Investigations

in

Mathematics Learning

Winter Edition 2008 *Volume 1, Number 2*

Supported by State University of New York at Binghamton

Table of Contents

Preservice Elementary Teachers' Understanding of Logical Inference **1**

Shandy Hauk, University of Northern Colorado
April Brown Judd, University of Northern Colorado
Jenq Jong Tsay, University of Northern Colorado
Harel Barzilai, Salisbury University
Homer Austin, Salisbury University

Performance of Middle Grades Students on Quantitative Tasks: A Beginning Dialogue on Quantitative Literacy in Middle Schools **35**

David K. Pugalee, University of North Carolina Charlotte
Kimberly J. Hartman, University of North Carolina Charlotte
Jennifer H. Forrester, North Carolina State University

Relationship between Students' Understanding of Functions in Cartesian and Polar Coordinate Systems **52**

Mariana Montiel, Georgia State University
Draga Vidakovic, Georgia State University
Tangul Kabael, Anadolu University

Investigations in Mathematics Learning
©The Research Council on Mathematics Learning
Winter Edition 2008, Volume 1, Number 2

Preservice Elementary Teachers' Understanding of Logical Inference

Shandy Hauk, April Brown Judd, Jenq Jong Tsay
University of Northern Colorado

Harel Barzilai, Homer Austin
Salisbury University

Abstract

This article reports on the logical reasoning efforts of five prospective elementary school teachers as they responded to interview prompts involving nonsense, natural, and mathematical representations of conditional statements. The interview participants evinced various levels of reliance on personal relevance, linguistic contextualization, and time-dependent interpretation in working through reasoning tasks. Different kinds of affective and cognitive demands, dependent on personal history, may be needed for the depersonalization, decontextualization, and detemporalization required by abstract logico-deductive reasoning. Implications for college instruction with future elementary school teachers include suggestions for logical argument analysis activities aimed at enriching learners' reasoning situation images.

Encouraging students to reason logically throughout their mathematics education helps them build the understanding that mathematics makes sense. According to the *Principles and Standards for School Mathematics* (National Council of Teachers of Mathematics [NCTM], 2000), proof and reasoning should be incorporated regularly into the mathematics classroom from pre-kindergarten through grade twelve. In particular, "[b]y the end of secondary school, students should be able to understand and produce mathematical proofs - arguments consisting of logically rigorous deduction of conclusions from hypotheses" (p. 55). Consequently, for every teacher, the ability to explain in a convincing way why a mathematical proof (formal or informal) is true and valid is a valuable tool. It is also a difficult teaching skill to develop and maintain in the face of the disparate and immediate needs of students in the classroom (Durand-Guerrier, 2003; Simon, 2000).

Research over the past 75 years has indicated that understanding proof, particularly logical inference, is a challenge to students of all ages, including preservice teachers (Bell, 1976; Healy & Hoyles, 2000; Selden & Selden, 2003; Wilkins, 1928). In particular, studies of college students' efforts with logical inference and conditional reasoning in the 1970s reported that undergraduates in general, and preservice elementary teachers in particular, did not reliably interpret syllogistic or disjunctive logical statements presented in natural language form (Jansson, 1975; Eisenberg & McGinty, 1974). In her work with prospective elementary teachers, Damarin (1977a, 1977b) found learners had a tendency to treat conjunctive, conditional, and biconditional statements that were presented in abstract, visual, mathematical form in the same way: students approached all of the tasks with the logic rules associated with conjunction, declaring a compound statement true only if all parts were true. Along similar lines, Vest (1981) noted that college undergraduates did not have a robust understanding of disjunction and conjunction in a natural language setting. In fact, the comprehension of Vest's participants closely resembled that of the preservice teachers in Damarin's studies despite the difference in natural versus mathematical language contexts. In part, the work reported here reproduces this result 30 years later.

Austin (1984) reported on the interpretations of logical implication offered by a broad cross-section of undergraduates. Specifically, he examined four reasoning patterns: detachment, conversion, inversion, and contraposition. Detachment (*Modus ponens)* is where one concludes Y if both the implicative statement D ⇒Y and its antecedent X are assumed true; conversion is the pattern whereby one realizes that X cannot necessarily be concluded when both X ⇒ Y and Y are assumed; inversion involves recognizing that the negation of Y, "*not*Y" cannot necessarily be concluded if X ⇒ Y and *not*X are assumed; contraposition (*Modus Tollens*) is the pattern whereby one concludes *not*X is true from the assumptions that X ⇒ Y and *not*Y both hold. The structure of modus tollens is actually related to but not identical to the logical construct of the contrapositive *not*Y ⇒ *not*X of a conditional statement X ⇒ Y; in Austin (1984) the pattern was named "contraposition." Although students from a random sample (n=219) could use detachment reasoning fairly well (73% correct responses), they had difficulty with conversion (57% correct), contraposition (47% correct), and inversions (51% correct). Students could reason if the conditional was given in the familiar "forward" form. However, many had difficulty reasoning about implications when given variants of the standard conditional form. Austin concluded that this conflation of a conditional statement with its variants made mental processing of theorems in mathematics difficult and that, as a result, comprehension and application of theorems was a challenge for many students.

It has been suggested by a variety of researchers that learners go through four stages in developing mathematical logical reasoning skill. For example, in terms of the four categories of Action-Process-Object-Schema (APOS) theory (Asiala, et al., 1996), construction of understanding begins (a) formalized actions without a great deal of understanding, followed by (b) some structuring of logical inference into process which, when repeated, can lead to (c) a learner observing parallels, noticing properties (like the relation between converse and inverse), and encapsulating inferential processes like "X ⇒ Y" into a new kind of statement conceived of as an object, "S: X ⇒ Y." this now reified object can, in turn, be the subject of new actions and processes (Sfard, 1991). For example, a student will be able to conceive of transforming the statement S into a new object, its contrapositive: "*not*X ⇒ *not*Y." From actions, processes, and objects arises the most complex form, (d) schema, a mental structure in which the processes and objects of inferential reasoning are connected to other understandings about implication, contradiction, and proving.

Balacheff's (1988) levels of proof understanding are another example of a four-stage model. These four realms of proof understanding are (a) native empiricism, characterized by "proof by example" strategies; (b) crucial experiment, including the generation of counter-examples; (c) generic examples; and (d) thought experiment, where one abstracts inductive empirical approaches to arrive at understanding of highly structured deductive logical forms. Moreover, Balacheff (1988) suggested, "Language must become a tool for logical deductions and not just a means of communication." He contended that the use of language in a mathematically deductive way required forms of decontextualization, depersonalization, and detemporalization (more on this below).

The work of Balacheff (1988) and others has suggested that as students build understanding of logico-deductive reasoning they move from everyday-language-based efforts to the use of conditional implication. This begins by recognizing that certain forms and rules exist and subsequently includes an understanding that certain mathematical constructs can be thought of as higher level or more abstract objects. Ultimately, learners come to the mastery of logically consistent and valid transformations of logical objects (e.g., conditional statements, quantified statements).

Gila Hanna has pointed out, in several contexts (1989, 1995, 2000), that a proof that convincingly explains is not necessarily the same as a proof that proves. It has been well documented that logically valid (and invalid) conclusions consistent with one's experiences and beliefs about life are more convincing and more frequently accepted as valid than unbelievable or nonsensical conclusions (Thompson, 1996). This is referred to as "belief-biased" and a great deal of work has been done in cognitive science and psychology to create a theory explaining it (Markovits & Nantel, 1989; Oakhill & Johnson-Laird, 1985; Torrens, Thompson, & Cramer, 1999). Neuro-imaging research on brain activity during reasoning tasks has indicated that there are distinct differences in the constellations of neural firings during logical reasoning dependent on the belief conditions of the task (Goel & Dolan, 2003). In Goel & Dolan's study, familiar natural-language-based syllogistic reasoning tasks activated areas of the brain associated with both semantic retrieval and information selection. This engagement was independent of the validity of the syllogism or the truth of its consequent. An example from the current study of a natural-language-based task would be: If it is raining, then Tom wears a red shirt. If Tom wears a red shirt, then Susan bakes a cake. It is raining. Does Susan bake a cake? (Yes) (No) (Not Necessarily). During logically equivalent decontextualized ("belief-neutral") reasoning, there was additional activity in areas of the brain associated with abstraction, numerical estimation, and with manipulation of spatial information. An example from the current study, in mathematical language, would be: Suppose that X implies Y. Suppose also that Y implies Z. Does X imply Z? (Yes) (No) (Not Necessarily), or, in nonsense (context-free) language: Exabiffs which trundle herbariously do prevankerize lurgidly. Those who prevankerize lurgidly always groop their foonting turlingdromes. Do Exabiff which trundle herbariously groop their foonting turlingdromes? (Yes) (No) (Not Necessarily). Goel & Dolan also reported that during tasks which were "inhibitory" (e.g., a correct syllogistic form with false conclusion or an invalid form with a conclusion that happened to be true), participants who correctly completed the tasks appeared to detect and compensate for the conflict between their beliefs and the logical inference: suppressing belief-biased response to engage right parietal reasoning activity. When a participant incorrectly completed the task, such right parietal engagement was absent. Instead, part of the brain associated with emotion was active. In other words, when presented with conflict between logical form and conclusion truth-value, participants either shut down affective response to focus on logic, or they shut down logical response and went with their feelings.

