more comprehensive description of practicing scientists’ epistemological views of science. The present results also introduce "reproducibility" and "prediction" as important components of the scientific endeavor.

Contemporary Science Practice: Multiple Contexts, Multiple Perspectives

This study significantly contributes to overall understanding of scientific practices and scientific perspectives. The research examples and perspectives from contemporary practices place scientific inquiry in real time and real places, being conducted by real people. The variability within and among science disciplines is evident, exemplifying scientific inquiry and the communities of science as dynamic.
The examples, quotations, and results from this study may be useful for teachers and teacher educators to portray perspectives from contemporary science.

**Implications of Methodology: The Importance of Facilitating Reflection**

Reflection on experiences with science, either in the classroom or authentic setting, has been identified as a critical element necessary for formalizing views of NOS (Schwartz & Crawford, 2004). However, as several participants reported here, scientists do not typically reflect on their practice (Glasson & Bentley, 2000), and this has been posited as a reason for their typically holding more traditional, absolutist, objective views of science (Kimball, 1967-68; Pomeroy, 1993). Furthermore, a prior study of preservice teachers in a science research internship suggested those individuals most closely aligned within the science community had the most difficulty responding to philosophically-focused questions (Schwartz et al., in press). For the present study it was necessary to facilitate the participants' reflection through the questionnaires and interview prompts. Given the individualized nature of data collection, the participants in this study were encouraged to reflect on their work in a philosophical sense that was challenging and novel to most. Probing follow up questions often resulted in further details and reflections not written initially. Some stated they had continued to think about the issues after filling out the questionnaires, or gave thought between their initial reading and filling them out or giving the interview. The two-part approach may have helped these scientists to become reflective and voice their views where previously they had not considered such issues. Finally, given that there were 25 average years research experience per participant, these scientists were experienced.
Perhaps that longevity was a luxury or a comfort that was the key to transition ability (Schwartz et al., in press) needed for reflection. Their reflection yielded details of their overall sophisticated, yet variable views. Without the reflection, the details and level of sophistication, such as the connections among categories, would probably have been lacking. It is also possible that the act of reflection may act as a treatment itself. Where many of these scientists had not previously reflected upon and formalized their views, the opportunity to do so, through the guidance of the questionnaires and interview, may have influenced the views they eventually reported.

Limitations of the Study

There are several limitations to this study. First, the sample was one of volunteers, recruited from references and prior contacts. They are not necessarily representative of other scientists within their broad disciplines, subdisciplines, specialty areas, or who utilize similar investigative approaches. There are myriad subdisciplines within the broad disciplines that are not represented in the present sample. For example, the Earth and space science group did not have representation from geology. The physics group was primarily theoretical scientists. Furthermore, the participants were selected based on specific criteria of expertise. The results, therefore, are limited to this group of scientists. Caution should be exercised when attempting to extrapolate any of the findings to other populations.

Second, the criteria of expertise lead to a limitation in data collection and attrition. Experts and leaders in the scientific community have great time demands
professionally. Participation in this study put another demand on an already tight schedule for many of these scientists. As such, several scientists had to decline, even after initially agreeing. The very factors that made them attractive as participants in this study (actively researching, leaders in their community) made participating difficult for them (no spare time). In the essence of maintaining as many participants as possible, data collection had to be flexible. Interviewing was less arduous than writing out responses for several of the scientists. They gave hours of their time to talk about their work and ideas of science. Through whatever means of collection, the researcher is confident that the ideas the scientists delivered were sincere and complete.

Third, specific descriptions of the scientists’ research practices are based on self reports. No observations of the scientists in practice were conducted. The approach groups were assigned partly based on these self reports. In vivo observations of scientists’ authentic practice would have added additional contextual information that may have been omitted from the scientists’ descriptions. Additionally, what one actually does and what one thinks one does are not always the same. However, a goal of the study was to gain information from the perspective of the scientists.

Fourth, the different numbers of participants within the groupings, as well as the small total sample size, made statistical use for identifying trends problematic. This study is exploratory in nature, with general patterns emerging through qualitative methods. Differences are based on relative comparisons, not statistical significance.
Fifth, relevant topics of NOS and NOSI in K-12 science education guided the development of the questionnaires. As such, what the scientists discussed was guided by these perspectives. There may be additional features of epistemological views not elicited in this study. Nonetheless, the perspectives sought and gained through the present study were those relevant for K-12 science education.

