EXPLORING THE VAN HIELE LEVELS OF PROSPECTIVE MATHEMATICS TEACHERS

Carole Simard and Todd A. Grundmeier
Cal Poly, San Luis Obispo
csimard@calpoly.edu tgrundme@calpoly.edu

This research project aimed to explore the van Hiele levels of undergraduate mathematics majors, most of whom were prospective mathematics teachers. The participants were taking a proof-intensive Euclidean geometry course that relied on technology, The Geometer's Sketchpad (KCP Technologies, 2006), to help them make and prove conjectures. Data was collected from classes in consecutive years, the first with 21 participants, the second with 24 participants. Data collection included a pre- and a posttest of participants' van Hiele levels. Data analysis suggests similar results for both sets of participants in that the course had greater influence on the van Hiele levels of female participants. Results also suggest that the van Hiele test instrument used for this study operated well with university students.

Keywords: Geometry, van Hiele levels, teacher preparation, secondary

In 1957, Pierre Marie van Hiele and Dina van Hiele-Geldof, mathematics educators in the Netherlands, developed a learning model for geometry as their doctoral thesis. The van Hieles defined what are now known as *the van Hiele levels of development in geometry*, which, according to van Hiele-Geldof's thesis, as cited in Fuys, Geddes, and Tischler (1984), are hierarchical. In other words, one cannot reach a higher level before mastering the lower levels first. Furthermore, van Hiele-Geldof's dissertation indicates that the language acquired at one level differs from that at a higher level (Fuys, Geddes, & Tischler, 1984).

Altogether, there are five van Hiele levels (VHLs), originally numbered 0 through 4 by the van Hieles and renumbered 1 through 5 in most American publications. In this study, as discussed by Senk (1989), the 1 through 5 numbering scheme will be used to allow an extra level, level 0, when the first VHL, called *basic* by Pierre Marie van Hiele, is not attained. The five van Hiele levels are summarized in Table 1 from several authors' interpretations of the definitions (Burger & Shaughnessy, 1986; Mayberry, 1983; Mistretta, 2000). Exact definitions can be found in van Hiele-Geldof's doctoral thesis (Fuys et al., 1984) and a more detailed list of behaviors at each level can be found in Usiskin (1982).

Table 1. Van Hiele Levels of Development in Geometry

Level[1]	*Name*	*Characteristics*[2]
0	Pre-VHL	Student does not meet van Hiele basic level, that is, VHL 1.
1	Visualization[3]	Students visualize geometrical figures as a whole and recognize them by their particular shape.
2	Analysis	Students recognize the geometric properties of the

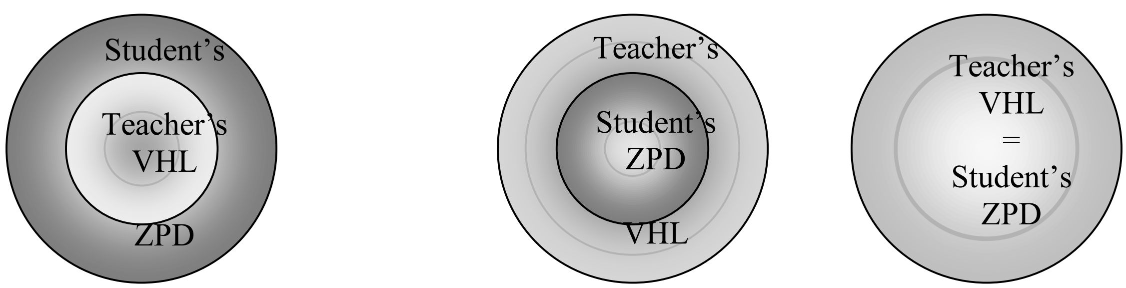
		different figures and are able to analyze the figures separately, but do not yet make connections between figures.
3	Abstraction	Students recognize relationships between figures and between properties of different figures.
4	Formal deduction	Students can write proofs and should provide justifications for each step in the proof.
5	Rigor	"… student understands the formal aspect of deduction… [and] should understand the role and necessity of indirect proof and proof by contrapositive" (Mayberry 1983, p. 59). Students can understand non-Euclidean geometries.

Notes. 1: This choice of levels was made to account for any student who did not meet the first van Hiele level.
2: The characteristics are drawn from Burger and Shaughnessy (1986), Mayberry (1983), and Mistretta (2000).
3: Level 1, visualization, corresponds to van Hiele's level 0 which he called the basic level.

As seen in past research, the van Hiele level or the level of competence in geometry of some teachers (male and female) is not at the highest level (Mayberry, 1983; Sharp, 2001; Swafford, Jones, & Thornton, 1997; van der Sandt & Nieuwoudt, 2003) thus possibly hindering the learning of geometry of some students. The conflict may arise when there is a discrepancy between the van Hiele level of the teacher and the zone of proximal development (ZPD; Vygotsky, 1987, p. 209) of the student as depicted in Figure 1.

