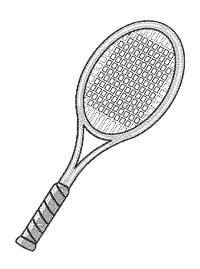
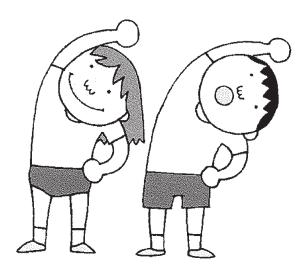
Is stretching beneficial to tennis players?

A statistical analysis of tennis and stretching





Introduction

Statement of Task

The game of tennis has always played a large role in my life, that is ever since I was ten years old and saw my first professional tennis match, and at Wimbledon no less. Ever since then I have loved the sport and just can't seem to stop playing it. When I'm not playing tennis, I am thinking about playing tennis, and therefore when my classmates and I were told that to pick a topic for a math project I immediately thought of tennis. Now we all know that stretching before doing physical activity prevents injury, or at least that is what the health professionals tell us, but does stretching before playing increase performance. I personally always feel better after I have stretched and feel as though I am better prepared to perform. When the muscles are loose and warm I believe that the number of unforced errors and reaction times will decrease. The question is to what extent? For this project I am going to be testing the ways in which stretching benefits tennis players. More specifically I will be attempting to answer three questions.

- 1) Is there a correlation between whether a player stretches before a match and the number of unforced errors (defined below) that the player makes during the match?
- 2) Does stretching before a match speed up reaction time, specifically, the time it takes to get to a drop shot (defined below)?
- 3) Is it more beneficial for advanced players to stretch before playing than for beginners?

An **unforced error** is committed when the player has had time to prepare and position him or herself to get the ball back in play but then makes an error. This is a shot that the player would normally get back into play. The real keys here are time and position. When the opponent takes away time by hitting the prior shot with extra pace this can result in a forced error. Also, when the opponent forces the player out of position with placement (depth and/or angle) this can result in a forced error.

A **Drop Shot** is a delicate shot that barely clears the net and falls short in the opponent's court.

Topspin (used later in the project) is the forward rotation of the ball in flight.

Plan

In order to collect my data I will be observing the play of numerous different tennis players of varying skill and ability. I will observe them while they play full tennis matches and write down various statistics of their play, such as number of unforced errors and number of seconds it takes them to reach drop shots. I will have each student play eight matches where no stretching is performed beforehand and then I will have each student play another eight matches where stretching *is* performed beforehand. After

A2

enough data has been obtained I will carry out numerous mathematical processes in order to answer the questions stated previously in the Task.

More specifically I will analyze the difference in data between days when stretching was not done and days when stretching was done. The differences in unforced errors and reaction times between the different days will be examined. The data will be examined from numerous angles and ability of the players will be taken into account.

I would also like to not here, as the reader may also notice later in the data tables, that there is only data for 15 sets. This is because for each student I threw out one set based on the student's own opinion on which set was the set most unrepresentative of their ability. As any experienced tennis player knows you never play the same every day, but there are some days when it is more than that, when you can't seem to make a single shot in play and nothing goes right. Due to this extreme deviation from normal ability that players experience every so often I made the decision to drop one set for each student.

Stretching Routine

Now I would like to give the stretching routine that each student will be going through before each of the sets where stretching is to be done. This procedure is a standard stretching routine and has been approved by a number of health professionals who are knowledgeable on the subject.

Start out lying on your back with both knees bent and your feet flat on the floor. Take one foot off the floor and bring your knee toward your chest, pulling it gently with both hands. Release that leg, return your foot to the floor, and repeat on the opposite leg. Every stretch I describe may be done as many as three times for up to 30 seconds each time.

Now place one leg flat on the floor, with the other knee bent and your foot flat. Bring the bent knee toward your chest as in the first exercise. You want to keep the leg on the floor as flat as possible, with the back of your knee flat. Repeat on the opposite leg.

