Exploring high school students’ use of theory and evidence in an everyday context: the role of scientific thinking in environmental science decision-making

Fang-Ying Yang, Department of Earth Sciences, National Taiwan Normal University, Taiwan, R. O. C, e-mail: fangyang@cc.ntnu.edu.tw

This study examined 10th-grade students’ use of theory and evidence in evaluating a socio-scientific issue: the use of underground water, after students had received a Science, Technology and Society-oriented instruction. Forty-five male and 45 female students from two intact, single-sex, classes participated in this study. A flow-map method was used to assess the participants’ conceptual knowledge. The reasoning mode was assessed using a questionnaire with open-ended questions. Results showed that, although some weak to moderate associations were found between conceptual organization in memory and reasoning modes, the students’ ability to incorporate theory and evidence was in general inadequate. It was also found that students’ reasoning modes were consistent with their epistemological perspectives. Moreover, male and female students appear to have different reasoning approaches.

Introduction

This study explored high school students’ use of theory and evidence in the context of learning about a socio-scientific issue: the use of underground water. There were three specific objectives in the study. One was to investigate students’ reasoning modes while formulating personal theories and to investigate to what extent they were able to evaluate or support their theories based on evidence. Another objective was to inspect students’ background knowledge about underground water in an attempt to draw associations, if any, between reasoning modes as assessed in the first objective and existing organization of knowledge in memory. Finally, an in-depth analysis was made of students’ reasoning patterns based on self-reflections, to elucidate possible meta-mechanisms that may have contributed to the thinking behaviours. Gender differences were also investigated in relation to this third objective.

Rationale

By nature, science is both a process of justifying knowledge (what we know) and of discovering knowledge (how we know) (Duschl 1990). However, traditional science classroom activities mainly focus on the part of what we know. In the science education community there is a growing consensus that we need to introduce learners to the another important facet of science; that is, how we create new knowledge. In other words, students should be better educated in the use of certain
established ways of thinking in science (for example, Duschl 1990, Lawson et al.
2000). The ‘certain established ways of thinking in science’ commonly referred to
as scientific thinking, are portrayed by philosophers of science as a process of
contexts, the scientific argumentation is the process of making assertions based on
observable and/or quantitative evidence. Seigel (1988) has noted that commitment
to evidence is an imperative trait of rational reasoning in many disciplines, although
the form it takes may vary with the disciplines. This raises the question whether
skills are always domain specific or can be general and widely applicable across
different contexts. In both developmental psychology and education there is a
strong belief that general cognitive skills do exist (Baron 1988, Lawson 1995,
Lawson et al. 2000, Perkins and Salmon 1989). The general cognitive skills are the
evaluative reasoning skills consisting of searching for possibilities and testing the
possibilities with accountable reasons (Baron 1988, Perkins 1989). Apparently, this
form of rational reasoning, when placed in science-related contexts, is congruent
with scientific thinking. Kuhn (1993), while recognizing science as argument, had
suggested that to describe scientific reasoning as argument is also sensible in
informal situations. Hence, although scientific reasoning is often discussed within
specific knowledge domains, this form of thinking is believed to be domain
independent.

Empirical studies on scientific thinking largely focus on a hypothesis testing
strategy (Kuhn 1991, Kuhn et al. 1988, Lawson et al. 2000), which highlights the
determining effect of evidence on the success of inference or theory-based hypothe-
sis (Baron 1988, Kuhn 1991). This hypothesis testing habit is important in daily
experiences for it is an important way to make rational and sound judgements on
controversial issues in a social context that increasingly have significant conse-
quences for the future of society. However, Kuhn (1991) reported that even adult
thinkers often do not have adequate ability to make scientific arguments regarding
social issues. One of the challenges presented to us, therefore, is how to create learn-
ing situations where these powerful forms of thought can be anchored in contexts
that may generalize beyond science classrooms.

As already mentioned, traditional science classrooms emphasize largely scien-
tific knowledge without paying much attention on the discovery process of science,
not to mention the prevailing neglect of science learning in a social context. This
kind of science teaching and learning emphasizing science outcomes rather than
processes has been seriously criticized by philosophers of science and science educa-
tors (Driver et al. 2000, Duschl 1990, Hodson 1986). To establish the link between
the nature of science and science education, Driver et al. (2000), based on a philo-
sophical point of view, advocate the practice of scientific argumentation in science
classrooms. On the other hand, developmental psychologists suggest that effective
classroom activities should include generating meanings, making inferences and
testing constructed theories (Carey 1986, Champagne et al. 1982, Osborne and
Wittrock 1983). Moreover, to promote the practice of thinking skills, classroom
activities must be placed in meaningful contexts (Glaser 1984, Perkins and Grotzer
1997).

Above all, the importance of incorporating situated, argumentative activities in
science classrooms is undisputed. To provide this kind of science instruction, a
deeper understanding is needed on students’ existing competence relevant to the
issue. To better understand students’ ability of making rational arguments about
social issues, this study examined high school students’ general hypothesis testing behaviour and underlying factors that promote or inhibit the behaviour.

**Literature survey**

This study used an environmental issue as the research context to inspect participants’ general hypothesis testing competence regarding making personal theories and testing the theories with evidence in a social setting. Contrary to the well-defined problems where criteria for hypothesis testing are straightforward, hypothesis testing on ill-defined problems, such as debates about many social and environmental issues, becomes difficult because evidence is by and large implied, sometimes of less certain origin, and information presented to thinkers is often multi-dimensional in nature. In such cases, consequent judgements might be biased or misled if there is not a clear understanding of the line of evidence being discussed or if there is a lack of understanding of which, among the several dimensions of evidence, is being considered at a given point. Given these complex issues associated with use of evidence in complex, everyday problem situations, we need more intensive investigation of students’ potential to use multiple sources of evidence and to critically judge claims based on evidence.