Figure 1. ZPD VS Teacher's VHL - Potential and Ideal Situations

Potential Situation | Ideal Situations



Teacher's VHL < 5 and Student's ZPD = VHL5 | Teacher's VHL = 5 and Student's ZPD < VHL5 | Teacher's VHL = 5 and Student's ZPD = VHL5

The authors believe that during the "'sensitive period of learning'", as cited by Vygotsky (1987, p. 212), the development of a student in geometry may be limited by the teacher's own ability to complete related tasks independently. We expect that the teacher should be knowledgeable and competent in geometry and we consider that the teacher's level of knowledge and competency may be equivalent to the teacher's van Hiele level of development. So, suppose that all five van Hiele levels of development are within the reach of the students, that is, within their zone of proximal development and that geometry instruction occurs during the sensitive period. Then, if the teacher's VHL is not at level 5, we argue that the van Hiele level of development of the teacher may limit the students' learning. Refer to the potential situation depicted in Figure 1. However, if the teacher's VHL is at the highest level, students' development is not compromised. Refer to the ideal situations depicted in Figure 1. It thus becomes ideal for all prospective mathematics teachers to achieve the highest van Hiele level of development, that is, level 5 or rigor.

While it is ideal for all prospective teachers to be at van Hiele level 5, gender differences favoring males are almost twice as large in geometry as in other areas of mathematics (Leahey & Guo, 2001). However, the findings reported in the literature suggest variations in gender differences. It also suggests that the differences are mostly in spatial visualization tasks. For example, in his study of 128 high school geometry students, Battista (1990) found that "males scored significantly higher than females on spatial visualization, geometry achievement, and geometric problem solving" (p. 54). In fact, on average, males had better spatial visualization skills than females and they always scored higher than females. However, Senk and Usiskin (1983) studied high school geometric proof writing abilities, which they consider as a high-level cognitive task requiring some spatial ability. While overall geometry performance has not been analyzed by gender, they found no gender differences in achievement in geometric proof writing at the end of a one-year geometry course even though females started the year with generally less geometry knowledge.

This review of literature only found a few peer-reviewed published studies involving the van Hiele levels in geometry of preservice or inservice teachers. Among them, one study has been conducted on VHLs of preservice elementary teachers (Mayberry, 1983), one on VHLs of inservice middle-grade teachers (Swafford et al., 1997), one on the geometry content knowledge of Grade 7 prospective and inservice teachers (van der Sandt & Nieuwoudt, 2003), two on three-dimensional geometric thinking of preservice teachers (Gutiérrez, Jaime, & Fortuny, 1991; Saads & Davis, 1997), and one on developing geometric thinking of inservice K-7 teachers (Sharp, 2001), but none specifically on the VHLs of prospective secondary mathematics teachers in a proof-intensive college Euclidean geometry course..

Since the two higher VHLs involve proof-writing abilities and all prospective secondary mathematics teachers, who are mathematics majors, have completed a high school geometry course, it makes sense that instruction that focuses on proof-writing activities may lead to an increase in VHLs. Therefore, the purpose of this study was to explore the van Hiele levels of prospective mathematics teachers before and after an upper-division undergraduate Euclidean geometry course that focused on geometric proof writing and to verify the choice of Usiskin's (1982) test to explore the VHLs of undergraduate students.

The Study

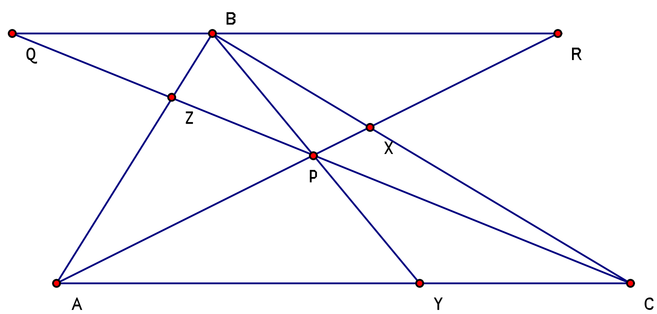
This research was conducted as a quantitative study using a pre- and posttest design with a convenience sampling as defined by Creswell (2005). Participants were enrolled in Euclidean geometry courses in consecutive years in a four-year Master's degree granting university located in the central coast of California. The first class included 21 participants, while the second class included 24 participants. Of the 45 participants, 40 were math majors and almost all were preparing to become a secondary mathematics teacher.

During both quarters the course was taught over a ten-week period, and met four times a week for fifty-minute sessions. The prerequisite is a course in methods of proof in mathematics, which focuses on instruction of logic and proof techniques. In addition, familiarity with the dynamic geometry software *The Geometer's Sketchpad* (GSP; KCP Technologies, 2006) is recommended and the geometry course is mandatory for math majors preparing to become secondary teachers.

Pedagogically, the classes could fully be described as proof-intensive (by their content) and as inquiry-oriented and technology-based (by their approach). A typical class meeting included working on instructor developed, inquiry-oriented activities using *The Geometer's Sketchpad*, and presenting and discussing solutions to the activities. A typical inquiry-oriented activity is shown in Figure 2.

Figure 2. A Typical Inquiry-Oriented Class Activity

1. (Based on Ch. 2 Activity 6 [Reynolds & Fenton, 2006]) Given with $\overleftrightarrow{QR} \parallel \overleftrightarrow{AC}$ and that $\overline{AX}, \overline{BY}$ and $\overline{CZ}$ are concurrent at *P*. Find all pairs of similar triangles in the picture below.