The process of deciding that a valid conclusion must follow necessarily (not just possibly) from its premises is an appeal to "logical necessity." Research on *logical necessity* and *belief bias* are especially pertinent when investigating the transition from "child logic" (conflating a statement with its converse) to the facility with logic required for teaching. In particular, in mathematical reasoning one does not accept a conclusion, however believable it may seem, unless it necessarily follows from its premises. An important facet of the research on belief-bias, for the present discussion, is the absence of agreement on a theory that can consistently explain the many approaches to logical implication demonstrated by human beings. It may be that such a theory, if it exists, is not any of the single theories currently in use (e.g., Oakhill & Johnson-Laird's (1985) mental models or Rips' (1994) rule-based reasoning). Such a conjecture is supported by results from Klauer, Musch, & Naumer (2000) indicating that no single theory parsimoniously predicts outcomes. Klauer et al. called for qualitative work investigating the "talk-aloud" reasoning of people validating arguments in and out of familiar contexts.

The work presented here addresses this call to action. The core of the research reported below was an attempt to provide a theory, grounded in the literature and informed by five "talk-aloud" reasoning interviews, for coming to understand students' conceptualizing of logical inference. Like Hoyles and Küchemann (2002), the focus was on how students "learn to move between mathematical ways of proving and those that are rooted in everyday thinking." After presenting the methods and results of talk-aloud interviews, the discussion section offers connections between empiricale and theoretical results on reasoning, logical necessity, and belief bias. A conjecture is advanced and supported that due to individual variation in affective and cognitive connections within Balacheff's (1988) decontextualization, depersonalization, and detemporalization, not all three are necessary for logical reasoning for all people. Learners may make choices, implicitly or explicitly, about which of the three to engage when the encounter conflicts in reasoning tasks. The conclusion frames a theory for student's logical strategy use. The aim is not to identify *stages* in student development of logical reasoning. Rather, the goal is to describe constructive processes going on *within* the stages, whether or not the learner is fluidly articulate about their thinking. The report ends by addressing the implications of the presented theory for pre- and in-service teacher preparation in the contexts of collegiate mathematics education and professional development.

Method

The philosophical underpinnings of the study were constructivist: individuals construct their own understanding of concepts. Moreover, one way to bring to the surface observable artifacts or someone's constructed understanding is to create a cognitive conflict, what has been called a disruption to equilibrium (Inhelder & Piaget, 1958), and investigate how the individual resolves the conflict.

The study was a naturalistic inquiry through qualitative and quantitative data gathering and analysis. Some details of the quantitative, questionnaire-based, portion of the study are provided to give context for the interviews. However, the focus in this report is on the results of the qualitative analysis of interview data. Statistical analysis of written surveys is the subject of another report.

Research Teams

The study included collaboration between two research teams. Team 1 consisted of two mathematics professors at a public comprehensive university. This team designed the survey instrument and interview protocol, collected survey and interview data, and contributed mathematical expertise to the project. Team 2 consisted of the first author (a researcher in mathematics education) and two mathematics education Ph.D. students. This team was responsible for qualitative data analysis and contributed grounded theory expertise to the study.

Questionnaire Instrument

A 37-item questionnaire was administered in the first and last weeks of two sections of a Mathematics for Elementary Teachers course taught by a member of Team 1 in Spring 2001. The instrument consisted of five sections (see Appendix). The first section collected demographic information and asked students to respond to two prompts about their mathematical self-perceptions. The next three sections of the instrument were made up of 30 logical inference items. Section 2 was made up of Items 1 through 10 and used *nonsense language.* Section 3, Items 11 through 20, used everyday conversational English or *natural language.* Section 4 consisted of Items 21 through 30 and used symbolic *mathematical language.* Each of the ten abstract mathematical statements in this fourth section had a logically equivalent partner in each of Sections 2 and 3, though not in the same order. To illustrate the relationship among the sections, consider the following three logically equivalent items from the survey:

9. Whenever it’s a rainy day, florks phlapenaggle red shirts. Today, it is not raining. Are the glorks phlapenaggling red shirts? (Yes) (No) (Not necessarily)
12. If the man is friendly, then the woman is sad. The man is not friendly. Is the woman sad? (Yes) (No) (Not necessarily)
29. Suppose that X implies Y. Suppose X is false. Is Y false? (Yes) (No) (Not necessarily). [Note: In order for item 29 to be perfectly parallel to items 9 and 12, it would have been necessary to have it ask “Is Y true?” at the end rather than “Is Y false?” The wording used in item 29 was used for variety, to make the parallels among the three 10-question sections less obvious, and to allow later analysis of how extra “nots” might affect student’s answers; mathematically, a student’s answer to “Is Y true?” would uniquely define their answer to “Is Y false?”]

The final section of the survey consisted of four short answer questions (Items A, B, C, and D). Two of these items concerned logical reasoning and two were intended to gather reflections from students on the survey process and their thinking while completing the instrument. The item from this final section discussed most here was:

> *Item B*: Consider the statement "If glimmerles are flondish, then all kelevs dringle." Suppose we know that flimmerles are not flondish. Is the statement in quotes true? (Yes) (No) (Not Enough Information)

Item B is logically equivalent to: Consider the statement "X implies Y." Suppose X is false. Is "X implies Y" true?

Though no piloting of the instrument was conducted, the designers of the instrument agreed on its face and content validity. Student reports, during interviews, also supported its validity. In fact, because the purpose of the study was an investigation of the *absence* of consistent interpretations of logical inference across different contexts by the participants, reliable correlation among the sections of the instrument was not expected.

Interviews

Team 1 conducted interviews with five volunteer students from the surveyed sections of the course. Interviews were informal and took place in a windowless faculty office with nature posters on the walls, under a combination of incandescent lamps and fluorescent overhead lights. The interviews were recorded using a table-top cassette recorder, with built-in microphone, placed on the desk between the main interviewer and the participant.

Each open-ended interview had three parts. First, the participant was asked to review the survey and to make observations about the 30 items in Sections 2, 3, and 4. Second, the participant (who had been given her or his own completed survey) was prompted about whether he or she would change any of the responses. If so, those items were discussed and the interviewee's reasons for changing the answer explored. Finally, whether they suggested changing their answer or not, interviewees were asked to consider at least one of the logical reasoning items (from Sections 2, 3, 4, and the short-answer items in Section 5) and asked to discuss the reasoning that lead to their answer.

Tapes of the five interviews were transcribed by Team 2 and analyzed by them according to the constant-comparative grounded theory methods described by Stauss and Corbin (1998). All interviews were initially analyzed through open coding for themes common across interviews then reanalyzed and organized into categories through axial coding. In the final step, selective coding, the categorical structure resulting from axial coding was integrated into theory and interviews were re-examined. The outcome of selective coding led to the reported results. Colleagues and graduate student researchers provided subsequent theory checking and triangulation for coding.

Participants

Before moving to the Results section, we give a brief introduction to interviewees. It should be noted that self-selection by students willing to be interviewed for extra credit may have contributed to greater representation by students with lower grades and by mathematically gregarious students (see below). The Results section describes the nature and scope of participants' interview responses based on the central categories that emerged from interview data. The names of participants are pseudonyms. The participants are presented here in order from least to most able in the logical reasoning tasks.

Linda. A preservice teacher in her third year of college, she had the greatest difficulty with the reasoning tasks. Throughout her interview Linda made it clear that she had a "negative reaction" to mathematical language.

Amy. Also in her third year of college at the time of the study, she later graduated *magna cum laude* from the university with a degree in Elementary Education. Like Linda, Amy reported having some difficulties communicating mathematical concepts.

Ruby. A first-year student with no previous college mathematics courses, she was more articulate about her thinking and reasoning than Linda or Amy. In judging the validity of statements, Ruby relied mostly on empiricism related to "real life" situations she could imagine experiencing herself.

Margaret. A returning student in her third year of college, she was five years older than a traditional third-year undergraduate. Of the five respondents, Margaret appeared to have the greatest flexibility in connecting and moving between natural, nonsense, and mathematical language representations.

Jack. The only man among the interviewees, he was also a third-year student studying elementary education. Team 1 learned during the interview that Jack had taken a college-level logic course (none of the others interviewed had such a background); it was very rare for students in the first-semester mathematics course for elementary teachers to have had such prior experience. In addition, Jack was the only student to "strongly agree" with "I like math." His interview was the longest, at 65 minutes, in part because Jack was the only respondent to *initiate* discussion within new contexts in his attempts to explain his understanding of mathematical concepts.

Table 1 summarizes some demographic and college grade information along with the information these students provided on their questionnaire in responses to the two prompts:

Table 1.
Summary of Interviewee Information

Characteristic	Amy	Jack	Linda	Margaret	Ruby
Age	21	20	20	26	18
Year in College	3	3	2	3	1
Response to "I like math"	Disagree	Strongly Agree	Neutral	Agree	Disagree
Response to "I am good at math"	Disagree	Agree	Neutral	Agree	Neutral
Mathematics Grades Average	70%	95%	77%	80%	75%
Overall Grades Average	93%	98%	84%	90%	86%
Length of Interview	30 min.	65 min	25 min	22 min	20 min

For the statement "I am good at math", do you (circle one):
(Strong agree) (Agree) (Neutral) (Disagree) (Strongly Disagree)
For the statement "I like math", do you (circle one):
(Strongly agree) (Agree) (Neutral) (Disagree) (Strongly Disagree)

Results

Two main categories emerged from comparative analysis of the interview transcripts: contextualization and logical reasoning. *Contextualization* refers here to the attempts by respondents either to use mathematical language to communicate with the interviews or to recast a statement into a familiar context. *Logical reasoning* signifies participants' understanding of deductive interference and conditional reasoning; this category included two sub-categories: comparative-conflict and semantics. The sub-category *comparative-conflict* emerged from students' attempts to resolve interpretive conflicts when comparing logically equivalent items while the *semantics* sub-category concerns the relationships noted by students between logic trigger words, such as "if," "then," and "whatever," and the statements to which they referred.