Most studies reported in the literature about scientific reasoning have been done in the context of school subjects rather than with everyday societal and environmental problems (for example, Newton 1999, Samarapungavan 1992, Schauble 1996, Vosniadou and Brewer 1992, Wong 1996). Only a few studies put emphasis on reasoning in areas of daily life situations (for example, Jiménez-Aleixandre and Pereiro-Munoz 2002, Kortland 1996, Kuhn 1991, Ratcliffe 1996, Yang and Anderson, 2003). Research on thinking within a school context seems to support the view that young students have the ability to practice scientific reasoning within the context of the subject domain. However, when placed in the life contexts, even adults displayed inadequate ability to make scientific arguments about social issues (Kortland 1996, Kuhn 1991, Ratcliffe 1996). In a previous study, Yang and Anderson (2003) showed that high school students’ use of scientific information in reasoning was often judged to be simple and could be easily affected by emotional factors that tinged the evaluation of the evidence. Superficially, studies concerning scientific reasoning in school and social contexts seem in conflict with each other. In some cases they show evidence of rational reasoning and critical thinking, while in other cases less conclusive evidence in support of critical reflection about everyday issues is reported. Part of the explanation for the differences may be attributed to the difficulty of skill transfer across different contexts for problem solving as has already been cited.

Given the apparent difficulties in educating students to generalize these reasoning skills, a review of pertinent literature sources dealing with psychological factors promoting or limiting such thought is presented. The factors discussed here are metacognition, personal epistemological belief and gender difference. From a cognitive perspective, metacognition is the high-order cognition that regulate thinking and problem-solving (for example, Flavell 1979, Garner 1987, Schraw and Moshman 1995). Empirical data indicate that metacognition predicts academic success and some metacognitive skills are domain general and can be improved through experience and practice (Hartman 2001). Thus, such metacognitive capacities may be good candidates for improving the kinds of generalized reasoning skills
that we hope to help students develop in learning to deal with environmental-based social issues.

Recent psychological research reveals an important underlying factor affecting the development of metacognition. Kitchener (1983) stressed that perhaps even higher than the metacognitive level in information processing is the epistemic level, which concerns people’s beliefs about knowledge. Although definitions vary in accordance with different research contexts, most scholars differentiate the realm of personal epistemology into three categories, which are consistent with three philosophical positions; namely, positivism, multiplist, and contextualism (or constructivism) (for example, Kuhn 1991, Perry 1970, Roth and Roychmoudhury 1994). Results of various epistemological studies often show that personal epistemological perspectives instigate different reasoning consequences (Belenky et al. 1986, King and Kitchener 1994, Perry 1970). As far as learning is concerned, the developmental epistemological stages appear to associate with learning strategies and school performance (Perry 1970, Schommer 1993). Apparently, status of personal epistemological development is an important factor guiding reasoning and decision-making. In the study reported here, this personal quality is examined to provide explanatory information for the students’ reasoning behaviour.

Another issue considered here that might also give weight to the reasoning performance is the factor of gender. In a previous study that examined students’ reasoning patterns about a nuclear energy issue (Yang and Anderson, 2003), significant differences between males and females were found in the preferred information used in reasoning and background knowledge. More generally, there is a large volume of educational research concerning gender differences, especially gender in relation to learning styles, academic achievements, interests in academic studies, and epistemological beliefs (for example, Belenky, et al. 1986, Jones et al. 2000, Richardson and King 1991, Severiens and ten Dam 1994, Speth and Brown 1990). Paradoxically, some of the research findings are not consistent, but show that gender effects vary across different situations and task demands. Severiens and ten Dam (1997) proposed that gender differences may vary according to the context of the task. Apparently, the role of gender differences in scientific reasoning in everyday problem contexts, such as the kind investigated here, needs to be examined more extensively.

In short, to have an in-depth understanding about learners’ reasoning behaviours, it is not sufficient to only document students’ reasoning behaviour, but it is also important to probe the mechanisms that underlie the surface performance. In this study, students’ use of theory and their perspective about the role of evidence was examined by placing them in a role of discussing controversial social issues. Concurrently, by asking them to reflect on their own thoughts, evidence was gained about their epistemological perspectives and its relationship to performance on the task variables in this study. Moreover, possible gender effects on patterns of reasoning were also examined by comparing differences in responses by male and female students.

Method

Subjects

There were 45 male and 45 female students who participated in the study. These students came from two intact, single-sex classes (one male and one female) in a
midsize academic high school located in an urban city in the east part of Taiwan. Except for grouping based on gender, which has been a traditional setting, these 10th-grade students can be regarded as relatively homogeneous on several relevant academic dimensions and represent a sufficiently random sample for the purposes of this study. First of all, students in Taiwan must take an entrance examine before admission into senior high schools (equivalent to the level of the 10th to 12th grades). Since students in general were appointed to a certain school based on their examination results, the students in this study, who were the freshmen in the target school, were assumed to have similar academic competencies by the time they received the designed instruction. Second, according to school records, most students in the study came from middle-class families. Accordingly, the participants were assumed to have a similar socio-economic background. Finally, students were assigned to classes without regard to academic qualifications or other stratifying variables; hence, for all practical purposes, they were randomly assigned to a particular class.

Data collection and analysis

To assess students’ use of personal theory and evidence, a questionnaire with open-ended questions was given to the participants. The beginning of the questionnaire was a news report describing some town residents’ resistance to a well-drilling project by the government water agency because a previous ground subsidence incidence had occurred during a major earthquake two years ago. The residents were afraid that the excessive use of underground water would induce further ground subsidence. Students were then given questions to explore their own thoughts. The questions used and the respective coding targets are presented in Table 1.