Finally, the researcher was the main instrument of data analysis. The analysis and results are a product of the researcher’s interpretation of the data. The interpretation was based on the researcher’s knowledge and experience in science and science education. The theory-laden basis for the investigation is a recognized limitation as well as strength. The section in Chapter III “About the Researcher” provides some information about the lens through which the data were analyzed. Additionally, the detailed quotations and associated discussions of the results expose the researcher’s rationale. This information may help the reader assess the validity of the findings. An exploratory study is the product of the researcher’s perspective, and it is recognized that a different researcher may identify different features of importance within the same data set.

Recommendations for Future Research

Continued exploration of scientists’ views, practices, and relationships between epistemological views of science and scientific inquiry context is important to help understand how “activity, concept, and culture are interdependent. No one can be totally understood without the other two. Learning must involve all three.” (Brown et al., p. 33) Within this perspective, knowledge is linked to activity and the
situation under which the knowledge is acquired. The results of this study
demonstrate that contextualized views may be very specifically tied to individual
experiences, rather than broader discipline and investigative contexts. However, at
broader levels, epistemological views have similar variance within and across these
areas. Both the broad generalities and specific contextually-based differences need
further exploration to establish the generalizability of these results among other
experienced practicing scientists.

The results that emerged from the theoretical physicists are in contrast to
those one might expect based on traditional views of this discipline (Knorr-Cetina,
1999; Mayr, 1997; Spieker, 1972). This apparent disparity may suggest that the
investigative method has more of an association with resultant epistemological views
than the disciplinary context or that there is an interaction among several factors or
that these participants just individually held consistent views. This question
regarding physicists as well as the other discipline groups warrants further
investigation. A larger sample within a single discipline representing multiple
approaches would address this question.

The participants in the present study averaged 25 years experience in
scientific research, and were still active conducting research and publishing. They
were clearly experienced scientists. As discussed earlier, many of the views
expressed are more sophisticated in nature by means of examples and
interconnections than are typical from teachers and students. Are the views
expressed here a function of their being scientists in general, or a function of their
being experienced scientists? A few of the participants discussed how their present
views are different from when they were graduate students or beginning scientists. Their suggestion was that through their experiences and comfort with the system, they had came to realize science wasn’t as certain or objective as they had once thought. Comparison studies of experienced and beginning scientists might shed some light on additional factors associated with views of NOS and NOSI and developments of views during authentic science practice.

Also, because of the sophisticated and/or novel perspectives reported in this study, it seems important to compare scientists and teachers’ views, and explore sources for those views. For example, even though these scientists saw theories and laws in a hierarchical relationship, the nature of that relationship was not that theories turn into laws once “proven true.” As discussed, this view of hierarchy is different from the typical naïve response from teachers and students (Lederman, Abd-El-Khalick et al., 2002). Is this difference typical among experienced scientists and other individuals? If so, why? How do their perceptions relate to their professional roles? All such explorations would aid in understanding factors associated with developing epistemological views of science.

Further information on authentic science practice and its relation to epistemological perspectives would be gained through direct observations of scientists “being” scientists. To what extend are the perceptions of how the scientists say they practice (e.g. justify claims, deal with anomalies, are or are not subjective) compare with how they really practice? To conduct such a study, observations of scientists “in practice” should be conducted prior to eliciting epistemological perspectives through questionnaire or interview methods. This approach enables
description of practice more objectively since the researcher's observations are not colored by knowing the scientist's perceptions.

Observations of authentic practice would also probe into the intriguing contrast raised by one participant of the difference between "being a scientist" and "being in the profession of science." What are the roles scientists play in their community? The associated question of "What role is prominent in the science classroom?" would need to be addressed through investigations of classroom-based inquiry. There is again a likely level of practicality for the classroom, but as yet, this level is unknown.

Reform in undergraduate education to target objectives for NOS and NOSI necessitate exploring scientists' views and related teaching practices toward NOS and NOSI. The present study suggests these scientists are generally similar in their views on broad levels, but views of NOS do not always align with teaching (Abd-El-Khalick et al., 1998; Lederman, 1998). How scientists' perceived views of NOS and NOSI are addressed instructionally, and if they are variable across subject area, is recommended for future study.

Teacher education and professional development should consider the inquiry experiences and associated explicit instruction provided to their students. How representative are these experiences of the conventions of authentic science? Is the instruction related to NOS and inquiry representative of a particular science discipline or approach? Are additional experiences and explicit instruction needed to encompass NOS and inquiry as advocated for scientific literacy? If one discipline or one approach is utilized to teach about NOS and NOSI, the resultant view may not
represent science as a complete domain. Further study is needed to explore the impact of single versus multiple experiences on epistemological views of science.

The developmental appropriateness of the 16 categories, interconnections, and awareness of “variability by context” needs to be explored. As learners develop cognitively and gain experiences within science, their abilities to develop more sophisticated NOS and NOSI knowledge may also advance.