2. Draw an arbitrary . Let X be a point on $\overline{BC}$ and Z be a point on $\overline{AB}$ with $\overleftrightarrow{AX}$ and $\overleftrightarrow{CZ}$ intersecting at P. Construct $\overleftrightarrow{BP}$ and label its intersection with $\overline{AC}$ as Y. Now calculate $\frac{m\overline{AZ}}{m\overline{ZB}}, \frac{m\overline{BX}}{m\overline{XC}}$ and $\frac{m\overline{CY}}{m\overline{YA}}$ and find the product of the ratios. Write a conjecture about this situation and vary some points to be sure the conjecture holds. Use your work from #1 to prove the conjecture.

3. Your conjecture in #2 is called Ceva's Theorem. Write the converse of Ceva's Theorem and use GSP to decide if you believe the converse is true. If it is not true find a counterexample, if it is true prove it.

The first question in the activity is based on an exploration from the book *College Geometry Using The Geometer's Sketchpad* by Reynolds and Fenton (2006) and the goal is to lead to the proof of Ceva's theorem that is discovered in the second problem. The third problem required students to explore the converse of Ceva's theorem and to either provide a counterexample or justification of the converse. Participants used GSP throughout this exploration as a tool to make conjectures for #2 and #3 and as a tool to inform their formal proofs for #2 and #3. Participants' conjectures and formal proofs for #2 and #3 and the solution to #1 were discussed as a class through presentation of their work.

Activities, similar to the one in Figure 2, were completed in groups during the quarter using one to three class periods each. These activities were generally inquiry-oriented, that is participants were engaged in sense making activities that required them to use GSP to make and formally prove conjectures. Every activity required the proof of at least one conjecture and formal proof writing was the main form of assessment for the course. Participants presented their work regularly and presentations often resulted in rich mathematical conversations and multiple avenues to prove conjectures being explored.

Other activities expected of participants in the course included reading assignments and working on problems assigned from the book *College Geometry Using The Geometer's Sketchpad* (Reynolds & Fenton, 2006) in the first class and *Roads to Geometry* (Wallace & West, 2004) in the second class. Among all the problems assigned, problems were selected to create homework assignments completed individually and projects, made up of more challenging problems than the homework assignments, could be completed in groups of up to three participants. In addition to homework and projects, two one-hour exams were administered each quarter and the participants completed a three-hour final examination.

The Van Hiele Test

After examining several documents written by the van Hieles and describing behaviors at each van Hiele level, Usiskin (1982) developed a test instrument to assess the van Hiele level of an individual. Although this instrument was primarily devised with high school students in mind, it has been used, with permission from the authors, for this study (S. Senk, personal communication, November 19, 2007).

Whether the subjects involved would constitute an appropriate reference base for the study using Usiskin's test was considered since the subjects involved have all completed a one-year high school geometry course. However, even though the van Hiele levels have been defined while studying high school students, Pierre-Marie van Hiele, as quoted by Usiskin, believed that the highest level is "'hardly attainable in secondary teaching'" (1982, p. 12). Furhermore, Mayberry (1983), who devised her own test instrument, found that "70% of the response patterns of the students who had taken high school geometry were below Level III" (equivalent to level 4 in this study), and "only 30% were at Level III" (pp. 67-68). Therefore, it seemed adequate to use Usiskin's test with the participants in this study.

Time constraints in preparing a van Hiele level test instrument and in-class time usage were also key factors in the selection of a test instrument. Prior to choosing Usiskin's test, other test instruments were considered. For instance, Burger and Shaughnessy (1986), as cited by Jaime and Gutiérrez (1994, p. 41), developed a testing instrument to assess van Hiele levels, but its administration, through an interview, would have required more time to conduct. Mayberry's (1983) 128-item test was discarded for the same reason. Usiskin's test was readily available and it is a timed-test limited to 35 minutes.

Therefore Usiskin's (1982) 25-item questionnaire was used to measure the van Hiele level of development of the participants before and after the college geometry course.

Results

Each participant was assigned a raw score (out of 25) and a VHL similar to what Usiskin (1982) calls a "classical van Hiele level" (p. 25). Each group of five questions in Usiskin's test corresponds to a different VHL. For example, questions 1 to 5 correspond to VHL 1, questions 6 to 10 to VHL 2, and so on. Each test was examined by groups of five items. For a participant to be assigned a level, say n, at least four items must have been answered correctly at level n and at each preceding level. If a participant answered less than four questions correctly at level 1, then level 0 was assigned. The 4-item criterion was chosen over the 3-item criterion because random guessing on the test was not expected from the participants in this study and a higher mastery level was expected considering all the participants had completed a high school geometry course.