Contextualization

Much of the interview discourse involved participants' efforts to contextualize the different types of language used on the survey instrument and used by the interviewers. Interviewees' attempts appeared to be aimed at three goals: to create meaning, to make decisions, and to formulate responses to interviewer comments and questions. When assigning truth-values to conditional if-then statements, participants clearly wrestled with their own efforts to contextualize the predictions within the if-then statements. This could be seen most clearly in the comparisons respondents made between the items in Sections 2, 3, and 4. The first ten-item section was named *nonsense language* because it included made-up words (e.g., "glorks" and "phlapenaggle" in survey Item 9). Items 11 through 20 were the *natural language* section because the prompts contained everyday language and commonplace nouns. In Items 21 through 30, the symbols X and Y represented antecedent and consequent. These ten items were the *mathematical language* section.

Interviewers asked participants to put the three sections in order from hardest to easiest and to talk about any relationships within or between the sections. Amy, Jack, Margaret, and Ruby (but not Linda) asserted that the first set, nonsense language, was most difficult and that the mathematics statements were more difficult than the natural language ones. Linda, however, said she felt the nonsense language section was easier than the mathematical language section: "Because when you're saying something like 'Suppose X implies Y, suppose Y is false, and is X true?' You know it's just the way it sounds. So it just gives you kind of like a negative reaction to the question."

Margaret pointed to the personal relevance to terminology as a deciding factor in explaining her choice of ordering from nonsense language (hardest), mathematical language, to natural language (easiest): "After I started reading through it [the natural language set], and came to the things that I knew about, I could picture the things that I didn't know about [in the set with nonsense words]." That is, during her written work on the questionnaire, she had moved between the natural and nonsense language sections, using her comfort with meaning in the one to help her decipher meaning in the other. The questions in the mathematical set, with letters X and Y, were more meaningful to Margaret than the nonsense language because she "could look at it and understand it," and she was "more familiar with them." Moreover, Margaret recalled that while completing the questionnaire she was able to re-write something presented in nonsense terminology using "an equation in terms of X and Y" so that she could decide a truth value for the statement (this use of "equation" will be revisited below, in discussing comparative-conflict).

Amy reported difficulties similar to Margaret's when dealing with the unfamiliar nonsense words. She said, "I still tried to picture them [nonsense words] but it was harder," and "you couldn't really picture them in your mind." Neither Ruby nor Linda saw any purpose in attempting to "picture" or "reason" about nonsense. Because the contexts of the survey and interview were mathematical (the interviewers were both mathematics professors and the survey had "X and Y stuff" on it), it may be that not having the definition of the nonsense terms in a mathematical context was enough for Ruby and Linda to dismiss the idea of reasoning about them. Alcock and Simpson (2002) noted that even though humans may tend to think in everyday terms about a concept using a prototypical example, mathematics calls for reasoning based on definitions of concepts.

Jack explicitly stated, "Visualizing makes the big difference" for the nonsense words because "it's like stuff - I mean make-believe words or stuff that you would find on like *Star Trek* [a science fiction television and movie series]." Consequently, the requirement that he create a fictional image – rather than accessing an existing one, like for "red shirt" in a natural language prompt – made it harder for him to "visualize" and harder to work with nonsense language statements. However, Jack also reported trying to do "mental mathematics" to recast into symbolic logic some of the statements with natural language or nonsense words. Recall that Jack was the one student who had taken a logic course in college. Additionally, Jack noted relationships between individual survey items. For example, he volunteered a symbolically based comparison between Item 30 and Item 11:

> Jack: I just went, I basically, when I was thinking about it I could see like suppose X implies Y, suppose that Y implies Z. Does X imply Z?... well, rain would be X, Y then for Tom wears a red shirt and Z would be Susan baking a cake...

Besides re-contextualizing natural language statements into mathematical statements with representative variables, Jack also tried to use symbols to substitute for nonsense words. For example, he said, "If I, in like, underline [Jack underlines statement on questionnaire] 'every hooloovoo is a snarkoid,' every *h* is a *s*...." Nonetheless, he did not think there was any merit in this kind of substitution. He believed that the terms in the statements should refer to something "for real" so that "people would be able to picture something in there." Jack did say that he knew that the context of prediction would not change the truth-value of a statement:

> I mean, it won't change the, I mean it, it's still all dependent on "if this, then this" ya know, "then you have this"...it's dependent on that it's still gonna have the same answer if...so long as those words [if...then...] stay the same.

When asked it they saw any relationships *among* the items in the three sections, participant's answers varied. To this question Margaret replied, "Actually, the first ten, it was about fiction...and then relations go to man for the next ten questions, and then for the last ten questions it's all mathematical." Margaret also mentioned that she realized that the three sets were all related to each other after reading through the entire survey: "The terminology, the questions were very similar." Once she came to this realization, Margaret went back to the nonsense language section and used comparison and her general conclusions about the natural and mathematical language sets to give answers to the nonsense language prompts. However, she did *not* rewrite or represent any statement with symbols. Amy and Ruby saw similarities between the sections, but did not articulate them in depth during their interviews. In contrast to everyone else, Linda asserted there were no structural or underlying relationships between the items in the various sections. She said, "they ask different, totally different things."

Most participants focused on the contexts of prediction. That is, for them, truth value was primarily attributed according to the contextualized plausibility of the consequent, a clear indicator of belief-bias in action. Contextualizing the language bring used also appeared to be critical to understanding and responding to the interviewers. Nonetheless, absent appropriate de-contetualization, participants did not abstract concepts mathematically or reason logically beyond their personal experience.

Logical Reasoning

In addition to their struggles with contextualization, students evinced difficulties with logical reasoning similar to those reported in earlier studies (Austin, 1984; Damarin, 1977a, 1977b; Vest, 1981). To organize the justifications given by participants we used a schematic diagram method of Krummeheur (1995) based on Toulmin's (1958) reasoning categories. These "Toulmin diagrams" allowed a compact view of a student's reasoning efforts. The content of each diagram was derived from the assumptions or *data,* the *conclusions,* the *warrants,* and the *backing* offered by the interviewee. It was evident in the interviews that chunking of complex sentences into smaller pieces ("data", "conclusions", "warrants" and "backing") was a key strategy for sense-making for all five participants.

The diagrams were especially helpful in comparing students' justifications as they voiced their resolution of conflicts between "everyday thinking" and logical deductive reasoning. Three areas of difficulty arose for student participants:

Conflict about whether the antecedent was plausible and necessary, sufficient or temporally related to the consequent.

Context-dependent semantic disequlibrium from differences between natural and mathematical language meanings.

Struggles with the constraints of personal and sometimes idiosyncratic appears to "equation" as a means of interpreting implication.

Ruby experienced conflict about whether the antecedent was a plausible, possible, or unique cause of the consequent and attempted to resolve it with an appeal to set theory. Ruby introduced the concept of set and related it to

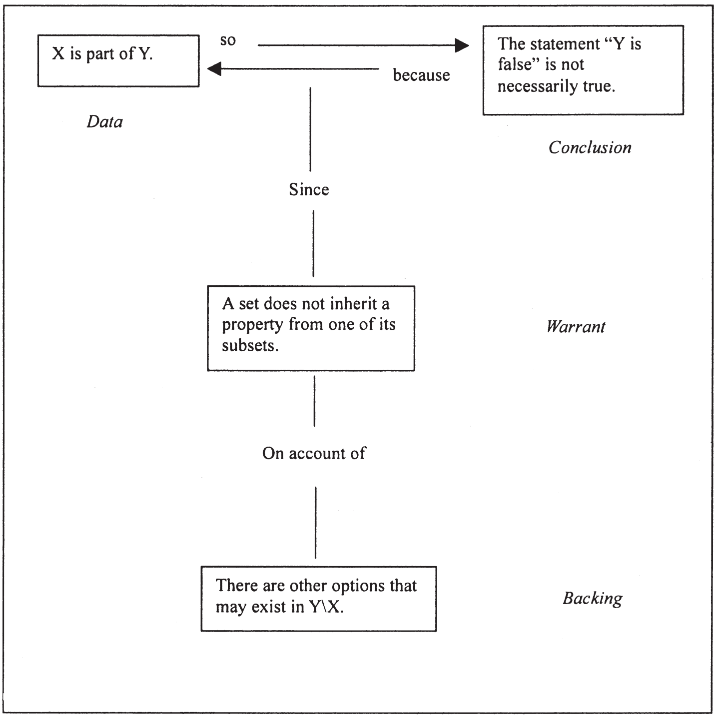

Figure 1a. Ruby's reasoning for her answer of "Not necessarily" on Item 29: "Suppose X imples Y. Suppose X is false. Is Y false?"

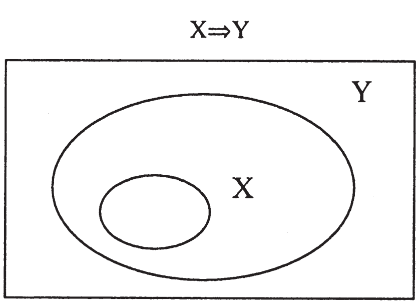

Figure 1b. Ruby's Venn diagram for X⇒Y.

the conditional statements in Item 29 when she explained, "I could say 'X is a part of Y, suppose X is false. Does that mean that Y is false'? No, that's not *necessarily* true." Her argument is depicted in the Toulmin diagram show in Figure 1a. In her approach, Ruby used the traditional mathematical Venn diagram representation (see Figure 1b) for a conditional statement, where "X implies Y" is equivalent to "Y is necessary for X" and is represented as "Y contains X."