Start out as you did in the previous stretch, one leg flat, one knee bent. Keeping the flat leg straight, raise it off the floor and bring it toward you as close as you can, without lifting your hips off the floor or pushing off of the foot that's on the floor. Hold this stretch at the highest point. Lower your leg all the way to the floor, while holding it as straight as possible. Repeat on the other side.

Still lying on your back, go back to the first position, both knees bent, both feet flat. Hold your arms out to the side with your palms facing up. Let both knees drop to one side while turning your head the opposite way. Don't try to adjust your legs to be perfectly on top of one another. Roll back up to the center and right on through to the opposite side. Hold this for as long as you like. Take a couple of deep breaths. Doesn't that feel good?

Back in starting position as in the previous stretch. Cross one ankle over the opposite knee. Bring that knee toward you, taking your foot off of the floor. Hold your thigh

behind your knee, and you should feel a stretch in your hip and butt area on the leg that's crossed over the one you are holding. Undo this pretzel, and repeat on the other leg.

Roll over onto your stomach, with your head down and your arms stretched out in front of you. Keep one arm stretched out in front of you while you reach back with one hand and grab the ankle on that side. If you had to move your knee away from your body to grab that ankle, move it back to the center once you have it. If you're too tight to be able to grab your ankle, grab your sock and pull that. This is a great quad stretch, and it really shows you how tight those muscles are, too. Go back and forth three times like you do with the other stretches.

Get up off the floor. (Finally!) Stand at arm's length from a wall with your arms straight and your hands on the wall at shoulder height, palms flat. Take a small step back with each foot, keeping both feet parallel and your knees straight, and lean toward the wall, keeping your heels flat on the floor. Play with this stretch. You can do one leg at a time, also. If you bend your knees while keeping your heels flat, you'll feel that stretch move down your calves and work into your Achilles tendon as well.

Information and Measurement

Results without Stretching

	Set Number	Student 1 (beginner)	Student 2 (intermediate)	Student 3 (advanced)
	Set 1	36	34	11
	Set 2	38	30	15
	Set 3	25	35	19
	Set 4	32	18	12
	Set 5	39	22	14
	Set 6	29	25	16
Number of	Set 7	36	28	13
Unforced	Set 8	38	32	18
Errors (per	Set 9	40	29	20
set of tennis)	Set 10	35	30	9
	Set 11	26	34	17
	Set 12	35	32	15
	Set 13	36	27	17
	Set 14	38	25	10
	Set 15	34	28	12
	Time 1	2.3	2.0	1.6
Reaction	Time 2	2.5	2.7	1.5
Times (in	Time 3	2.8	3.0	1.9
seconds)	Time 4	4.6	2.6	2.1
	Time 5	3.5	4.0	2.3

Results with Stretching

	Set Number	Student 1 (beginner)	Student 2 (intermediate)	Student 3 (advanced)
	Set 1	38	35	9
	Set 2	39	32	11
	Set 3	27	30	12
	Set 4	36	33	15
	Set 5	40	26	11
	Set 6	25	25	13
Number of	Set 7	28	26	8
Unforced	Set 8	32	21	14
Errors (per	Set 9	22	24	16
set of tennis)	Set 10	29	28	17
	Set 11	30	31	12
	Set 12	33	19	12
	Set 13	35	24	13
	Set 14	39	27	15
	Set 15	37	20	15
A.	Time 1	2.2	2.4	1.5
Reaction	Time 2	2.9	2.6	1.6
Times (in	Time 3	1.8	2.1	1.1
seconds)	Time 4	1.9	1.7	1.3
	Time 5	2.6	2.1	1.4

Mathematical Processes

The first set of mathematical processes that I put the data through were the calculations referred to as "One Variable Statistics". The ways in which I obtained all of the figures can be found below under the section "Calculations".

I also graphed all the data that is graph able from the One Variable Statistics calculations in box plots (or box-and-whisker diagram). Box plots are a convenient way to graph the "five-number summary". Furthermore box plots are able to visually show different types of populations, without any assumptions of the statistical distribution, which is why standard deviation is not included. The amounts of space between the different parts of the box help indicate variance, skew and identify outliers.