The questionnaire was given to two content experts to verify the surface validity, and both concurred the basic four open-ended questions were appropriate for the research objectives but more follow-up questions were needed to probe students’ thought in depth. According to the suggestion, questions three and four as presented in Table 1 were added with two follow-up questions in order to gain more information about students’ views toward evidence and what they believe would count as evidence. This questionnaire was further vetted through a pilot test. Content

<table>
<thead>
<tr>
<th>Question set</th>
<th>Coding target</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. What do you think caused the previous ground subsidence in the town? Why?</td>
<td>Formulating the personal theory</td>
</tr>
<tr>
<td>2. Do you think the residents’ resistance was reasonable? Why?</td>
<td>Referring personal theory to make a judgement</td>
</tr>
<tr>
<td>3. Do you believe the claim made by the water company that they had done careful investigation before the decision of well drilling was made? Why?</td>
<td>Identifying a need of the concrete and clear evidence to back up water company’s claim</td>
</tr>
<tr>
<td>• What could they do to make you believe?</td>
<td></td>
</tr>
<tr>
<td>4. Are you sure about your answers? Why?</td>
<td>Self-reflections on one’s own thoughts</td>
</tr>
<tr>
<td>• What information would you need to help making better judgement?</td>
<td></td>
</tr>
</tbody>
</table>
analysis of the subjects’ written responses to the questionnaire based on a standard coding target list (table 1) was used to interpret the data. Two independent coders analysed all subjects’ responses using the coding list, and the inter-coder agreement was 91%. The minor disagreement was also resolved through discussions.

To document the students’ understanding of the major ideas presented in the initial instructional unit on the basic scientific issues (as explained later in the section on Research strategy) and to determine the degree of knowledge networking in memory, a flow-map method was used (for example, Anderson and Demetrius 1993). The evidence of the amount of knowledge and its organization in memory as determined from the flow map analyses was also used to determine whether there was a relationship with the students’ use of theory and evidence. Prior research (Anderson et al. 2001) has shown that the amount of information networking in memory as measured by flow maps is positively correlated with higher-order thinking skills, but that research was within a subject domain of evolution, not a social-based, scientific issue as used here. Flow map analysis is a method of examining narrative to detect the sequential progression of ideas elicited from the respondent and the amount of ideational networking inherent within the narrative as evidenced by explicit cross-linking of ideas in the stream of discourse (Anderson and Demetrius 1993, Anderson et al. 2001, Bischoff and Anderson 1998, Yang 2002).

To construct a flow map, student narrative is elicited either as a tape-recording of spoken discourse or as a written narrative. In this study, the students were asked to practice assembling their thoughts about the cause of land subsidence and then write down sequentially all of the ideas that they could mobilize about this concept. As described in appendix 1, subjects were first asked to write down whatever concepts they can immediately think of about the cause of land subsidence, and then explain thoroughly each concept. The written narratives were inspected and transformed by the researcher into a sequential conceptual map, including the cross-linking arrows between statements containing cross-referenced ideas according to the standard procedures of Anderson and Demetrius (1993) (see examples in appendix 1). The number of sequential and network linkages was recorded as the indicators of the students’ concept achievement. Twenty examples of students’ responses were randomly selected for cross-checking, and the inter-coder agreement was over 88%.

Statistical analyses in the study include chi-square analyses, paired t-tests, one-way analysis of variance (ANOVA) and correlation analyses. Chi-square analysis tests the homogeneity of two nominal variables. For example, the chi-square test of homogeneity will reveal whether there is a difference between genders and types of theory formulated. The value of Pearson’s $\chi^2$ as well as its probability value ($p$), and the adjusted standardized residuals ($R$), are reported to indicate the statistical significance. The Pearson $\chi^2$ value indicates whether the two variables have an effect on the sample frequency distributions, while the residuals ($R$) will specify which levels or cells of the variables are the major contributors. The $R$ in the absolute value must be larger than 2 at the 95% significant level (Hinkle et. 1998). The paired-sample $t$-test in the study examines the significance of subjects’ concept improvement. The effect size ($d$) of the concept increase is also reported to indicate the effect of the improvement (Cohen 1986). According to Cohen, a value of effect size ($d$) ranging from 0.2 to 0.5 indicates a low effect, 0.5–0.8 a median effect, and over 0.8 a high effect. ANOVA was used to examine the association between nominal and continuous variables; for instance, gender versus concept achievement. The Pearson
correlation analyses were conducted to determine whether there are associations between students’ existing concepts and the amount of scientific information they expressed in their reasoning.

**Research strategy**

Before assessing students’ thinking modes, a Science, Technology and Society (STS)-oriented course with the use of underground water as the main issue was presented to the students. STS instruction introduces scientific knowledge by means of a discussion of some relevant social issue that promotes student interest and mobilizes relevant personal knowledge to help anchor subsequent science content that is the focus of the lesson (Yager 1996). The target issue frequently appeared in media has received nation-wide attention because several cities in Taiwan are suffering from the land subsidence caused by the excessive use of underground water. It further resulted in debates about whether economical development should surpass the environmental protection. Therefore, it is believed this issue provide a rather familiar context for students in general to practice rational reasoning.

The STS-oriented course used here, which lasted for three weeks (two hours per week) and instructed by the same teacher, contained information about the formation of underground water and the possible disasters that can arise due to the excessive use of the water. These disasters included ground subsidence, backward flow of seawater, and so forth, and were presented to students as the main discussion topics. At the end of the class activities, a forum was set up among the students to consider the future of some hypothetical residents who live in the area suffering the consequences of excessive use of underground water. To prepare the forum, students were asked to play different roles in groups, including geologists, government agents, politicians, and residents. Some students were selected to be independent judges. Each participating group was allowed two weeks of time to collect information and formulate arguments related to their pretending positions. All the arguments were then presented and discussed in the forum, and the independent judges rated the performance of each group according to the strength of information that backs the related arguments and role players’ reactions to the residents’ questions. The evaluations from the independent judges were reported at the end of the forum. The main purpose of the course was to enhance students’ background knowledge of the issues with the aim of creating a coherent knowledge base to facilitate further learning. Before and after the course, a flow-map analysis was used to examine students’ progress on concept learning. The questionnaire with open-ended questions eliciting student opinions and evidence of rational thinking was given last.

**Findings**

**Concept performance**

The paired-sample $t$-tests on the flow-map narratives showed that students’ concepts about land subsidence and the amount of ideational networking increased significantly after the STS instruction (table 2). Examples of flow maps appear in appendix 1.
As table 2 shows, a statistically significant increase in linear linked concepts was found after the STS instructional session. The effect size \( (d = 0.91) \) indicates a high effect from the instruction. Although the number of cross-linked ideas also increased significantly, the effect size \( (d = 0.46) \) showed only a low to median effect. As a matter of fact, the range of increased complex links was low (from 0 to 2). According to the data, only two subjects (out of 90) made one complex link before the instruction, while after the instruction 16 subjects were able to make one complex link and two made two complex links.