Analyzing the Test

Before analyzing the data with respect to our research purpose, we became interested in confirming our decision to use Usiskin's van Hiele test (1982) with mathematics majors. First, we noted that two participants did not meet the basic VHL and that none of the participants answered all the questions correctly on the pretest, thus allowing a measurement of growth on the posttest. Second, we implemented a Guttman scalogram analysis similar to that of Mayberry (1983) to determine whether the van Hiele levels as tested by Usiskin's test form a hierarchy. In order to do that, we created sequences in the same manner as Mayberry (1983) based on participants meeting the 4-item criterion for the set of five questions corresponding to each van Hiele level. For example, a sequence of 11101 means that a particular participant met the 4-item criterion at every van Hiele level except at VHL 4. In addition, this sequence is considered to have one error since the participant did not meet the level-4 criterion, but did meet the level-5 criterion. In this way, errors were counted as the number of 0s that appear before the last 1 in the sequence. Using the sequences generated, reproducibility factors (Reps), as described by Mayberry (1983), were calculated to give the percentage of VHLs that occur hierarchically in the sequences. The reproducibility factors were judged with the scale of 0.90 or better as suggested by Torgerson, as cited by Mayberry (1983). As can be seen in Table 2, these reproducibility factors confirm that the van Hiele levels, as tested by Usiskin's test, form a hierarchy in all cases except perhaps for male participants on the posttest in the first class. In our opinion, this Rep factor (0.867) in conjunction with the other 17 reproducibility factors over 0.90, and the results briefly mentioned previously, imply that Usiskin's van Hiele test operated adequately for our research population in terms of the hierarchical nature of the VHLs.

Table 2. Reproducibility Factors

Test	*Males*		*Females*		*Combined*	
	1st Class (12)	*2nd Class (8)*	*1st Class (9)*	*2nd Class (16)*	*1st Class (21)*	*2nd Class (24)*
Pretest	*0.96*	*0.975*	*0.91*	*0.90*	*0.9429*	*0.925*
Posttest	*0.86*	*0.925*	*0.95*	*0.9125*	*0.9048*	*0.916*
Combined Tests	*0.916*	*0.95*	*0.93*	*0.90625*	*0.9238*	*0.92083*

Once the functionality of Usiskin's test was confirmed, we turned to analyzing the data with our initial research purpose in mind. Essentially interested in exploring the van Hiele levels of prospective mathematics teachers before and after a proof-intensive geometry course, we began by analyzing the raw scores and the VHLs on the pretest and posttest. We then continued with analyses of changes in raw scores and VHLs, and changes by VHLs.

Raw Scores and van Hiele Levels

As suggested in past research, males outperformed females on the pretest in both classes (Battista, 1990; Hyde, Fennema, & Lammon, 1990; Leahey & Guo, 2001;

Halpern et al., 2007). However, females outperformed males on the posttest in the first class. The male and female average scores on the pretest and posttest can be seen in Table 3. With respect to the entire first class, the average raw score increased from 20.38 on the pre-test to 21.05 after the course and the second class saw an increase from 20.04 on the pretest to 21.71 on the posttest.

Table 3 also includes participants' average van Hiele levels as determined by the pretest and posttest and Table 4 includes counts of VHLs for each class.

Table 3. Participants' Averages of Raw Scores and van Hiele Levels

	1st Class				*2nd Class*			
	Males (12)		*Females (9)*		*Males (8)*		*Females (16)*	
Score Type	*Pre*	*Post*	*Pre*	*Post*	*Pre*	*Post*	*Pre*	*Post*
Raw Scores	21.08	20.92	19.44	21.22	20.75	22.375	19.69	21.375
Van Hiele Levels	3.67	3.25	2.56	3.44	3.75	3.875	2.875	3.67

Table 4. VHL Counts

	1st Class				*2nd Class*			
VHL	*Males (12)*		*Females (9)*		*Males (8)*		*Females (16)*	
	Pre	*Post*	*Pre*	*Post*	*Pre*	*Post*	*Pre*	*Post*
0	0	0	1	0	0	0	1	0
1	0	1	0	1	0	1	2	3
2	0	2	1	1	1	1	1	2
3	7	5	7	2	3	1	8	2
4	2	1	0	3	1	0	2	1
5	3	3	0	2	3	5	2	8

The VHL averages in Table 3 from the pretest imply that males in both classes began the course with higher VHLs than the females. Considering Table 3 and Table 4, the male VHL average from the pretest for the first class was 3.67 and all 12 male participants were at VHL 3 or higher, while the female VHL average was 2.56 and all female participants were at VHL 3 or lower. The data in the second class is slightly different: on the pretest 4 out of 8 males scored at VHL 4 or higher while 4 out of 16 females scored at VHL 4 or higher. Considering the first class, 67% of the participants started the course at VHL 3 and only 23% were at level 4 or 5. However, after the course 43% of the participants were at level 4 or 5. The second class began with 46% of the participants at level 3 and 33% at levels 4 and 5. However, after the course 58% were at level 4 or 5. As can be seen in Table 3, after the course the average female VHL was higher than the males' for participants in the first course and only slightly lower for participants in the second course.

Analyses of Changes

Counts of changes in raw scores and VHLs are summarized in Table 5. Some difference between the classes can be recognized from this data. In the first class only 3 males increased their raw score and 5 out of 12 decreased their raw score from pretest to posttest. However, in the second class 7 out of 8 males increased their raw score. These changes in raw scores translated to 4 males from the first class having a lower VHL after the course, but only 1 male from the second class has a lower VHL on the posttest. We see consistent changes between the females in both classes with the majority maintaining or improving both their raw score and VHL. This analysis will continue by class.