Finally, this study offers a starting point for further explorations of any of the 16 categories. Questions for future research are generated from these scientists’ perspectives and reports of authentic scientific inquiry and how these images relate to classroom science. One category is detailed here. The following example addresses the recognition, reaction toward, and handling of anomalous data.

Chinn and Brewer (1998) raised several questions concerning influences of subject matter knowledge and experience within the scientific community on how scientists would deal with anomalous data across their eight categories. Although the present study did not pose situations for these participants to demonstrate their reasoning, or observe them in authentic practice, the self reports suggest that once the anomalous data are identified are valid, they often lead to progress in the field. The notion of setting aside for later (abeyance) was not a frequent response, and ignoring was not a stated option. There is an apparent discrepancy between what happens when scientists identify an anomaly and what is typical in school science (Chinn & Malhotra, 2002). In school science anomalies typically mean an error was made and there is a need to redo or change the hypothesis (Chinn & Malhotra, 2002). True anomalies in the scientific community lead to progress (Kuhn, 1970).
Contributing such progress is how success is measured for the scientist, and how one becomes an established member of the community (Lave & Wenger, 1991; Wenger, 1998). In addition to the cognitive responses that lead to scientific progress, according to some of these scientists’ reports, anomalies also generate excitement. They are what make science fun! To what extent can the same be said for the typical school science experience? How alike is knowing from the beginning of an investigation the report must contain a section called “error analysis” to knowing there might be something interesting that would expand your knowledge base, leading to success and progress? In an effort to enhance meaningful learning, anomalies have been introduced into school science as an instructional component to promote conceptual change (Chinn & Brewer, 1998), but the extent to which teachers can effectively introduce and promote authentic responses, cognitive and affective, to anomalous data within the school setting remains to be explored.
REFERENCES


Central Association of Science and Mathematics Teachers (1907). A consideration of the principles that should determine the courses in biology in the secondary schools. *School Science and Mathematics, 7*, 241-247.


inquiry-enthusiast's approach to teaching science as inquiry. Paper presented at the annual meeting of the National Association for Research in Science Teaching, St. Louis, MO.


APPENDICES
Appendix A

Initial Query

Dear ______________:

My name is Renee' Schwartz and I am a doctoral student in Science Education at Oregon State University. I am beginning my dissertation research soon and am asking for the participation of scientists to help with my investigation.

My research project will focus on the work of research scientists and the views scientists have related to their practices. This investigation aims to help inform educators of real-world practices and people in scientific research. As such, I am seeking the participation of scientists who are actively conducting and publishing original scientific research.

If you choose to participate in the study, I will ask you to provide me with a copy of your vita and a written description of your current work. These documents will inform me of your educational and professional backgrounds and current research focus. I will also ask you to complete two questionnaires and one interview. The questionnaires both contain open-ended questions aimed to probe your views related to certain issues about science and scientific practices. There are no right or wrong answers to any of the questions. I am simply interested in your views, as a research scientist. Next I will ask to interview you about your responses to the questionnaires. The interview may be in person or over the phone. The interview can be done in one or two sittings, depending on time availability. Total time for the interview will be approximately 1-1.5 hours and will be audiotaped.

Additionally, I may ask to conduct on-site observations in your laboratory. In this portion of the study I aim to describe more in-depth the daily practices of various scientific research settings. The extent and nature of these observations will depend on availability and research schedule of the scientists. Ideally, observations will span a 4-6 week period, about 5 hours per week, during which I will have opportunities to observe daily interactions, lab meetings, research procedures, and any other relevant happenings within the setting that would help me understand the processes involved in conducting the type of research that is specific to that setting. These observations are not intended to be obtrusive or disruptive to the normal practices of the scientists.

Confidentiality will be maintained through the use of coding, rather than names, on questionnaires, interview cassette tapes and transcripts, and field notes. Additionally, these data sources will be kept in a locked location at all times and be available only to my major professor and me. Any publications that result from this investigation will use pseudonyms to maintain the anonymity of the participants. Please note that your participation in this research is voluntary and you may withdrawal from the investigation at any time without consequences.

The information that you will provide through your participation in this project is extremely valuable to my research. I would like to thank you in advance for your consideration.

Please contact me at schwarre@ucs.orst.edu or (541) 737-2545, or return the enclosed card, to indicate your interest in participating in this study or for further information. Please indicate if you are also willing to open your research setting to the observation phase of this study.

You may also contact my Major Professor, Norm Lederman (ledermann@iit.edu; 312-567-3658) or my Co-Chair, Larry Flick (larry.flick@orst.edu; 541-737-3664) for information.