Data collected from the open-ended questionnaire showed that most students (over 98%) were able to identify and explain the causes of ground subsidence even though the quality of the explanations varied. Following are two response examples:

**Question set #1:** What do you think caused the previous land subsidence in the town? Why?

**Student 1, male:** The previous land subsidence was caused by the fault movement induced by the earthquake.

**Student 2, female:** I think it’s the earthquake which caused the land subsidence in that crustal movement would result in the land subsidence.

**Modes of reasoning: the use of theory and evidence**

**Use of theory** About 70% of students used theories formulated in question 1 to evaluate the town residents’ action. When gender is taken into account, with respect to references to personal theories, males used this reasoning pattern more frequently than females when they were asked to judge whether the residents’ actions were
reasonable (table 3). The gender difference was statistically significant ($F = 10.95$, $p < 0.01$).

Some examples of student narrative illustrating gender differences in the use of theory are presented:

Question set #2: Do you think the residents’ resistance was reasonable? Why?

Male, Student 1: No. I don’t think they were reasonable because it was the earthquake that resulted in the previous land subsidence. The land subsidence would not take place just because some holes were drilled into the ground. Besides, before any well drilling, some field investigations should have been done.

Female, Student 2: Yes. They were reasonable. No matter the underground water has been overly used or not, the amount of underground water should be always maintained. There were three wells within area of 1 square km, a fourth will only add problems to the land.

As presented in this selected narrative, the male subject used his personal theory to make some inferences while the female did not.

Based on students’ written responses, three types of hypotheses regarding the cause of land subsidence that the students identified were tallied. One was the previous earthquake as evidenced by the number ($n$) of students who proposed this cause ($n = 26$), another the excessive use of underground water ($n = 35$), and the other hypothesis the combination of the former two cases ($n = 24$). According to the chi-square tests, no difference was found between gender and the type of personal hypothesis made by the subjects. However, the type of hypotheses posited by the respondent had an effect on whether or not the respondent used theory to judge the rationality of the residents’ protests (Pearson $\chi^2 = 6.7$, $p < 0.05$). Further cell analysis indicated that those subjects who thought that a combination of earthquake effects and use of ground water might be a cause were more likely to relate their judgements to a previously stated theory ($R = 2.4$). Furthermore, this effect was evident in the female group (Pearson $\chi^2 = 9.78$, $p < 0.05$) while in the male group no association was found (Pearson $\chi^2 = 0.72$, $p > 0.1$).

**Use of evidence** Regarding the role of evidence, only 57% of the participants pointed out that ‘evidence’ should be provided by the water company to support their claim. The following are two illustrative responses given by a student who did not offer evidence (Student 3) and one who did (Student 4):

Question set #3: Do you believe the claim made by the water company that they had done careful investigation before the decision of well drilling was made? Why? What could they do to make you believe?

Student 3: I don’t believe their claim. There existed three wells already. No need to drill one more … I think the water agency should hold a public meeting for the residents to participate and understand their work.
Student 4: I don’t trust them completely because they did not provide any evidence. I think they should provide clear evidence. For example, they should test or simulate if there would be further land subsidence ... The most important thing is that they should reveal both the advantages and disadvantages regarding the well drilling to the public.

The first response was coded as ‘No evidence identified’ since the student largely argued that a public information sharing session should be held. This session was to be convened not to provide evidence, but to help the residents ‘Understand their work’. The second response was coded as ‘Evidence identified’ since the student clearly called for evidence to be presented and suggested some ways it may be obtained. No statistically significant difference was found between male and female students (table 4).

The chi-square test of homogeneity showed no association between the use of personal theory and the identification of the role of evidence (Pearson $\chi^2 = 0.042$, $p > 0.1$).

**Associations between existing concepts and reasoning modes**

One-way ANOVA showed that there seemed to be a statistical association between students’ post-test sequential concepts as analysed by the flow map technique and their use of personal theory ($F = 8.56$, $p < 0.01$), but no association was found between subjects’ post-test sequential concepts and the use of evidence ($F = 1.97$, $p > 0.05$). However, interestingly, the male groups’ responses showed a statistically significant association between the number of network linkages in the flow maps and the respondents recognition of the importance of evidence in judging claims made by the residents ($F = 5.65$, $p < 0.05$).

When students were asked what information they would need to be able to make better judgements, over 65% of the students were able to point out at least one statement that was deemed to be scientific. Female students, moreover, referred to scientific information more than males. The average number of scientific statements cited by female students was 1.78, while it was 0.84 for males ($F = 11.36$, $p < 0.01$). Based on a correlation analysis, the more students know about the causes of ground subsidence, the better they are able to quote what specific scientific information they would need to make critical judgements. Nevertheless, the correlation was not strong ($r = 0.23$, $p < 0.05$). Further analysis showed that such tendency was more apparent in the female group ($r = 0.34$, $p < 0.05$). However, the male respondents exhibited a relative strong correlation between the number of network linkages in their flow map narrative and their ability to identify scientific information required to make critical judgements ($r = 0.38$, $p < 0.05$).

**Table 4. Percentages of female and male students who identified evidence in making judgements**

<table>
<thead>
<tr>
<th>Sex</th>
<th>Evidence identified</th>
<th>Evidence not identified</th>
</tr>
</thead>
<tbody>
<tr>
<td>Female (valid $n = 40$)</td>
<td>55% ($n = 22$)</td>
<td>45% ($n = 18$)</td>
</tr>
<tr>
<td>Male (valid $n = 44$)</td>
<td>59% ($n = 26$)</td>
<td>41% ($n = 18$)</td>
</tr>
</tbody>
</table>
Although 65% of students were able to mention at least one scientific statement in response to the questionnaire items, no subjects recommended that information should be sought outside of the scientific domain, such as examining social concerns in making judgements about the adequacy of the arguments presented by the residents. For those who did not point out scientific information, they were either sure about their judgement or stressed that they did not care for further information since they did not possess sufficient knowledge to make judgements.