Table 5. Counts of Changes from Pretest to Posttest

Change Type	*1st Class*				*2nd Class*			
	Raw Scores		*van Hiele levels*		*Raw Scores*		*van Hiele levels*	
	Male	*Female*	*Male*	*Female*	*Male*	*Female*	*Male*	*Female*
Did worse	5	1	4	2	0	2	1	3
Maintained	4	0	7	1	1	3	5	5
Did better	3	8	1	6	7	11	2	8

First Class

While four male raw scores remained the same on the posttest as on the pretest and three increased, it is important to note that every male participant missed at least one item that they had answered correctly on the pretest. However, while all the females' raw scores changed from the pretest, three females did not commit any new errors on the posttest.

Examining the data from the perspective of total correct answers for males and females also helps describe the changes from pretest to posttest. As a group, the 12 male participants missed 47 items on the pretest while the 9 female participants missed 50 items on the pretest. Thus the male participants had the potential to increase their total raw score by 47 points on the posttest and in actuality their total raw score decreased by 2 points. By contrast, the 9 female participants increased their total raw score by 16 of the 50 potential points. These changes in raw score correspond to the male VHLs changing by –0.42 or almost the equivalent of half a VHL on average, while the female VHLs changed by 0.88, the equivalent of almost one full VHL.

Interestingly, two participants, one male and one female, increased their raw score on the van Hiele test from pretest to posttest, but saw their VHL decrease. Both participants were at a van Hiele level of 3 after the pretest. The male's raw score increased from 18 to 21 yet his VHL decreased to 1, while the female's raw score increased from 19 to 21 and her VHL decreased to 2. For the male participant, the VHL decrease was a result of answering two questions incorrectly at VHL 2 that he had answered correctly on the pretest. The increase in raw score for this participant can be attributed to answering four questions correctly at VHL 5 and one additional question at VHL 3 that he did not answer correctly on the pretest. The decrease in VHL for the female participant is due to answering one additional question incorrectly at VHL 3 that she answered correctly on the pretest. Her increase in raw score is attributed to two questions at VHL 4 and one question at VHL 2 that were answered incorrectly on the pretest.

Second Class

While all the males improved or maintained their raw scores from pretest to posttest only three male participants did not commit any new errors on the posttest. Fourteen of the sixteen females improved or maintained their raw score from the pretest to posttest and similarly only three females did not commit any new errors on the posttest. Examining the data from the perspective of total correct answers for males and females also helps describe the changes from pretest to posttest. As a group, the 8 male participants missed 34 items on the pretest while the 16 female participants missed 85 items on the pretest. Thus the male participants had the potential to increase their total raw score by 34 points on the posttest and in actuality their total raw score increased by 13 points. Similarly, the 16 female participants increased their total raw score by 27 of the 85 potential points. These changes in raw score caused the male VHL average to increase from 3.75 to 3.875 and the female VHL average to increase from 2.875 to 3.67, almost the equivalent of one VHL.

Interestingly, there were 3 participants, 1 male and 2 females, in the second class who actually increased their raw score from pretest to posttest and decreased their VHL on the posttest. The male was at level 3 on the pretest and at level 1 on the posttest, but his raw score increased from 19 on the pretest to 21 on the posttest. The decrease in VHL was due to the male incorrectly answering a question at VHL 2 that he had previously answered correctly on the pretest. The increase in raw score was due to answering two questions correctly on the posttest that he had answered incorrectly on the pretest, at VHLs 3 and 5.

The first female increased her raw score from 12 on the pretest to 15 on the posttest, while her VHL decreased from 2 to 1. As for the male, this decrease was due to the female answering one question at VHL 2 incorrectly on the posttest that she had answered correctly on the pretest. She also answered six questions correctly on the posttest that she had missed on the pretest (one at VHL 1, one at VHL 3, three at VHL 4, and one at VHL 5) and answered two additional questions incorrectly that she had formerly answered correctly, at VHLs 3 and 5. The other female scored 18 and 22, respectively, on the pretest and on the posttest, and saw her VHL decrease from level 3 to level 2. This was due to this female participant incorrectly answering one question in the VHL 3 section of the posttest that she had previously answered correctly. The change in her score was due to correctly answering six questions that she had not answered on the pretest (one at VHL 2, three at VHL 4, and two at VHL 5), and answering two questions incorrectly that she had previously answered correctly at VHLs 3 and 5.

Performance at Each van Hiele Level

The fact that some participants in both classes had an increase in raw score accompanied by a decrease in van Hiele level suggested that it may be interesting to look at participants' performance by groups of questions relating to each van Hiele level. Tables 6 through 9 show the performance on each set of questions for the male and female participants respectively by class.

Table 6. Analysis at Each van Hiele Level – Males in 1st Class

	Van Hiele Level									
	1		*2*		*3*		*4*		*5*	
Statistic	*Pre*	*Post*	*Pre*	*Post*	*Pre*	*Post*	*Pre*	*Post*	*Pre*	*Post*
Average	5	4.83	5	4.58	4.67	4.33	3.25	3.25	3.25	3.92
4-item criterion met	12	12	12	11	12	10	5	4	5	9
Paired *t*-	0.1661		0.0538		0.2199		1.0		0.0433	

test	Not stat.sign.	Not quite stat.sign.	Not stat.sign.	Not stat.sign.	Stat.sign.