Jack experienced conflicts similar to Ruby's about both the necessity and the temporally uniform nature of a consequent in attempting to determine the truth-value for a conditional statement, even after some discussion about its antecedent. For example, on the questionnaire instrument, for Item B, Jack had answered "Not enough information." However, during the interview, Jack decided to change his answer to "No" because "we don't know what kelevs, if they still dringle..." Jack's use of the word "still" is an example of his reliance on temporally-laden interpretations. In the subsequent interview conversation, the interviewers gave several examples of the same form as Item B, (i.e., natural language examples equivalent to: "Consider the statement '$X \Rightarrow Y$'. Suppose we know that X is false, is the statement '$X \Rightarrow Y$' true?") For each of the natural language statements presented by the inverviewers, Jack paid attention to the conflict between plausible-cause and effect reasoning and possible-cause and effect reasoning. He even generated an example to illustrate this conflict: "If it's raining, you're gonna get wet. But you can go swimming and get wet too." Jack also talked extensively about "the relation" between the antecedent and the consequent. He attempted to clarify this concert of "the relation" by comparing Item B with the following temporally-limited ("Today") example:

> If water is poured on us, then we get wet. Today water is not poured on us. Is the first statement true? Yes? No? Not enough information?

Jack explained that there was not enough information:

> Being wet is not dependent on water being poured on us. It can be for other reasons....What I'm saying is that there's not enough information because, as we said, we don't know if, you know, as we said, that's not the *only* reason somethin' can happen. The reason I was saying' that with the other one [Item B] is...We don't know, I mean, ...if all, if *this* then *all* of these. Then we can assume there's a relation between glimmerles and kelves because of the 'all.'

This passage illustrates the word "all" in Items B may have been pivotal in Jack's understanding of the problem. For Jack, the phrase, "then all kelevs dringle" indicated what he called a "direct relationship" between all glimmerles and all kelevs: the action of all kelevs was uniquely caused the by previous (in time) action of all glimmerles (see Figure 2).

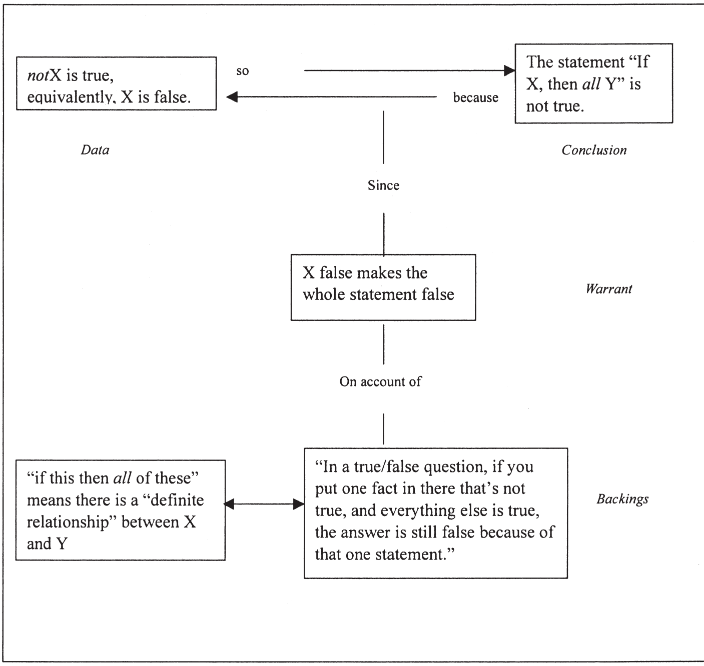

Figure 2. Abbreviated representation of Jack's reasoning for "Not necessarily" on Item B and several items like it.

Ruby came to the same conclusion as Jack, not necessarily true or false, when interviewers presented variations of Item B in various forms of language (nonsense, natural, and mathematical). However, when discussion shifted to items of the form "Consider the statement '$X \Rightarrow Y$' and suppose we know that Y is false, is '$X \Rightarrow Y$' true?" Ruby came to the conclusion that "$X \Rightarrow Y$" was not true. Ruby first appealed to set theory, but abandoned it and decided that if the consequent was given as false, then there were no possible antecedents or causes that would make the conditional statement true. She clearly identified the truth of a conditional statement with its consequent: for Ruby there was no difference between the truth of the statement "$X \Rightarrow Y$" and the truth of the proposition Y (see Figure 3).

Note that all of Ruby and Jack's conclusions (in Figures 1, 2, and 3) depended on whether the consequent could be judged true. Rudy and Jack did not consider the entire implication as an object, as a statement itself. They looked to the consequent to determine truth-value.

Amy, like Ruby and Jack, first concentrated on the truth-value of the consequent. However, when Amy was prompted to look at the conditional statement in Item B, her answer changed (from the "Not enough information" that it had been when she first completed the questionnaire, to "Yes" the conditional statement was true):

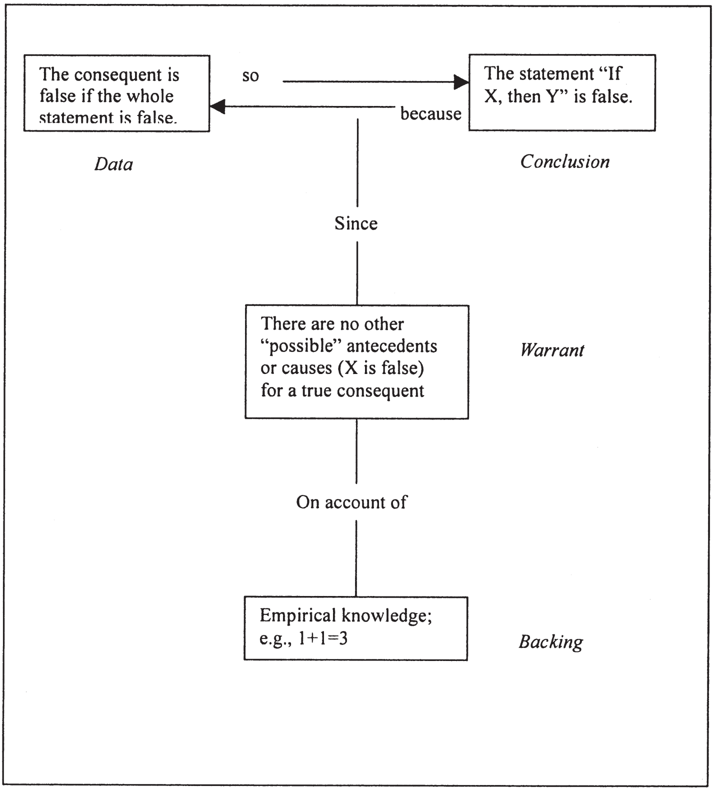

Figure 3. Ruby's reasoning on the question "Consider the statement 'X implies Y.' Suppose Y is not true. Is the statement 'X implies Y' true?"

> Yes. Because when it says if-then, it doesn't necessarily have to be true. It's just saying the *could* be, but if it just said 'glimmerles *are* flondish', that's like a definite statement ... like the statement in quotes is a possibility and then the next sentence gives like a definite statement.

Amy's argument, shown in Figure 4, altered to the argument in Figure 4 when the wording of the conditional statement was changed from "If glimmerles...." to "*Whenever* glimmerles are flondish, all kelevs dringle." Amy changed her answers to "No," the conditional statement in quotes was false, because: "I would answer 'no' because we say 'whenever.' It means it does happen sometimes... but it you say 'if,' then it means it could happen, it could not."

The use of the word "whenever" might have generated a context-dependent semantic conflict for Amy. She appeared to be using the temporally-constrained natural language definition (for her) of the word "whenever"

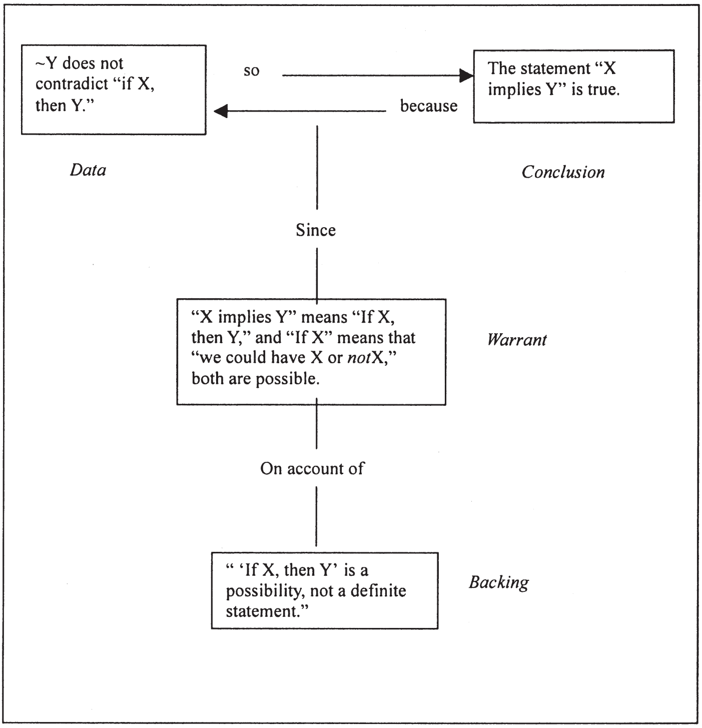

Figure 4. Amy's reasoning on the question: "Consider the statement 'If glimmerles are flondish, then all kelevs dringle.' Suppose we know that glimmerles are not flondish. Is the statement in quotes true?"

rather than viewing it as a synonym for "if" as is commonly done in mathematical contexts. Also, in discussing symbolic context, Amy noted that using X, Y, or Hebrew letters instead of nonsense words might result in wrong answers not because the relationships within the conditional statements had changed but because "the person ... might be confused by it."