**Self-reflections**

The analysis on students’ self-reflections showed that most subjects were not confident about their thoughts, and the majority tended to attribute their uncertainty to insufficient information provided to them and insufficient knowledge about the scientific issues. Table 5 presents evidence of students’ self-judgements about whether they were sure about their own answers.

As presented in table 5, subjects seemed to believe that ‘correct’ knowledge can provide answers to all problems. Also, male students seemed to be more confident about their thoughts.

**Discussion**

**Concept performance and reasoning modes**

This study analysed students’ reasoning behaviour regarding the making of personal theory, and use of theory and evidence in the context of the underground water issue. The basic research strategy was to present a course first that provided basic scientific knowledge and then tested subjects’ reasoning behaviours targeted in the study. The flow map method showed that students’ sequential concept organization improved significantly after the STS instruction. And, the analysis of effect size on the linear concepts indicated a strong effect from the designed instruction. However, as far as the complex linkage is concerned, students’ performance indicated a low to

<table>
<thead>
<tr>
<th>Answer type</th>
<th>Number of males (valid n = 39)</th>
<th>Number of females (valid n = 38)</th>
</tr>
</thead>
<tbody>
<tr>
<td>The information provided was insufficient or false</td>
<td>12</td>
<td>13</td>
</tr>
<tr>
<td>My knowledge or experience is insufficient</td>
<td>9</td>
<td>11</td>
</tr>
<tr>
<td>I am not an expert</td>
<td>7</td>
<td>2</td>
</tr>
<tr>
<td>There were alternative explanations to the situation</td>
<td>2</td>
<td>4</td>
</tr>
<tr>
<td>I was not the resident so could not feel how they felt</td>
<td>0</td>
<td>4</td>
</tr>
<tr>
<td>I am confident with my own thought</td>
<td>8</td>
<td>3</td>
</tr>
<tr>
<td>There is no right or wrong answer to the issue</td>
<td>2</td>
<td>1</td>
</tr>
</tbody>
</table>
median effect from the instruction. This result reveals the difficulty of constructing complex knowledge structures among students.

It was found that most students were able to formulate and use personal theory in reasoning. Male students were more successful than female students in using personal theory. However, regardless of gender, subjects seemed to have difficulty in incorporating theory and evidence. In other words, they did not have an adequate sense that, in terms of scientific reasoning, a theory or hypothesis needs to be supported by evidence. Such a finding suggests that high school students do not have proper understanding about the role of theory in scientific investigations and the ways evidence is used to test theories and the hypotheses derived from them. This outcome is consistent with Kuhn's (1991, 1995) findings who showed that young adults have limited ability to make scientific arguments.

The results suggest that, although this group of students exhibited some evidence of the rudiments of scientific thinking, overall most students lacked a coherent sense of the importance of evidence in evaluating claims and seemed not to have a thoroughly rationalized stance on the role of theory and hypotheses in scientific analyses. Thus, within this context, it appears students in this age range, and given the opportunities presented here to learn about scientific thinking, may not possess strong scientific reasoning abilities. This finding is contrary to the conclusions of some researchers who claim even young children can think as scientists do (for example, Samarapungavan 1992, Vosniadou and Brewer 1992). However, it is important to clearly define terminology when making claims about scientific thinking. In some cases, the indicator used for scientific thinking is quite different than used in this study. Moreover, it is important to determine the situation established for the problem-solving and the knowledge context in which the skills are elicited. Students may be indeed equipped with scientific reasoning skills learned within a domain-specific knowledge framework, but throughout their learning experiences they may have had little opportunities to practice such an ability in non-scientific contexts such as the social and environmental contexts used in this study. As a consequence, the power of scientific thought that a student may bring to a problem in laboratory science, for example, may not be mobilized in another less scientific situation. In other words, these students might have difficulty of skill transfer across contexts. One must also be aware that different rates of cognitive development may account for some differences among students and that special care needs to be taken to ensure, if at all possible, that the generality of scientific thinking skills is adequately planned during learning in school settings. On the other hand, young adults’ difficulty in relating theory and evidence further leads to a speculation that the development of scientific reasoning, however potentiated at an early age, may reach a plateau in early adulthood, especially if no further opportunities are given to enhance the development of this kind of reasoning skill. This could be the consequence of usual classroom instructional practices that emphasize presenting final end products of scientific knowledge without stressing the process of how the ends are achieved (Duschl 1990).

A weak association was found between students’ knowledge about the causes of ground subsidence and their ability to identify the kind of scientific information they need to make better judgements. This suggests that some forms of classroom-learned knowledge might not be applied when situations require it. A major problem of traditional classroom settings is the lack of correspondence between curriculum experiences and the demands of daily life. Although the instructional presentation
used in this investigation was STS oriented, the short period of training might not be sufficient to make a significant difference in effect on the outcome variables examined here. However, the incorporation of STS-oriented instruction in the classrooms is promising for the purpose of enhancing scientific thinking since a moderate association between students’ existing concepts and ability to use relevant concepts was found in the female group after the instruction.

When asked to reflect on their thoughts, students attributed their uncertainty to insufficient background information and insufficient knowledge. In addition, some students, especially males, tend to think that the word of experts guarantees its validity. Only a few subjects stated that there might be alternatives, or there is no absolute right or wrong. It seems that most students believe knowledge can provide certain answers to all problems. From an epistemological point of view, such a view toward knowledge indicates the students have not developed a constructive or relative epistemology. In terms of the ‘Perry Scheme’ (Perry 1970), they are in an early stage of ‘multiplicity’ where thinkers still believe all things are knowable, but start to realize some knowledge domains are fuzzy. This may add some explanation as to why students in the study have difficulty incorporating theory and evidence, and why the association between students’ existing knowledge and their ability to identify evidence is not strong. For most participants, since all is knowable, information provided by the water company should be right or wrong. And, it is experts who can tell the correctness of information. In this case, the important role of evidence could be easily ignored.