Table 7. Analysis at Each van Hiele Level – Males in 2nd Class

	Van Hiele Level									
	1		*2*		*3*		*4*		*5*	
Statistic	*Pre*	*Post*	*Pre*	*Post*	*Pre*	*Post*	*Pre*	*Post*	*Pre*	*Post*
Average	5	5	4.62	4.75	4.0	4.63	3.63	3.89	3.5	4.25
4-item criterion met	8	8	8	8	7	7	4	6	4	6
Paired *t*-test	1.0 Not stat.sign.		0.6168 Not stat.sign.		0.2600 Not stat.sign.		0.6875 Not stat.sign.		0.1759 Not stat.sign.	

Tables 6 and 7 show data for the males from both classes. The data suggests that both before and after the course the male participants performed well on the questions related to van Hiele levels 1 to 3 with slight decreases at each level on the posttest in the first class. The data also suggest little change at VHL 4, none in the first class and a slight increase for the second class. The analysis suggests a statistically significant change in the performance at level 5 for the first class and a similar change for the second class, although not statistically significant. VHL 5 is "rigor" and relates to the proof-based nature of the course. In the first class, the average number of questions correct at VHL 5 increased from 3.25 on the pretest to 3.92 on the posttest with only 5 participants meeting the 4-item criterion on the pretest and 9 participants meeting the 4-item criterion on the posttest. A similar change at level 5 is evident in the second class as well with the average score increasing from 3.5 to 4.25 and the number of participants meeting the 4-item criterion increasing from 4 to 6.

The data in Tables 8 and 9 suggest a similar situation for the female participants; however, both classes saw either no change or slight increases in the averages at van Hiele levels 1 through 3. Unlike the male participants, the female participants showed no change at level 5, but statistically significant change at VHL 4. VHL 4 relates to formal deduction and mathematical proof, again a part of the focus of the proof-intensive geometry course. For the first class the number of correct answers increased on average by more than one (out of five) question at level 4. Also, no females met the 4-item criterion at level 4 on the pretest, but six out of nine met the criterion on the posttest. For the second class there was an increase of one and one-half questions on average from pretest to posttest with only 3 participants meeting the 4-item criterion on the pretest and 11 meeting the 4-item criterion on the posttest.

Table 8. Analysis at Each van Hiele Level – Females in 1st Class

	Van Hiele Level									
	1		*2*		*3*		*4*		*5*	
Statistic	*Pre*	*Post*	*Pre*	*Post*	*Pre*	*Post*	*Pre*	*Post*	*Pre*	*Post*
Average	4.67	5	4.78	4.77	4.22	4.56	2.44	3.56	3.33	3.33
4-item criterion met	8	9	9	8	8	8	0	6	3	3
Paired *t*-	0.1950		1.0		0.2815		0.0304		1.0	

test	Not stat.sign.	Not stat.sign.	Not stat.sign.	stat.sign.	Not stat.sign.

Table 9. Analysis at Each van Hiele Level – Females in 2nd Class

	Van Hiele Level									
	1		*2*		*3*		*4*		*5*	
Statistic	*Pre*	*Post*	*Pre*	*Post*	*Pre*	*Post*	*Pre*	*Post*	*Pre*	*Post*
Average	4.69	4.75	4.5	4.56	4.44	4.44	2.5	4.0	3.5	3.69
4-item criterion met	15	16	14	14	13	13	3	11	8	10
Paired *t*-test	0.7855 Not stat.sign.		1.0 Not stat.sign.		1.0 Not stat.sign.		0.0031 stat.sign.		0.6384 Not stat.sign.	

Discussion

Although change patterns in raw scores and VHLs between the pretest and the posttest are slightly different for the two classes, there are some interesting parallels when examining participants' performance by VHL. As shown in Tables 3 and 4, both raw score and VHL averages decreased for males (0.16 and 0.42 respectively) and increased for females (1.78 and 0.88 respectively) in the first class. To the contrary, both raw score and VHL averages increased for both males (2.125 and 0.125 respectively) and females (1.685 and 0.795 respectively) in the second class.

Looking at the combined male and female results after a ten-week proof-intensive geometry course, the influence on the van Hiele levels of development of prospective mathematics teachers cannot be stated as statistically significant. The difference in the average van Hiele level before and after the course is only 0.14 for the first class and 0.5 for the second class.

However, some influence has been observed on the VHLs by gender, more notably for females. Comparing pre- and post-VHLs shows an average increase of 0.88 for females with 6 out of 9 females improving their VHL in the first class and an average increase of 0.795 with 8 out of 16 improving their VHL in the second class. For males in the first class the VHL average decreased by 0.42 and only 1 out of 12 males improved his VHL. In the second class the VHL average increased by 0.125 and only 2 out of 8 males increased their VHL.

It should be noted that results related to the female participants were statistically significant for both classes. A paired *t*-test of the female raw scores showed that the change between pre- and posttest results is statistically very significant in the first class (□=0.0035, □=*0.05*) and similarly in the second class (p =0.0032, α =*0.05*). The overall change in female VHLs is statistically significant in the first class (p =0.0379, α =*0.05*) only when the results of one candidate (female participant number 6) are removed from the female population. Female participant 6 is the participant whose raw score increased and van Hiele level decreased. In addition, paired *t*-tests ran on female VHLs for each VHL separately showed a statistically significant change (p =0.0304, α =*0.05*) for the first class at level 4 and similarly (p =0.031, α =*0.05*) for the second class.