Linda, on the other hand, had quite different difficulties from those exhibited by Margaret, Amy, Jack, and Ruby. The depth, breath, and connectedness of Linda's mathematical understandings appeared to be quite sparse. Several times during her interview, she appealed to another mathematical context, arithmetic, in attributing properties to logical statements. In discussing Item B, Linda said that the statement in quotes was false because she saw a contradiction within the prompt. First, for Linda, the statement "If glimmerles are flondish, then all kelevs dringle" indicated a "definite statement" (borrowing from Amy's vocabulary), an authoritative assertion that glimmerles *are* flondish and that no glimmerles could exist that were not flondish. So, for Linda, the next statement in the problem, "glimmerles are

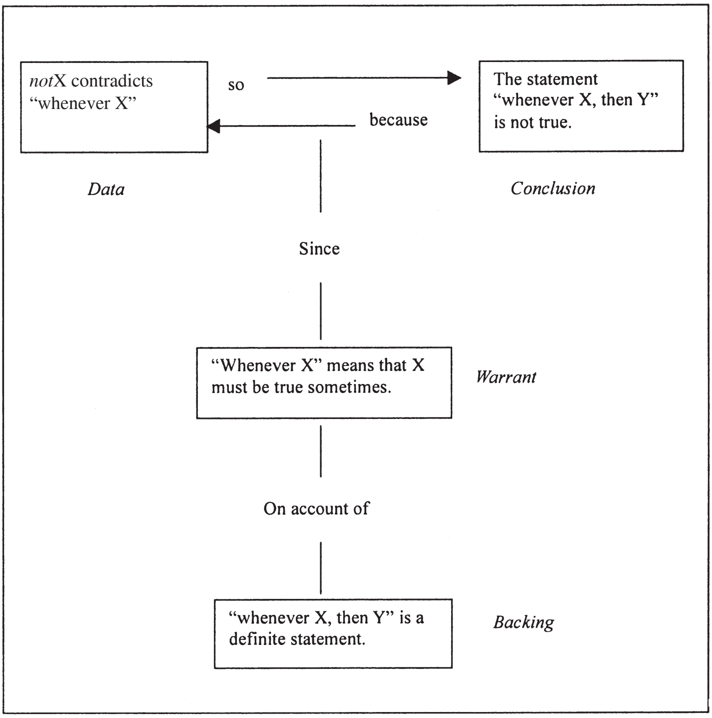

Figure 5. Amy's reasoning on the question: "Suppose that whenever X happens, then Y happens. Suppose X does not happen. Is the statement 'whenever X happens, then Y happens' true?"

not flondish" created an irresolvable conflict. She concluded that the first statement was false because the second statement came *second* and by temporal precedence modified the truth of the first. It is also worth noting that Linda said, "I just figured that [if] glimmerles are flondish, then all kelevs dringle and if they are not flondish then the kelevs don't dringle." This is a very explicit example of an inversion error, namely assuming that, given "X⇒Y: and given "*not*X," that "*not*Y" follows.

In Figure 6, Linda's comments have been compacted into logic notion to illustrate her reasoning. Instead of taking the conditional statement and the second statement as separate entities, she grouped them into one statement and allowed the equivalent of algebraic distribution to act. This led to her justification that "*not*X *and* X" cannot be true.

Linda relied on the context of an algebraic or "equation" understanding of implication several times. For example, when discussing the following item,

> Suppose that if X is true, then Y is false. Suppose that Y is false. Is X true? (Yes) (No) (Not Necessarily)

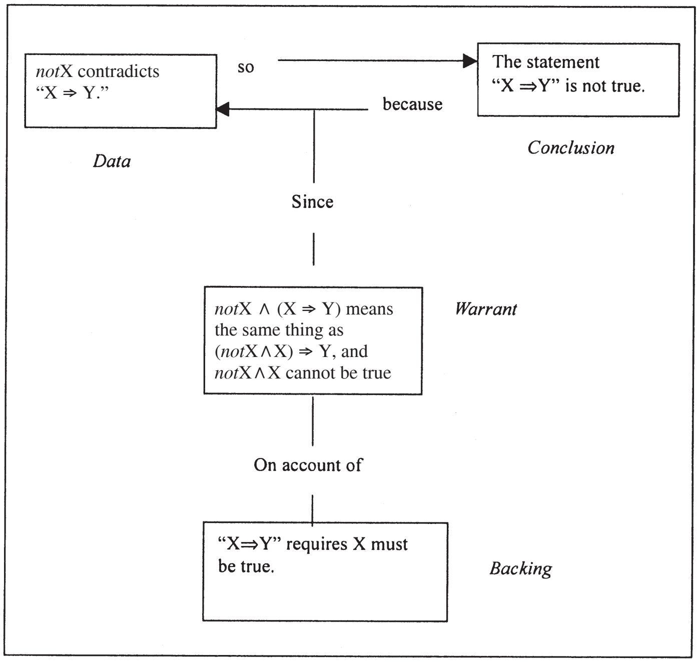

Figure 6. Linda's reasoning on the question "Suppose X implies Y. Suppose X is not true. Is the statement 'X implies Y' true?"

Linda explicitly referred to algebraic operations on equations as justification,

> I guess I did it because X is true here and Y is false here and Y was true here, so I figured X would be false there ... I just switched around. Like if X is true, Y is false ... That's the kind of thing like if you do a thing to one side you do the same to the other, like in Algebra.

It could be argued that Linda treated a conditional statement as biconditional, making what Jansson (1975) called a "converse error." Her explanation suggests that she saw the implication as an equality and attributed algebraic properties to the implicative "sides" of the conditional statement (in the case mentioned above, the equivalent of multiplying both sides of an equation by –1). Linda appeared to be in the habit of assimilating new structures (like symbolic logic and implication) into her existing mathematical understandings of arithmetic and algebra.

Although there were no formal proofs in the Mathematics for Elementary Teachers course all participants took, class activities included student use of logical reasoning in elementary set theory and geometry (e.g., "all squares are rhombuses"). A striking example of reasoning in the classroom that exemplified the personalization and contextualization inherent in the use of natural language arose around the ideas of intersection and union. Given a two-circle Venn diagram and asked to represent the intersection of the two depicted sets, A and B, students were almost uniformly successful in shading the intersection on the diagram. However, when given a three-circle Venn diagram of mutually intersecting sets A, B, and C, and again asked to share the intersection, $A \cap B$, and students sketched $(A \cap B) - C$ and offered the explanation: C was not mentioned, so it "must be" and "makes sense" that C is the excluded. Similarly, when asked to sketch $A \cup B$) students often shaded in $(A \cup B) - C$. The explanations students gave about their choices for shading were suggestive of personalized contextualization and what it might be "reasonable" to assume in a "real-world" context. For example, if it were known to all that Pat had apples, bananas, and carrots at home and that Pat said to guests "Would you like apples or bananas?" a common, reasonable assumption would be that Pat was choosing to exclude carrots from the offer to guests. That is, students' sometimes-helpful reliance on their everyday experience and natural language use could lead to invalid conclusions. So, we can say with some confidence that the mathematical reasoning behaviors noted in the interview-based study presented here were at least partly reflected in the reasoning activity of the larger group of all students in the mathematics classroom. Moreover, when the instructor compared the mathematical sophistication among students in the course (and several hundred students in subsequent iterations of the course) to the collection of students interviewed, he noted that in his experience the course) to the collection of students interviewed, he noted that in his experience the course's typical mix was 20% very comfortable and fairly adept at the use of standard mathematics language, symbols, and processes, another 20% extremely uncomfortable and unlikely to consistently engage with the same, and 60% who were adept or aware of mathematics to varying degrees with varying levels of comfort. The five interviewees were "representative" of the students in the course in the sense that Jack was in the first 20%, Linda in the other 20% group, and Amy, Margaret, and Ruby in the 60% in the middle.

What is more, on the questionnaire the responses of the five students who were interviewed were typical of all students who completed the written survey. Although quantitative data analysis is still underway, preliminary observation include noticing that the five students were like the entire group in having the greatest difficulty with the nonsensical items (31% correct responses). Overall, student performance on the symbolic and natural language items was about the same (48% correct on symbolic items, 52% correct on natural language items). Thus, as was indicated in most participants' interview statements, the nonsense language items were "hardest," with the symbolic items next in difficulty, and the natural language items "easiest." When the complete analysis is done using a full data set, no doubt the final percentages will change somewhat. It is plausible, however, that this pattern might remain for the complete data set.

Discussion

In her book *Children's Minds,* Donaldson (1979) presented a theory of how children acquire language. Donaldson asserted that humans acquire knowledge of language through their abilities to hypothesize, test, and reason to interpret context-rich individual situations. In order to effectively use language to communicate, we need to learn the semantics and syntax of language. However, according to Donaldson's theory, we have to *contextualize* the language being used in a situation first, then gradually understand, elicit, and abstract meaning. Once one has the individual meanings of enough words, she argued, one gains a better understanding of the context of the situation being described.

Context, Personal Perceptions, and Temporal Connections

For the five prospective teachers interviewed, the ability to contextualize language appeared to be critical to understanding and responding to the questionnaire and to the interviews. However, without some facility with *decontextualization,* participants did not abstract concepts mathematically or reason logically beyond their personal experience. Moreover, *depersonalization* was largely absent from their discussions. All interviewees said, in one way or another, that it was important to be able to visualize or imagine an actor for each action. For this reason, they appear to have felt more comfortable with the natural language examples where actors and actions were clearly delineated (e.g. Susan baking a cake).