About gender differences

Further analysis showed several apparent gender differences. As presented in table 2, female students’ pre-test score in terms of the total number of statements enumerated in the flow map analysis was significantly lower than that of males. In other words, prior to the treatment instruction, female students possessed less scientific information about underground water. A previous study (Yang and Anderson 2003) also showed that female subjects possessed less prior knowledge about nuclear energy than males. Seemingly, in terms of socio-scientific issues, male students had more access to scientific information. Is such a phenomenon a natural consequence due to the apparent lack of interest by females in scientific information or a consequence of unequal learning opportunity? More investigation is needed to answer this question.

Although female students did not possess as much prior relevant knowledge as the males, they learned as much as the males following the initial instruction. Moreover, as displayed in the previous section, female students showed better ability in referring scientific information for making judgements. In a previous study, Yang and Anderson (2003) found that after providing relevant nuclear knowledge, female students were as willing as males to consider scientific information in making informed judgements about environmental issues. These findings suggest that gender differences in reasoning found in the current study could have resulted from differences in thinkers’ existing domain-specific knowledge. It is probable that the reason females are usually thought of as more ‘emotional’ or ‘unscientific’ when confronted with critical socio-scientific issues is because they are not exposed enough to the scientific aspects of the issues. However, it also should be noted that the unequal performance and differences in knowledge background may be due to
females’ approaches to information processing related to their epistemological developmental status (Belenky et al. 1986).

Another interesting finding is that the male students in the current study had statistically significantly more network linkages in their post-test narrative compared with the females. In other words, female students made more progress in acquiring new information and presenting it in a linear sequence, while males developed more cross-linkages among ideas within their sequence of statements. Moreover, interestingly, for the male students, it was the number of network linkages that was somewhat correlated with performance of pinpointing scientific information. Since the greater the number of network linkages observed in a communication may indicate better development of multirelational information structures in memory (Anderson and Demetrius 1993), it could be that male students’ reasoning depends more on mobilizing networks of ideas in a multi-relational perspective, while females depend more on mobilizing a linear or sequential form of information. However, this is only a conjecture based on fragmentary evidence from this study, and more research focused on this specific issue is needed.

As far as self-reflection is concerned, there is not much difference between males and females (table 5). However, when examining individual performances, it was found that males seemingly displayed higher confidence and exhibited greater dependence on experts to provide ‘correct’ information. Females, on the other hand, would consider the residents’ feelings as being more important, although the difference appears not to be highly significant. In terms of epistemological development, Perry (1970) did not find gender differences. However, Belenky et al. (1986) proposed that females tend to initiate reasoning more from personally relevant experiences, and, they also care more about other people’s feelings. The theoretical arguments that Belenky and colleagues make may explain our research outcome as well. Overall, students are in the same epistemological development stage but female students may have different ways of approaching information or may possess different forms of knowledge structures compared to many males.

Reflections on the research methods

This study used the flow map method to access students’ concept structure. The number of accurate statements increased in the students’ flow maps after the STS-oriented instruction. However, the number of network linkages (i.e. cross-linkages between concepts) was not particularly impressive based on evidence from changes induced by instruction in other studies (Anderson and Demetrius 1993, Anderson et al. 2001). Anderson et al. (2001) had reported that the complex network linkages predict academic performance and are positively correlated with level of cognitive thinking. The inherent knowledge organization of the topics may have had an effect on the student’s use of cross-linking references in their narrative. Previous studies used the flow map method to assess thinker’s domain-specific knowledge structure but in this study more then one knowledge domain was examined. Instead of essays, as used in Anderson’s study, this study employed open-ended questions placed in the questionnaire to probe students’ memory. The posited questions, which mainly asked for explanations on the causes of land subsidence, could have constrained the subjects’ expressions of view. Moreover, as far as the causes of land subsidence are concerned, factors such as excessive use of underground water and earthquakes (or
other types of plate movements) have obviously different mechanisms. Accordingly, cross-linkages between concepts are expected to be rare.

On the other hand, a previous study using the same questioning technique showed that network cross-linkages do not usually appear until the number of discourse statements have accrued to a certain amount (Yang 2002). In this study, it was also found that respondents who displayed network linkages in their post-test flow maps had at least six linearly linked statements where, as table 2 shows, the maximum sequential linkage is 11. Hence, it is probable that the students’ sequential concept structure implied by the amount of linear linkages was still too short to generate rich network linkages. However, it also should be pointed out that the STS-oriented lesson presented to the students at the beginning of the study did not emphasize linking of concepts. So there was no explicit attempt to influence the students’ ideational network of ideas in this treatment communication.

This study used a questionnaire instead of interviews to explore reasoning. This research technique is limited in the inferential power in that follow-up questions could not be given individually to probe in-depth the individual thoughts. However, with the method, the study was able to present and interpret data in both qualitative and quantitative ways. The gender differences found in the study, which in some cases were not statistically meaningful enough, could have been due to the limitation of the data collection technique. As listed in the literature section, theories about gender difference were mostly constructed based on intensive interviews. Even though, it is still encouraging that a theoretically matched trend was found within a small amount of samples. This research finding has inspired a further large-scale investigation.

Conclusion

The research finding that high school students do not have adequate hypothesis testing ability, although consistent with some psychological studies, is contradictory to the findings of other studies indicating that even young children reason like scientists. The two apparently contradictory findings may be resolved by considering the role of context as a possible explanatory variable. It is well known that skill transfer is not easy across different contexts even in one single knowledge domain, not to mention transfer from a scientific domain to a life context that contains a variety of domains (including personal and social dimensions) that transcend those found in the natural sciences. Besides, as far as social issues are concerned, personal value systems must have played a determining role in the decision-making process, which would definitely add more complexity to the skill transfer process. Hence, further research concerning skill transfer from one scientific domain to relevant social domains should receive more attention.