Pre- and posttest results for males and combined genders in raw scores and VHLs have not produced any statistically significant changes. However, paired *t*-tests ran on male VHLs for each VHL separately showed a statistical significance (p =0.0433, α =*0.05*) at level 5 for the first class only.

The unexpected result of the male participants, an overall decrease of almost one half

VHL in the first class and a minimal increase of one-eighth of a VHL in the second class, is hard to explain. Perhaps the test instrument was not adequately testing the levels at which some male participants were working? Usiskin (1982) mentioned that his test instrument could not be very well defined to test the highest level since the van Hiele theory itself is not defined at the highest level in enough detail. Van Hiele, as cited by Usiskin (p. 14), expressed doubts on the theory at the highest level. However, these arguments seem to be contradicted by the performance of the male participants on the level 5 questions on the posttest. Male participants in both classes averaged almost one more question correct on the posttest at level 5 and the number of males meeting the 4-item criterion at level 5 changed from five on the pretest to nine on the posttest in the first class and from four on the pretest to six on the posttest in the second class. It is possible that the male participants in the first class hurried through the "easy" questions on the posttest, but, like the male participants in the second class, demonstrated by their performance at VHL 5 that they are able to use "rigor" in their geometric reasoning. This ability can likely be attributed to the focus on mathematical proofs and the rigor required during the course.

While the participants had to brush up on general geometric concepts more independently (levels 1 through 3), the course in this study focused on proof-writing abilities (levels 4 and 5). The results show that the females had significantly more room for improvement at VHLs 1-3 from the pretest to the posttest. Also, based on the change in female performance at level 4 and male performance at level 5, it may be plausible to assume that Usiskin's test was nevertheless operating reasonably well at the two higher VHLs. The gain obtained by the male participants at level 5 was likely offset by mistakes that were not made on the pretest at VHLs 1-3. However, the female participants who had more room for improvement at lower VHLs and demonstrated significant improvement at level 4 might have made their gain more easily and accurately measurable by Usiskin's test. For Usiskin, this was not an issue, as he knew his participants would mostly be working at one of the first three levels.

Finally, some findings in this research are consistent with the findings of prior research. For instance, the results on the pretest are consistent with Leahey's and Guo's (2001) findings where male students did better than female students in geometry at the end of high school. As in Senk and Usiskin (1983), females and males performed (almost) equally well in geometric proof writing at the end of a geometry course. Additionally, as in this study where, in general, the females' performance has improved substantially, Ferrini-Mundy and Tarte, as cited by Leahey and Guo, found that girls' performance improved after learning spatial-related strategies. This may correspond to the use of *The Geometer's Sketchpad* in this course as a tool to visualize geometric objects and form conjectures about specific situations.

Furthermore, in a study of middle school students, Kenney-Benson, Pomerantz, Ryan, and Patrick (2006, p. 16) found that girls are more mastery oriented than boys are. Similarly, in a study of college students, Zusman, Knox, and Lieberman (2005) found that "women [are] significantly more likely than men to attend class, take notes, buy/read the book" (p. 625). Perhaps these findings were also present here and the female students were more dedicated to their learning. This may partially explain the significant increase in female scores and VHLs.

Conclusion

In this specific situation the instructor entered the course with the hope that both males and females would perform equally well and all participants would be operating at the highest VHL by the end of the course. Some of the results of this research were surprising and somewhat unexpected. The males' advantage on the pretest was consistent with past research, but unexpectedly, females finished with more significant increases in raw scores and VHLs. Additionally, one might have expected that after such an intensive course in geometric proof-writing, the van Hiele levels would have been at their highest for every participant. Surprisingly, it was not the case: males' VHL averages were 3.25 and 3.875 and females' was 3.44 and 3.67 at the end of the study. These averages are not at the highest VHL. Combining the two classes, only 3 out of 20 males improved their VHL while 14 out of 25 females improved theirs. Furthermore, only 18 out of the 45 participants are at the highest VHL, which is only 40% of the participants in the two classes. This allows one to wonder why. Could it be that, as forewarned by Dr. Senk (S. Senk, personal communication, November 19, 2007), the test instrument was not perfectly suited for this investigation? Alternatively, could it be that the higher the VHL, the higher the difficulty to reach it (Usiskin, 1982)? In turn, it makes one wonder what would have to be taught, and how, for prospective secondary mathematics teachers to achieve the highest van Hiele level.

After reviewing the results of this study, one legitimate question arises: are these prospective mathematics teachers well prepared to teach geometry at the high school level even if none of them scored perfectly on Usiskin's van Hiele test? Swafford, Jones, and Thornton (1997) showed that after increasing the geometry content knowledge of middle-grade (4-8) teachers as well as increasing their awareness of how students learn geometry, their teaching practices, when giving geometry lessons, were substantially and positively modified (pp. 480-482). Consequently, their students were exposed to more geometry activities and concepts, and their teaching performance moved closer to the ideal situations depicted in Figure 1.