I look forward to the opportunity to work with you.

Sincerely,

Renee' Schwartz
appendix B
informed consent form

Dear [blank]:

My name is Renee' Schwartz and I am a doctoral student in Science Education at Oregon State University. I am beginning my dissertation research soon and am asking for the participation of scientists to help with my investigation.

If you choose to participate in the study, I will ask you to provide me with a copy of your vita and a written description of your current work. I will also ask you to complete two questionnaires and one interview during the spring of 2003. The questionnaires both contain open-ended questions aimed to probe your views related to certain issues about science and scientific practices. The interview will focus on your responses to the questionnaires. There are no right or wrong answers to any of the questions. The interview can be done in one or two sittings, depending on time availability. Total time for the interview will be approximately 1-1.5 hours and will be audiotaped.

Additionally, I may ask to conduct on-site observations in your laboratory. The extent and nature of these observations will depend on availability and research schedule of the scientists. Ideally, observations will span a 6-8 week period, about 5 hours per week, during which I will have opportunities to observe daily interactions, lab meetings, research procedures, and any other relevant happenings within the setting that would help me understand the processes involved in conducting the type of research that is specific to that setting. These observations are not intended to be obtrusive or disruptive to the normal practices of the scientists.

Interview audio-tapes will be transcribed for subsequent analysis. Confidentiality will be maintained through the use of codes, rather than names, on data sources (questionnaires, cassette tapes and transcripts, and field notes). All data will be kept in a locked location at all times. The only people who will have access to the data will be my major professor and myself. Any publications that result from this investigation will use pseudonyms to maintain the anonymity of the participants.

There are no foreseeable risks to participants in this project. Other than making a valuable contribution to science education, there are no direct benefits to participating in this project. Please note that your participation in this research project is voluntary and you may withdraw from the investigation at any time.

If you have any questions about the study or specific procedures, please contact me at schwarre@ecs.orst.edu (phone 541-737-2545), my major professor, Dr. Norman Lederman, at ledermann@iit.edu (phone 312-567-3658), or the Co-Chair of my doctoral committee, Dr. Larry Flick, at larry.flick@orst.edu (phone 541-737-3664). If you have questions about your rights as a research participant, please contact the Oregon State University Institutional Review Board (IRB) Human Protections Administrator at IRB@oregonstate.edu (phone 541-737-3437).

Thank you for your consideration. I look forward to the possibility of working with you.

I agree to participate in this research project and understand the general intent of the study, the types of data that will be collected, and the anticipated time commitments.

Signature ______________ Date ______________

(Please return one signed copy of this form with a copy of your vita and a description of your current research projects. Retain the second copy of this form for your information.)
Appendix C

VNOS-Sci

Name: ________________________________

Date: ________________________________

Science research area/discipline: ______________

Instructions

☐ Please answer each of the following questions. Include relevant examples whenever possible. You can use the back of a page if you need more space.

☐ There are no “right” or “wrong” answers to the following questions. I am only interested in your opinion on a number of issues about science.

These questions aim to elicit your views concerning science as it is practiced within your own research area. Please consider this authentic context in your responses.

1. What, in your view, is science? What makes science (or a scientific discipline such as physics, biology, etc.) different from other disciplines of inquiry (e.g., religion, philosophy)?

2. How are science and art similar? How are they different?

3. Science textbooks often represent the atom as a central nucleus composed of protons (positively charged particles) and neutrons (neutral particles) with electrons (negatively charged particles) orbiting that nucleus.

   (a) How certain are scientists about the structure of the atom?
   (b) What specific evidence, or types of evidence, do you think scientists used to determine what an atom looks like?
4. Is there a difference between a scientific theory and a scientific law? Illustrate your answer with examples from your own research, if appropriate. If not appropriate, explain why and provide examples from another area of science.

5. (a) After scientists have developed a scientific theory, does the theory ever change?

- If you believe that scientific theories do not change, explain why. Defend your answer with examples.

- If you believe that scientific theories do change:
  (a) Explain why theories change?
  (b) Explain why we bother to learn scientific theories. Defend your answer with examples.

(b) After scientists have developed a scientific law, does the law ever change?

- If you believe that scientific laws do not change, explain why. Defend your answer with examples.

- If you believe that scientific laws do change:
  (a) Explain why laws change?
  (b) Explain why we bother to learn scientific laws. Defend your answer with examples.

6. (a) What is a scientific model?

(b) What is the purpose of a scientific model?
(c) Describe a scientific model from your own area of research, if appropriate. If you do not use scientific models, describe a scientific model from another area of research. Describe why your example is a scientific model.