Implication of the study

Students in the study showed even though they made progresses in knowledge achievement relative to the basic scientific information, their ability to relate theory and evidence in reasoning was unsatisfactory. Boys displayed better ability in using theories while girls performed better in referring scientific information for making judgements. Males and females also seemed to have different information approaches. Overall, the students’ reasoning behaviour was found consistent with
their personal epistemological perspectives. Accordingly, as far as the enhancement of scientific thinking ability is concerned, other than emphasizing learning in authentic environments as for learners to see where, when and how relevant skills could be used, additional attention should be paid to the personal epistemological belief in relation to the level of maturation of the learners. This idea has been actually advocated by some renowned scholars (for example, Hofer 2001, Kuhn 1999), but has not yet received wide attention.

As for the situated instructional design, gender differences in information approaches should be taken into consideration. This study has specifically indicated that female students need to be exposed more to the scientific aspects of to-be-discussed issues and encouraged to use scientific information in reasoning. In addition, a consideration should be noted on the knowledge structures exhibited by students of different genders. It was showed that male and female students seem to reason differently depending on cross-linked and sequential knowledge structures, respectively. Accordingly, information presented in the classroom may need to adapt different orders to satisfy the difference. For example, for male students, relations between a few sources of information might be better discussed before all relevant information is presented, while for females it might be more effective to present all relevant information before discussing the relations. Moreover, given the preferential orientation of females to consider socio-emotional factors, the instructional approach may start from personal experiences or include an analysis of the subjective aspects and personal feelings of the individuals involved in a given situation.

Finally, as already mentioned, more information is in need about the transfer of thinking skills between specific and general knowledge domains in order to clarify the apparent discrepancy between competence in domain-specific problem-solving and performance in broader domains of daily problem-solving where modern citizens are expected to be equally competent. Studies of the kind may also consider the effects of personal epistemological beliefs, and the gender difference on the knowledge structure for this study suggests these variables may play a fundamental role in determining the success of skill transfer.

References


USE OF THEORY AND EVIDENCE


Appendix 1: examples of the concept flow map

Mapping procedure

Students were asked to respond in writing to the following questions and then researchers constructed their concept flow maps.

Questions

What concepts can you think of regarding the causes of land subsidence?

Please explain more of each concept you mentioned above.

Note

Concepts in the first level of each map are extracted from students' responses to question 1, while the rests of concepts from responses to question 2. Straight arrow lines indicate the linear linkage, and curved, dashed arrow lines indicate the complex linkage.

Example 1

(before instruction)

| Excessive use of underground water | Whether change |
| Space appears in soils | reduces rain falls |
| Source not filling not on time | Source not filling not on time |
| Support for the ground is reduced |

(six sequential links and one cross-link)

(after instruction)

| Excessive use of underground water |
| Underground water supports ground |
| Without water, space in the soil increases |
| Ground subsides to fill up the space |

(eight sequential links and two cross-links)
Example 2

**Before instruction**

- Excessive use of underground water
  - Water in aquifer dried out
  - Space appears in soil
  - The land receive too much overlying weight

*(three sequential links and zero cross-links)*

**After instruction**

- Excessive use of underground water
  - Water is reserved in aquifer
  - There is certain pressure
    - To prevent land subsidence
    - Lack of water causes unbalanced pressure
    - The soil is loosen
  - Crustal movement
    - Plate movement
    - results in faults

*(eight sequential links and one cross-link)*
Annotations from tsed100744.pdf

Page 2

Annotation 1; Label: Author Query; Date: 9/6/2004 12:06:57 PM
AQ1 – p.2
Text states Salmon and reference list states Salmon – which spelling is correct?

Annotation 2; Label: Author Query; Date: 9/6/2004 12:07:24 PM
AQ2 – p.2
Confirm Lawson et al. (2000) is correct citation

Annotation 3; Label: Author Query; Date: 9/6/2004 12:08:03 PM
AQ3 – p.2
Confirm Kuhn et al. (1988) is correct citation

Annotation 4; Label: Author Query; Date: 9/6/2004 12:08:31 PM
AQ2 – p.2
Confirm Lawson et al. (2000) is correct citation

Annotation 5; Label: Author Query; Date: 9/6/2004 12:10:01 PM
AQ4 – p.2
Text states Carey and reference list states Carry – which spelling is correct?

Page 6

Annotation 1; Label: Author Query; Date: 9/6/2004 12:11:54 PM
AQ5 – p.6
Provide full details of Cohen (1996) for reference list

Page 8

Annotation 1; Label: Author Query; Date: 9/6/2004 12:17:13 PM
AQ5 – p.8
Provide full details of Cohen (1996) for reference list

Page 16

Annotation 1; Label: Author Query; Date: 9/6/2004 12:14:12 PM
AQ4 – p.16
Text states Carey and reference list states Carry – which spelling is correct?

Page 17

Annotation 1; Label: Author Query; Date: 9/6/2004 12:15:45 PM
AQ1 – p.17
Text states Salmon and reference list states Salmon – which spelling is correct?
# PROOF CORRECTION MARKS