It is easy to argue that the geometry content knowledge of the females in this study was increased and it is possible that this will influence how they teach geometry in the future. However, based on the overall results from the van Hiele test, a similar argument is hard to make about the male participants. It is argued, though, that based on the performance of the males on the posttest at level 5, the proof-intensive geometry course changed their content knowledge as well.

Finally, for teacher educators, the outcome of this research may raise questions about the geometry preparation that will best serve prospective teachers in their future practice. The course described here, while inquiry-oriented and proof-intensive, focused on content beyond the level of a typical high school Euclidean geometry course. It is possible that this choice in content coverage influenced the performance at VHL levels 1 through 3. Maybe a more axiomatic development of Euclidean geometry from scratch would better serve prospective teachers and help move them towards VHL 5. Alternatively, maybe the most likely scenario is that an undergraduate course develops prospective teachers thinking up to VHL 4 with the hope that, as they continue to learn beyond their undergraduate education, they will reach VHL 5.

References

Battista, M. T. (1990). Spatial visualization and gender differencs in high school geometry. *Journal for Research in Mathematics Education, 21*(1), 47-60.

Burger, W. F., and Shaughnessy, J. M. (1986). Characterizing the van Hiele levels of development in geometry. *Journal for Research in Mathematics Education*. *17*(1), 31-48.

Creswell, J. W. (2005). *Educational research: Planning, conducting, and evaluating quantitative and qualitative research* (second ed.). Upper Saddle River, New Jersey: Pearson Prentice Hall.

Fuys, D., Geddes, D., and Tischler, R., Eds. (1984). *The didactics of geometry in the lowest class of secondary school*. New York: New York. (Original work by Dina van Hiele-Geldof published 1957).

Gutiérrez, A., Jaime, A., & Fortuny, J. M. (1991). An alternative paradigm to evaluate the acquisition of the van Hiele levels. *Journal for Research in Mathematics Education. 22(3)*. 237-251.

Halpern, D. F., Benbow, C. P., Geary, D. C., Gur, R. C., Hyde, J. S., & Gernsbacher, M. A. (2007, August). The science of sex differences in science and mathematics. *Psychological Science in the Public Interest, 8*(1), 1-51.

Hyde, J. S., Fennema, E., & Lammon, S. J. (1990). Gender differences in mathematics performance: A meta-analysis. *Psychological Bulletin, 107*(2), 139-155. Retrieved May 31, 2008, from EBSCO Host database.

Jaime, A., and Gutiérrez, A. (1994). A model of test design to assess the van Hiele levels. *Proceedings of the 18th PME Conference, 3*, 41-48.

KCP Technologies (2006). The Geometer's Sketchpad (Student Ed., Version 4) [Computer software]. Emeryville, CA: Key Curriculum Press.

Kenney-Benson, G., Pomerantz, E., Ryan, A., and Patrick, H. (2006). Sex differences in math performance: The role of children's approach to schoolwork. *Developmental Psychology*, *42*(1), 11-26.

Leahey, E., and Guo, G. (2001). Gender differences in mathematical trajectories. *Social Forces, 80*(2), 713-732.

Mayberry, J. (1983). The van Hiele levels of geometric thought in undergraduate preservice teachers. *Journal for Research in Mathematics Education,* 14(1), 58-69.

Mistretta, R. (2000). Enhancing geometric reasoning. *Adolescence, 35*(138), 365-379.

Reynolds, B. E., and Fenton, W. E. (2006). *College geometry using the Geometer's Sketchpad* (preliminary ed.). Emeryville, CA: Key College Publishing.

Saads, S., and Davis, G. (1997). Spatial abilities, van Hiele levels, & language use in three dimensional geometry. *Proceedings of the 21st PME Conference 4*. 104-111.

Senk, S. (1989). Van Hiele levels and achievement in writing geometry proofs. *Journal for Research in Mathematics Education*, 20(3), 309-321.

Senk, S., and Usiskin, Z. (1983). Geometry proof writing: A new view of sex differences in mathematics ability. *American Journal of Education*. *91*(2), 187-201.

Sharp, J. M. (2001). Distance education: A powerful medium for developing teachers' geometric thinking. *Journal of Computers in Mathematics and Science Teaching*. *20*(2), 199-219.

Swafford, J. O., Jones, G. A., and Thornton, C. A. (1997). Increased knowledge in geometry and instructional practice. *Journal for Research in Mathematics Education*, 28(4), 467-483.

Usiskin, Z. (1982). *Van Hiele levels and achievement in secondary school geometry* (CDASSG Project Report, 231 pages

Van der Sandt, S., and Nieuwoudt, H. D. (2003). Grade 7 teachers' and prospective teachers' content knowledge of geometry. *South African Journal of Education. 23(3)*, 199-205.

Vygotsky, L. S. (1987). *Problems of general psycholgy, Thinking and speech* (N. Minnick, Trans.) New York: Plenum Press. (Original work published 1934).
Wallace, E. C., and West, S. F. (2004). *Roads to geometry* (3rd edition). Upper Saddle River, NJ: Pearson Education.
Zusman, M., Knox, D., Lieberman, M. (2005). Gender differences in reactions to college course requirements or "Why females are better students". *College Student Journal, 39*(4), 621-626.