7. Scientists perform experiments/investigations when trying to find answers to the questions they put forth. Do scientists use their creativity and imagination during their investigations?

- If yes, then at which stages of the investigations do you believe that scientists use their imagination and creativity: planning and design; data collection; after data collection? Please explain why and how scientists use imagination and creativity. Provide examples from your own work.

- If you believe that scientists do not use imagination and creativity, please explain why. Provide examples from your own work.

8. It is believed that about 65 million years ago the dinosaurs became extinct. Of the hypotheses formulated by scientists to explain the extinction, two have enjoyed wide support. The first, formulated by one group of scientists, suggests that a huge meteorite hit the Earth 65 million years ago and led to a series of events that caused the extinction. The second hypothesis, formulated by another group of scientists, suggests that massive and violent volcanic eruptions were responsible for the extinction. How are these different conclusions possible if scientists in both groups have access to and use the same set of data to derive their conclusions?

9. Some claim that science is infused with social and cultural values. That is, science reflects the social and political values, philosophical assumptions, and intellectual norms of the culture in which it is practiced. Others claim that science is universal. That is, science transcends national and cultural boundaries and is not
affected by social, political, and philosophical values, and intellectual norms of the culture in which it is practiced.

- If you consider science to be reflective of social and cultural values, explain why and how. Defend your answer with examples from your own work.
- If you consider science to be universal, explain why and how. Defend your answer with examples from your own work.
- If you view some science as universal and some as reflective of social and cultural values, explain why and how. Defend your answer with examples from your own work.
Appendix D

VOSI-Sci

Name: ____________________________

Date: ____________________________

Science research area/discipline: ______________________

Instructions

☐ Please answer each of the following questions. Include relevant examples whenever possible. You can use the back of a page if you need more space.

☐ There are no “right” or “wrong” answers to the following questions. I am only interested in your opinion on a number of issues about science.

These questions aim to elicit your views of scientific inquiry as it is practiced within your own research area. Please consider this authentic context in your responses.

1. What scientists choose to study and how they learn about the natural world may be influenced by a variety of factors. How do scientists decide what and how to investigate? Describe all the factors you think influence the work of scientists. Be as specific as possible.

2. (a) How would you define “scientific inquiry” as it is conducted in your field of research? Give an example from your work that represents your view.

   (b) Do you think your approach to scientific inquiry is representative of all scientific research, some fields of research, or specific only to your field? Explain with examples.

3. (a) What do you think the purpose(s) of scientific inquiry is? Describe how your work represents this purpose?

   (b) Do you think all scientific inquiry has the same purpose? Why or why not?
4. (a) Write a definition of a scientific experiment?

A scientific experiment is.....

(b) Give an example from something you have done or heard about in science that illustrates your definition of a scientific experiment. Explain why your example is an experiment.

(c) Do you think that the generation of all scientific knowledge requires the use of experiments? If so, explain why. If not, explain why not.

5. A person interested in animals looked at hundreds of different types of animals who eat either meat or plants. He noticed that those animals who eat similar types of food tend to have similar teeth structures. For example, he noticed that meat eaters, such as lions and coyotes, tend to have teeth that are sharp and jagged. They have large canines and large, sharp molars. He also noticed that plant eaters, such as deer and horses, have smaller or no canines and broad, lumpy molars. He concluded that there is a relationship between teeth structure and food source in the animals.

(a) Do you consider this person’s investigation to be scientific? Please explain why or why not.

(b) Do you consider this person's investigation to be an experiment? Please explain why or why not.

(c) Compare the investigative approach used in the teeth example with the types of investigations you conduct in your own research. How are they similar? How are they different?

6. The person in the teeth investigation (#5) identified a correlation between teeth structure and food source.

(a) Do you think this person is also justified in concluding that natural selection shapes the teeth of animals to fit specific food resources? Please explain why or why not.
(b) What type of information is critical for scientists in your field of research to justify and accept a scientific claim?

(c) Do you think all fields of science hold the same conventions as you stated in (b) for justifying and accepting scientific claims? Explain.

7. (a) What does the word “data” mean in science? Provide an example from your own work.

(b) Is “data” the same or different from “evidence”? Explain and provide an example from your own work.

(c) Compare how you identify and use of evidence in your research to that in other types of science. Do you think all scientists identify and use evidence in the same manner? Why or why not?

8. Scientists sometimes encounter inconsistent findings (anomalous information). Consider such a case in your own research to answer the following questions:

(a) How are anomalies identified? (i.e. What is considered “inconsistent” in your field of research?) Provide an example.

(b) What does the scientist do when an anomaly is identified?

(c) Do you think all scientists identify and handle anomalous information this same way? Why or why not?