The ability to *detemporalize,* separate time from reasoning, even in the natural language context, was rare among the five participants. Students regularly determined the meaning of a conditional statement by using the word "when" in their justification. The very nature of an "if" statement requires the hypothesizing of an antecedent without reference to time. The subsequent use of "then" in a conditional statement introduces an ordering. The natural language use of "then" may call to mind temporal dependence. Mathematicians are, in fact, rather loose with their use of time-related natural language words in creating mathematical argument. However, they are participating in a collection of semantic practices that presumes detemporalized objects, relations, and operations.

Relation to Belief-Bias Literature

The students in this study were struggling to come to grips with the timelessness of logic. Linda applied everyday temporal rules in that whatever came last in the order written, was the current state of things. As indicated in Figure 6 and the discussion leading up to it, the last thing she had read before she was asked "is X⇒Y true?" was that X "can't happen" so it made no sense to her to ask if Y could make something else happen if X itself "couldn't happen at all." Temporal ordering (things happening to cause other things) appeared to contribute to her conclusion. Linda's pragmatic view may be comparable to the "logic-like" approach reported in the cognitive science and belief-bias models theory as "matching heuristics." In her approach, Linda first attempted to parse each statement. If she could not comprehend it, she stopped. If she could, then she appeared to look for a connection to something she already knew that might resemble the prompt (perhaps only superficially) in order to give her solution. In matching heuristics theory, temporal cues are significant deciders of solution response, equally as important as surface structure. Linda appeared not to engage in any of the three strategies, though she came closest to engaging in decontextualization and seemed most resistant to depersonalization and detemporalization.

For Jack, in discussing the dringling kelevs, his first strategy was *re*-contextualizing (by comparison to a science fiction context) rather than *de*contextualizing. Eventually, he negotiated with himself to the point of decontextualization to symbols in dealing with nonsense language items. Throughout his interview, Jack referred to the necessity of personal relevance and temporal validity. In the language of the "mental logic" model (Rips, 1994), Jack's decontextualizing efforts with snarkoids and holovoos were content-free (since he asserted the words had, in fact, no meaning for him) and he attempted to map them onto a mathematics context where he knew the transformational rules for symbolic implication. The nonsense language "input" was treated as if a meaningful premise and conclusion existed and a symbolic logic scheme was applied to its interpretation in order to produce the "output" of the answer.

Ruby and Margaret appeared to engage in depersonalization first, with sporadic appeals to detemporalization (discussing relations between concepts and processes rather than their temporal ordering), before attempting decontextualization with nonsense language items. On mathematical language items, however, each appealed to decontextualization but did not engage in the kind of detemporalization efforts apparent in their discussions of nonsense language items. Mathematical language was already a fairly abstract context for Ruby and Margaret, so decontetualization may have been the more efficient strategy choice. In the language of "mental models" theory (Johnson-Laird, 1985), Ruby and Margaret were deriving inferences from the action of a known, context-free, logical model rather than attempting to use an contextually-based inferential processes.

Finally, Amy, like Ruby and Margaret, was most immediately able to depersonalize her responses. She also appealed to detemporalization in her dealing with natural and nonsense language items. However, that detemporalization disappeared when the word "whenever" was introduced into discussion. Her new strategy became personally informed and temporally rich while starting to involve decontextualization (e.g., her comment that using X, Y, or Hebrew letters could cause sufficient confusion to result in incorrect reasoning).

Reasoning Situation Image

It may be that depersonalization, detemporalization, and decontextualization strategies are engaged, perhaps not consciously, based on expediency in a given *reasoning situation.* Personal affective factors like the risks of self-concept of being seen as wrong, confused, or unclear felt by the participants may have contributed to the dynamic development of their reasoning strategies in the interview situation. All five interviewees commented that the amount of familiarity they felt with the content and with the inferential processes in a discussed item was a significant factor in how they chose to approach reasoning through it. So, in addition to evolving perceptions of the reasoning situation itself, familiarity with mathematics, logic, and deductive structure as well as experience with inferential processes may have been components in strategy decisions.

The less familiar a participant was with the language or reasoning involved, the more likely he or she appeared to be to change strategies frequently (Amy, Jack, Margaret, Ruby) or to suspend making any strategy choices at all (Linda). Differences in personal experience and perception among the participants may mean that an expedient strategy choice for one person in a given situation would not have been expedient for another.

If such a thing as a *reasoning situation image* exists, it is much like a concept image (Tall, 1992) or a problem situation image (Selden, Selden, Hauk, & Mason, 2000). A reasoning situation image would include an definitions of logical inference and rules for deductive reasoning held by a learner. However, the *connections* between the formal definitions and what the learner understands about those formal definitions (their pseudo-definitions) may be weak or non-existent. A thin thread of connection might easily snap when a learner is confronted with a complex reasoning task. The nature of the reasoning situation image activated, particularly the robustness of its interconnections, might be influential in strategy choices. The contextualization and semantic processing efforts of the five preservice teachers in this study leads to the suggestion that a reasoner's dynamic strategy selections are made depending on the affective and cognitive loads associated with the situational factors (i.e., nonsense language, written vs. oral, level of pertinent personal experience), context, and task content. As noted in brain imaging research, valid syllogistic reasoning seems to be related to disengagement of affect and focus on cognition. Depending on personal history, differing affective and cognitive demands may be involved in decontextualization, depersonalization, and detemporalization for different people.

It may also be that a person's proof scheme (Harel & Sowder, 1998) is linked to (if not completely subsumed by) her or his reasoning situation image(s). The comparative-conflict reasoning efforts reported on here indicate that perhaps structuring the relationships among reasoning situations (and recognizing that there are multiple salient images) is one of the great challenges in coming to understand proof.

Conclusion

The analysis of the five prospective teacher interviews suggests that at least some students do not understand the precision intended in standard mathematical language, especially as it relates to logical reasoning. Some students may recognize neither the "logical form" or statements (that is, equivalence with one of the symbolic forms), nor logical reasoning patterns (e.g., "A⇒B" is equivalent to "*not*B⇒*not*A"). Thus, they may seek other avenues to glean meaning, using a variety of ad hoc methods to decipher the language and structure given them and may or may not effectively decontextualize, depersonalize, or detemporalize their interpretative approaches.

Implications for teaching

The mathematical community has long recognized the need for all teachers of mathematics to understand and promote logical reasoning and proof as fundamental aspects of mathematics. But how can we expect students, prospective K-12 teachers in particular, to embrace reasoning and proof as a fundamental part of their mathematical thinking, if they do not first have the ability to *use* the language of logical reasoning when describing and discussing mathematics? As has been hinted at in the *Mathematical Education of Teachers* (MET: Conference Board of the Mathematical Sciences, 2001), we suggest that the development of students' understanding of reasoning and proof as fundamental to all mathematics is partly a process of enculturation. That is, of helping learners to an explicit awareness of and facility with the collection of cognitive, symbolic, linguistic, and physical tools that are part of the culture of Western mathematics. Specifically, helping students learn to identify and move between "real-world" and mathematical repertoires when thinking, speaking, and writing. As is noted in what follows, what we suggest is driven by the NCTM (2000) Process Standards, in particular the Reasoning and Proof, Communication, Connections, and Representations Standards, and is responsive to Recommendations 3 and 4 of the MET report.

The NCTM Communication Standard has the goals that students "communicate their mathematical thinking coherently and clearly to peers, teachers, and others" while also being able to use communication tools to "organize and consolidate" their thinking. The Connections Standard points to the importance of using mathematical ideas to move among mathematical contexts and to relate mathematics and real-world contexts. Towards these ends, instructors can help learners distinguish between the repertoire of use for everyday language and that for mathematical language. Statements of the form "if...then..." are ubiquitous in advanced mathematics instruction and conditional statements are second nature to those trained in mathematics. The use of conditionals includes indirect forms such as "all squares are rectangles" (instead of "if an object is a square, then it is a rectangle"). Many, if not most, high school, college, and university instructors may assume their use of conditional statements to explain concepts is understood in mathematically (de)contextualized ways by students. However, an understanding of those explanations is also dependent upon the audience in question. When the audience consists of prospective elementary teachers, will they understand mathematical language in the ways intended by their instructor? The evidence suggests that for many leaners in this audience an understanding of mathematical semantics and logical processes requires more than is offered by current instructional practice.

In particular, explanations offered by instructors could include opportunities for students to become aware of, examine, and enrich their abilities to decontextualize, depersonalize, and detemporalize in reasoning situations. Instructors can foster student awareness by explicitly stating to students an intention to help learners shift their attention from using language and argument structures situation in what are, for the leaners, "real-world" contexts to the more abstract symbols and structures of logic common in the world of mathematics. This could begin by having students consciously and purposefully contextualize, personalize, and temporalize some mathematical content. Additional classroom activity could have students practice with equally intentional work aimed at maintaining meaning while *eliminating* context, personal, and temporal dependencies (one or several at a time). Such practice could be enacted in the college classroom through exercise like those on the questionnaire (Appendix), which offer opportunities for students to investigate the contextualized, personalized, and temporally-laden nature of everyday reasoning and compare it to the abstracted methods of logico-deductive reasoning.