<table>
<thead>
<tr>
<th>Instruction</th>
<th>Mark in text</th>
<th>Mark in margin</th>
</tr>
</thead>
<tbody>
<tr>
<td>Leave unchanged</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Extraneous marks or damaged letter</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Delete</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Insert or replace text</td>
<td></td>
<td>text to be added</td>
</tr>
<tr>
<td>Add ( / ) or substitute ( / ) : full stop</td>
<td></td>
<td></td>
</tr>
<tr>
<td>decimal point</td>
<td></td>
<td></td>
</tr>
<tr>
<td>comma</td>
<td></td>
<td></td>
</tr>
<tr>
<td>semi-colon</td>
<td></td>
<td></td>
</tr>
<tr>
<td>colon</td>
<td></td>
<td></td>
</tr>
<tr>
<td>apostrophe or quotation mark</td>
<td></td>
<td></td>
</tr>
<tr>
<td>superscript</td>
<td></td>
<td></td>
</tr>
<tr>
<td>subscript</td>
<td></td>
<td></td>
</tr>
<tr>
<td>hyphen</td>
<td></td>
<td></td>
</tr>
<tr>
<td>short or long rule</td>
<td></td>
<td></td>
</tr>
<tr>
<td>oblique</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Wrong typeface or size</td>
<td>encircle</td>
<td></td>
</tr>
<tr>
<td>Change to: roman (upright)</td>
<td>encircle</td>
<td></td>
</tr>
<tr>
<td>italic</td>
<td>underline</td>
<td></td>
</tr>
<tr>
<td>capital letters</td>
<td>underline three times</td>
<td></td>
</tr>
<tr>
<td>small capitals</td>
<td>underline twice</td>
<td></td>
</tr>
<tr>
<td>bold type</td>
<td>wavy underline</td>
<td></td>
</tr>
<tr>
<td>lower case letters</td>
<td>encircle</td>
<td></td>
</tr>
<tr>
<td>Greek letters</td>
<td>encircle</td>
<td></td>
</tr>
<tr>
<td>Delete and close up</td>
<td>copy</td>
<td></td>
</tr>
<tr>
<td>Reduce space</td>
<td>copy</td>
<td></td>
</tr>
<tr>
<td>Close up space</td>
<td>copy</td>
<td></td>
</tr>
<tr>
<td>Insert space</td>
<td>copy</td>
<td></td>
</tr>
<tr>
<td>Make space in line equal</td>
<td>copy</td>
<td></td>
</tr>
<tr>
<td>Insert space between lines</td>
<td>copy</td>
<td></td>
</tr>
<tr>
<td>Reduce space between lines</td>
<td>copy</td>
<td></td>
</tr>
<tr>
<td>New paragraph</td>
<td>copy</td>
<td></td>
</tr>
<tr>
<td>Run on, no new paragraph</td>
<td>copy</td>
<td></td>
</tr>
<tr>
<td>Transpose letters or words</td>
<td>copy</td>
<td></td>
</tr>
<tr>
<td>Transpose lines</td>
<td>copy</td>
<td></td>
</tr>
<tr>
<td>Take character to next line</td>
<td>copy</td>
<td></td>
</tr>
<tr>
<td>Take character to previous line</td>
<td>copy</td>
<td></td>
</tr>
<tr>
<td>Raise text on page</td>
<td>copy</td>
<td></td>
</tr>
<tr>
<td>Lower text on page</td>
<td>copy</td>
<td></td>
</tr>
<tr>
<td>Check vertical alignment</td>
<td>copy</td>
<td></td>
</tr>
<tr>
<td>Check horizontal alignment</td>
<td>copy</td>
<td></td>
</tr>
</tbody>
</table>
AN AGREEMENT FOR THE TRANSFER OF COPYRIGHT

IN RELATION TO THE CONTRIBUTION OF YOUR ARTICLE (‘THE ARTICLE’) ENTITLED:

................................................................................................................................................

BY:........................................................................................................................................

WHICH WILL BE PUBLISHED IN T-SED

................................................................................................................................................

In order to ensure both the widest dissemination and protection of material published in our Journal, we ask authors to assign the rights of copyright in the articles they contribute. This enables Taylor & Francis Ltd (‘us’ or ‘we’) to ensure protection against infringement. In consideration of the publication of your Article, you agree to the following:

1. You assign to us with full title guarantee all rights of copyright and related rights in your Article. So that there is no doubt, this assignment includes the assignment of the right to publish the Article in all forms, including electronic and digital forms, for the full legal term of the copyright and any extension or renewals. Electronic form shall include, but not be limited to, microfiche, CD-ROM and in a form accessible via on-line electronic networks. You shall retain the right to use the substance of the above work in future works, including lectures, press releases and reviews, provided that you acknowledge its prior publication in the journal.

2. We shall prepare and publish your Article in the Journal. We reserve the right to make such editorial changes as may be necessary to make the Article suitable for publication, or as we reasonably consider necessary to avoid infringing third party rights or law; and we reserve the right not to proceed with publication for whatever reason.

3. You hereby assert your moral rights to be identified as the author of the Article according to the UK Copyright Designs & Patents Act 1988.

4. You warrant that you have at your expense secured the necessary written permission from the appropriate copyright owner or authorities for the reproduction in the Article and the Journal of any text, illustration, or other material. You warrant that, apart from any such third party copyright material included in the Article, the Article is your original work, and does not infringe the intellectual property rights of any other person or entity and cannot be construed as plagiarising any other published work. You further warrant that the Article has not been previously assigned or licensed by you to any third party and you will undertake that it will not be published elsewhere without our written consent.

5. In addition you warrant that the Article contains no statement that is abusive, defamatory, libelous, obscene, fraudulent, nor in any way infringes the rights of others, nor is in any other way unlawful or in violation of applicable laws.

6. You warrant that wherever possible and appropriate, any patient, client or participant in any research or clinical experiment or study who is mentioned in the Article has given consent to the inclusion of material pertaining to themselves, and that they acknowledge that they cannot be identified via the Article and that you will not identify them in any way.

7. You warrant that you shall include in the text of the Article appropriate warnings concerning any particular hazards that may be involved in carrying out experiments or procedures described in the Article or involved in instructions, materials, or formulae in the Article, and shall mention explicitly relevant safety precautions, and give, if an accepted code of practice is relevant, a reference to the relevant standard or code.

8. You shall keep us and our affiliates indemnified in full against all loss, damages, injury, costs and expenses (including legal and other professional fees and expenses) awarded against or incurred or paid by us as a result of your breach of the warranties given in this agreement.
9. You undertake that you will include in the text of the Article an appropriate statement should you have a financial interest or benefit arising from the direct applications of your research.

10. If the Article was prepared jointly with other authors, you warrant that you have been authorised by all co-authors to sign this Agreement on their behalf, and to agree on their behalf the order of names in the publication of the Article. You shall notify us in writing of the names of any such co-owners.

11. This agreement (and any dispute, proceeding, claim or controversy in relation to it) is subject to English law and the jurisdiction of the Courts of England and Wales. It may only be amended by a document signed by both of us.

Signed

Print name

Date

_______________________________