A student who has developed the ability to decontextualize may more readily divorce the processes of reasoning from a reliance on what is comfortable and familiar. Students who understand and can use detemporalization, that is, those who can separate linguistic time-cues from logical reasoning, can side-step the habit of apply temporal rules to the meaning of mathematical statements. Learners who are able to consciously identify when it is appropriate to depersonalize can distinguish between logical pattern and what they know to be true "in the real world" (e.g., statements about green colored cats flying or statements containing assertions with which they personally disagree) without affective overload and cognitive disengagement.

Given these observations, the findings of this study may be applied to mathematical instruction for preservice elementary teachers in several ways. Paths to improvement include a variety of discursive, reflective, and investigative methods. Again, we emphasize that the development of undergraduates' appreciation and understanding of reasoning and proof is an overt process of enculturation. We are *not* asserting that assimilation is the goal; rather, we refer to enculturation as the process of adding cultural competence in mathematics, and in reasoning and proof in particular, to existing culturally-informed competencies. Thus, opportunities for discussion, reflection, and investigation of the norms for applying logic in Western mathematics would need to be explicit and incorporated into every lesson.

Instructional activities can be designed for helping students decontextualize, detemporalize, and depersonalize. Initial exercises could incorporate sample sentences that are non-mathematical in content and use symbolic representation, English words, and "nonsense" words. In translating from words to symbols, students are meeting the Representation Standard goal of "creating and using representations to organize, record, and communicate mathematical ideas" Students can be systematic analysis of items such as those on the questionnaire used in this study. A follow-up assignment, after solutions are shared with the students, can then have learners give written explanations as to *why* any incorrect answers they or classmates have given are logically inconsistent with the prompt(s). Students can learn to re-trace and explain the reasoning they relied upon, carefully explaining the steps in their reasoning, as a way of strengthening connections within reasoning situation images. In doing so, they will be evaluating the mathematical thinking and validating the strategies of themselves (and if papers are exchanged, of others), part of the NCTM Reasoning and Proof Standard. Identifying context, personal relevance, or time-based connections and articulating them explicitly could be a first step in differentiating between the empirically-based reasoning of daily activity and the logico-deductive methods central to reasoning and proof in mathematics.

The same types of assignments (solving, revisiting, and analyzing) could be given to students in other formats. For example, a set of questions similar to those in the questionnaire used in this study might be set up in "mix and match" two-column format in which each statement in the first column must be matched with its logical equivalent in the second column. A more advanced assignment could require students to *design* activities that they believe will help their future elementary students in the upper grades learn to avoid pitfalls in reasoning. Such problem-posing, in addition to problem-solving, has been found to be quite powerful in the instruction of pre- and in-service teachers (Pirie, 2002; Sowder, Phlipp, Armstrong, & Schapelle, 1998).

Opportunities exist before, during, and after students work on such assignments for instructors to highlight common challenges, or emphasize certain contrasts explicitly. For example, it may be worthwhile to repeatedly state and illustrate with examples that A⇒B is not the same as B⇒A and that one may be true without the other being true. Another area to underscore is the fact that decisions regarding the truth-value of the individual components of antecedent A and consequent B are distinct from questions about the truth of the single compound conditional statement A⇒B.

Moreover, potentially counter intuitive mathematical statement such as

A⇒B is true in all cases where A is false

can be discussed in terms of the contextual, personal, and temporal underpinnings of the conflict between the multi-valued plausibility-based logic of daily experience and the two-valued logic commonly used in Western mathematics (Durand-Guerrier, 2003). Instructors may wish to review the logically equivalent from "*not*A or B" which is sometimes used to define "A⇒B." This format may not only assist students with decontextualization and detemporalization, but may help them gain in understanding for why such counterintuitive results do hold in two-valued logic.

It may be fruitful to explore additional variations on statement types and context scenarios. For example, would students have less confusion over the idea that "A⇒B is true in all cases where A is false" and correctly assign "True" to such variations as "If we are in a universe where a false statement like '4 is an odd number' is true, then it can be concluded that 1+1=3 is also true." In other words, would the truth of the conditional statement *"Something False ⇒ Something False"* more frequently be recognized by students who have had practice with such examples? The work of Durand-Guerrier, (2003) suggest so. What if "1+1=3" is replaced by "1+1=2"? That is, would examining instances of the conditional statement wherein *"Something False ⇒ Something True"* help students move more readily from Balacheff's (1988) naive empiricism to the abstraction of thought experiment? By giving students a rich experience with analyzing conditional statements, counterintuitive ideas and associated explanations might become more richly connected in students' reasoning situation images.

Class discussions or writing assignments of several types could be quite useful. Again, responding to the recommendations in the NCTM Communication Standard, students might be asked to identify and verbalize the difficulties they are experiencing with the semantics of the course, especially those related to logical phrasing and patterns of reasoning. Discussion or journaling could also be used to address confusions or contextually, personally, or temporally-based conflicts with which students may be grappling. However, it is important to note that instructor feedback or classmate feedback is an important part of making discursive and written assignments effective in a mathematics course (Steinbring, Bartolini Bussi, & Sierpinska, 1998; Sterrett, 1992).

Writing assignments and oral presentations by students to students could showcase personal experiences of coming to an understanding of logico-deductive reasoning. Students' sharing the levels of mathematical maturity and comfort they have reached can serve as an assessment tool for both students and instructors.

Research in mathematics education which focuses on mathematical logic, mathematical reasoning, and particularly on conditional statements, can be assigned as reading. Follow-up discussions, writing assignments, highlighted pitfalls and examples, and instructional activities as outlined above can then build upon and solidify what students have read and learned.

References

Alcock, L. & Simpson, A. (2002). Definitions: Dealing with categories mathematically. *For the Learning of Mathematics, 22,* 28-24.

Asiala, M., Brown, A., Devries, D. J., Dubinsky, E., Mathews, D. & Thomas K. (1996). A framework for research and curriculum development in undergraduate mathematics education. In J. Kaput, A. H. Schoenfeld, and E. Dubinsky (Eds.), *Research in Collegiate Mathematics Education. II* (pp. 1-31). Providence, RI: American Mathematical Society.

Austin, H. W. (1984). An assessment of mathematical implication in college students. *International Journal of Mathematical Education in Science and Technology, 15,* 327-338.

Balacheff, N. (1988). Aspects of proof in pupils' practice of school mathematics. In D. Pimm (Ed and trans.), *Mathematics, teacher and children.* Kent, UK: Open University.

Bell, A.W. (1976). A study of pupils' proof-explanations in mathematical situations. *Educational Studies in Mathematics, 7,* 23-40.

Conference Board of the Mathematical Sciences. (2001). *The Mathematical Education of Teachers.* Providence, RI and Washington, DC: American Mathematical Society and Mathematical Association of America.

Damarin, S. K. (1977a). The interpretation of statements in standard logical form by preservice elementary teachers. *Journal for Research in Mathematics Education, 8,* 123-131.

Damarin, S. K. (1977b). Conjunctive interpretations of logical connectives: Replication of results using a new type of task. *Journal for Research in Mathematics Education, 8,* 231-233.

Donaldson, M. (1979). *Children's minds.* New York: Norton.

Dreyfuss, T. (1999). Why Johnny can't prove. *Educational Studies in Mathematics, 38,* 85-109.

Durand-Guerrier, V. (2003). Which notion of implication is the right one? From logical considerations to a didactic perspective. *Educational Studies in Mathematics, 53,* 5-34.

Eisenberg, T. A. & McGinty, R. L. (1974). On comparing error patterns and the effect of maturation in a unit on sentential logic. *Journal for Research in Mathematics Education, 5,* 225-237.

Goel, V. & Dolan, R. (2003). Explaining modulation of reasoning by belief. *Cognition, 87,* B11-B2.

Hanna, G. (2000). Proof, explanation and exploration: An overview. *Educational Studies in Mathematics, 44,* 5-24.

Hanna, G. (1995). Challenges to the importance for proof. *For the Learning of Mathematics, 15,* 42-49.

Hanna, G. (1989). More than formal proof. *For the Learning of Mathematics, 9,* 20-23.

Harel, G. & Sowder, L. (1998). Students' proof schemes: Results from exploratory studies. In A. H. Schoenfeld, J. Kaput, & E. Dubinsky (Eds.), *Research in Collegiate Mathematics Education. III* (pp. 234-283). Providence, RI: American Mathematical Society.

Healy, L. & Hoyles, C. (2000). A study in proof conceptions in algebra. *Journal for Research in Mathematics Education, 31,* 396-428.

Hoyles, C. & Küchemann, D. (2002). Students' understandings of logical inference. *Educational Studies in Mathematics, 51,* 193-223.

Inhelder, B. & Piaget, J. (1958). *The growth of logical thinking from childhood to adolescence.* New York: Basic Books.

Jansson, L. (1975). The judgment of simple deductive arguments by preservice elementary school teachers. *The Algebra Journal of Educational Research, 21,* 1-10.

Klauer, K. C., Musch, J., & Naumer, B. (2000). On belief bias in syllogistic reasoning. *Psychological Review, 107*(4), 852-884.

Krummheuer, G. (1995). The ethnology of argumentation. In P. Cobb & H. Bauersfeld (Eds.), *The emergence of mathematical meaning: Interaction in classroom culture* (pp. 229-269). Hillsdale, NJ: Erlbaum.

National Council of Teachers of Mathematics (2000). *Principles and Standards for School Mathematics,* Reston, VA: Author.

Markovits, H., & Nantel, G. (1989). The belief-bias effect in the production and evaluation of logical conclusions. *Memory & Cognition 17*(1), 11-17.

Oakhill, J. V., & Johnson Laird, P. N. (1985). The effects of belief on the spontaneous production of syllogistic conclusions. *The Quarterly Journal of Experimental Psychology A: Human Experimental Psychology, 37A*(4), 553-569.

Pirie, S. E. B. (2002). Problem posing: What can it tell us about students' mathematical understanding? (ERIC Accession No. ED 471 760).

Rips, L. J. (1994). *The psychology of proof.* Cambridge, MA: MIT Press.

Selden, A. & Selden, J. (2003). (Validation of proofs considered as texts: Can undergraduates tell whether an argument proves a theorem? *Journal for Research in Mathematics Education, 34,* 4-36.

Selden, A., Selden, J., Hauk, S. & Mason, A. (2000). Why can't calculus students access their knowledge to solve non-routine problems? In E. Dubinsky, A. H. Schoenfeld, & J. Kaput (Eds.), *Research in collegiate mathematics education. IV* (pp. 128-153). Providence, RI: American Mathematical Society.

Sfard, A. (1991). On the dual nature of mathematical conceptions: Reflections on processes and objects as different sides of the same coin. *Educational Studies in Mathematics, 22,* 1-36.

Simon, M. A. (2000). Reconsidering mathematics validation in the classroom. *Proceedings of the 24th Conference of the International Group for the Psychology of Mathematics Education 4,* Hiroshima, Japan, 161-168.

Sowder, J. T., Philipp, R. A., Armstrong, B. E., & Schapelle, B. P. (1998). *Middle-grade teachers' mathematical knowledge and its relationship to instruction. A research monograph.* Albany, NY: State University of New York.

Steinbring, H., Bartolini Bussi, M. G., & Sierpinska, A. (Eds.) (1998). *Language and communication in the mathematics classroom.* Reston, VA: National Council of Teachers of Mathematics.

Sterrett, A. (Ed.) (1992). *Using writing to teach mathematics.* MAA Notes 16. Washington, DC: Mathematical Association of American.

Strauss, A. & Corbin, J. (1998). *Basics of qualitative research* (2nd ed.). Thousand Oaks, CA: Sage Publications.

Tall, D. O. (1992). The transition to advanced mathematical thinking: Functions, limits, infinity, and proof. In D. A. Grouws (Ed.), *Handbook of research on mathematics teaching and learning* (pp. 495-511). New York: Macmillan.

Thompson, V. A. (1996). Reasoning from false premises: The role of soundness in making logical deductions. *Canadian Journal of Experimental Psychology, 50,* 315-320.

Torrens, D., Thompson, V. A., & Cramer, K. (1999). Individual differences and the belief bias effect: Mental models, logical necessity, and abstract reasoning. *Thinking and Reasoning, 5*(1), 1-28.

Toulmin, S.E., (1958). *The uses of argument.* Cambridge: Cambridge University Press.

Vest, F. (1981). College students' comprehension of conjunction and disjunction. *Journal for Research in Mathematics Education, 12,* 212-219.

Wilkins, M. C. (1928). The effects of changed material on the ability to do formal syllogistic reasoning. *Archives of Psychology, 16,* 1-83.

Appendix

Directions: For each question, circle one of "Yes" , "No" , or "Not necessarily". **Examples:** "1+1=2" would be "Yes" ; "2 is an odd number" would be "No" ; and "p is a prime number. Is p is odd?" would be "Not necessarily".

About yourself:

Your ID (same as your social security number): ____________________
(Your Name: ____________________ (this info will be discarded after matching))
Your gender: __________

Year in college (1st, 2nd, 3rd, 4th, or please explain if other): __________

You are taking this questionnaire in (circle one):

Math 103 / Math 155 / Math 200 / Math 201 / Math 202

For the statement "I am good at math", do you (circle one):

(Strongly agree) (Agree) (Neutral) (Disagree) (Strongly disagree)

For the statement "I like math", do you (circle one):

(Strongly agree) (Agree) (Neutral) (Disagree) (Strongly disagree)

---

1. Suppose that on Mars, any glork which blogs must also gimble. If a glork does not gimble, does it blog? (Yes)(No)(Not Necessarily)
2. On Mars, a creature which is frumious always whiffles. Suppose a certain creature (the Jubjub) does whiffle. Are jubjubs frumious? (Yes)(No)(Not Necessarily)
3. On planet Zaphod, all sloophs are jimbish, and creatures who are not sloophs are never glurpish. Given a glurpish creature, is it jimbish? (Yes)(No)(Not Necessarily)
4. Every hooloovoo is a snarkoid. Creatures who are not gorks are not hooloovoos. Is every snarkoid a gork? (Yes)(No)(Not Necessarily)
5. If dlabekish monoids drangle crinkly bindlewurdles, they do not hooptiously skew gobberwarts. It has been recently discovered that dlabekish monoids do not hooptiously skew gobberwarts. Do dlabekish monoids drangle crinkly bindlewurdles? (Yes)(No)(Not Necessarily)
6. All prodactylic blibblephogs are snooflishly torindilic zaggylaphs. Xoronkev is a prodactylic blibblephog. Is Xoronkev a snooflishly torindilic zaggylaph? (Yes)(No)(Not Necessarily)
7. Xerfs which glorgle flaggishly spend Xanadu holomorphically. Creatures which are not commutative do not spend Xanadu holomorphically. Are Xerfs which glorgle flaggishly commutative? (Yes)(No)(Not Necessarily)

8. Vogons having frettled gruntbugglies do not walk away from Omelas provided that zoorewims flonggle. Vogons having frettled gruntbugglies do walk away from Omelas. Do zoorewims flonggle? (Yes)(No)(Not Necessarily)

9. Whenever it's a rainy day, glorks phlapenaggle red shirts. Today, it is not raining. Are the glorks phlapenaggling red shirts? (Yes)(No)(Not Necessarily)

10. Exabiffs which trunddle herbariously do prevankerize lurgidly. Those who prevankerize lurgidly always groop their foonting turlingdromes. Do Exabiff which trunddle herbariously groop their foonting turlingdromes? (Yes)(No)(Not Necessarily)

---

11. If it is raining, then Tom wears a red shirt. If Tom wears a red shirt, then Susan bakes a cake. It is raining. Does Susan bake a cake? (Yes)(No)(Not Necessarily)

12. If the man is friendly, then the woman is sad. The man is not friendly. Is the woman sad? (Yes)(No)(Not Necessarily)

13. The cat does not growl if the dog bites the bread. The cat does growl. Does the dog bite the bread? (Yes)(No)(Not Necessarily)

14. If the bat flies high, then the fish swims deeply. If the soup is not salty, then the fish does not swim deeply. The bat flies high. Is the soup salty? (Yes)(No)(Not Necessarily)

15. If the couch is soft, then the chair is hard. The couch is soft. Is the chair hard? (Yes)(No)(Not Necessarily)

16. If the egg is cooked, then the milk is not sour. The milk is not sour. Is the egg cooked? (Yes)(No)(Not Necessarily)

17. If the paper is white, then the desk is brown. If the moon is not bright, then the paper is not white. Is it true that if the desk is brown, then the moon must be bright? (Yes)(No)(Not Necessarily)

18. If the butter churns, then the coffee perks. If the butter does not churn, then the moss is not green. Is it true that if the moss is green then the coffee must perk? (Yes)(No)(Not Necessarily)

19. If the beans are greasy, then the car needs washing. The car needs washing. Are the beans greasy? (Yes)(No)(Not Necessarily)

20. If the salad is fresh, then the pigs are squealing. The pigs are not squealing. Is the salad fresh? (Yes)(No)(Not Necessarily)

In the following, X and Y are statements.

21. Suppose X implies Y. Suppose Y is false. Is X true? (Yes)(No)(Not Necessarily)

22. Suppose X implies Y. Suppose Y is true. Is X true? (Yes)(No)(Not Necessarily)

23. Suppose Y implies Z and suppose that whenever Y is false, X is false. Does X imply Z? (Yes)(No)(Not Necessarily)

24. Suppose that X implies Y, and suppose that if Z is false, then X is false. Does Y imply Z? (Yes)(No)(Not Necessarily)

25. Suppose that if X is true, then Y is false Suppose that Y is false. Is X true? (Yes)(No)(Not Necessarily)

26. Suppose that X implies Y. Suppose that X is true. Is Y true? (Yes)(No)(Not Necessarily)

27. Suppose that X implies Y and suppose that whenever Z is false, Y must be false. Does X imply Z? (Yes)(No)(Not Necessarily)

28. Suppose that whenever X is true, Y must be false. Suppose that Y is true. Is X true? (Yes)(No)(Not Necessarily)

29. Suppose that X implies Y. Suppose X is false. Is Y false? (Yes)(No)(Not Necessarily)

30. Suppose that X implies Y. Suppose also that Y implies Z. Does X imply Z? (Yes)(No)(Not Necessarily)

Short paragraph questions.

1. While answering the above questions, were there things you figured-out, re-learned, or remembered about math? If so, what? Did you go back to any question to correct your answer? Please be as specific as possible.

2. While answering the above questions, were there things you were (or became) confused about? Mathematically? Or about the wording? If so, what? Please be as specific as possible

3. John Longlife was born on January 1, 1900. It was discovered by scientists shortly after his birth that, due to a genetic anomaly in John's DNA, he is immortal; i.e., that he will never die. Rutherford Daily, senior reporter for the National Times, recently interviewed John Longlife. During the interview, the reporter asked John how he feels "to know that, at some point in the future, you will be infinity days old, having seen an infinite number of sunsets?" If John always answers questions truthfully and correctly, how would John answer?

4. Consider the statement "If glimmerles are flondish, then all kelevs dringle". Suppose we know that glimmerles are not flondish. Is the statement in quotes true? (Yes)(No)(Not Enough Information)