Calculations

Calculation for Mean: The mean is the most common measure of central tendency and therefore is a way to show the center of a distribution of data. In order to calculate the mean (or arithmetic average) number of unforced errors and reaction times of each student I found the sum of all the given elements and divided by the total number of elements.

For example: Student 1 was found to have an average of 34.47 unforced errors without stretching and the calculation for this figure follows.

$$\frac{36+38+25+32+39+29+36+38+40+35+26+35+36+38+34}{15} = 34.7$$

Standard Deviation:

$$S_n = \sqrt{\frac{\sum_{i=1}^k f_i (x_i - \overline{x})^2}{n}}$$

Where f_i is the frequency, X_i is the mark, $\overline{\mathcal{I}}$ is the mean and n is the sum of the frequencies.

Using this formula we are able to find the standard deviation which is the square root of the variance. Standard deviation, measures the spread or dispersion around the median of a data set. It is the most widely-used measure of spread.

However, there is another way to find standard deviation other than through the use of a formula and that is by using a frequency table. An example of which follows for the unforced errors for Student 1.

G1

Ranges	Mark	Frequency	Mark - Mean (34.47)	Frequency (Mark - Mean) ²
(24-26]	25	2	-9.47	179.36
(26-28]	27	0	-7.47	0
(28-30]	29	1	-5.47	29.921
(30-32]	31	1	-3.47	12.041
(32-34]	33	1	-1.47	2.1609
(34-36]	35	5	0.53	1.4045
(36-38]	37	3	2,53	19.203
(38-40]	39	2	4.53	41.042
		Total: 15		Total: 285.1335

To find the Variance of this data you take the total of the last column and divide it by the total number of data items. In this case it would be $\underline{285.1335} = 19.00$

After finding the variance it is quite simple to find the standard deviation. Just take the square root of the variance since earlier we squared the deviations.

Standard deviation = $\sqrt{19.00} = 4.36$

Median: The median is used to show a central tendency just as the mean is. It is just another way of showing where the data is centered and in contrast to the mean is not affected by outliers. The median is found by putting the data in order from least to greatest and the number in the middle of the set is the median. If there is an even number of elements in the set then the two middle numbers are added together, and the sum is divided by two in order to find the average.

Lower and Upper Quartiles: To find the lower quartile all you need to do is find the number that is the 25th percentile of the data. For example with Student 1 to find the lower quartile of the number of unforced errors without stretching you would put the data in order from least to greatest and find the number that is halfway to the median. If there are an even number of values below the median then you would just find the average of the two numbers in the middle, just as was done for the median. For the upper quartile the same method is used except the number at the 75th percentile of the data is found.

Min Lower Quartile Median Upper Quartile Max

Inter Quartile Range: The inter quartile range is merely the difference between the upper quartile and lower quartile. It is used in the same way as the range; that is it is a measure of the spread of the data and shows if it is relatively close together or far apart.

In the case of the data above the Inter Quartile Range (IQR) would be 38 - 32 = 6

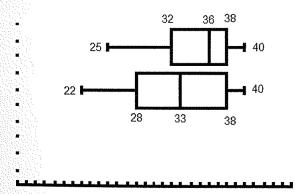
Minimum and Maximum x values: The smallest and largest values in the set of data respectively. By looking at these two numbers you are able to decide whether there are outliers in the data and knowledge of this can tell you whether or not your mean is a good measure of central tendency.

Note all box plots have the "without stretching" plot on top and the "with stretching" plot on bottom. Also the numbers (from left to right) represent the Minimum x value, Lower Quartile, Median, Upper Quartile, and Maximum x value

Student 1

One Variable Statistics (for unforced errors)

	Without Stretching	With Stretching
Mean	34.47	32.67
Standard Deviation	4.36	5.45
Lower Quartile	32	28
Median	36	33
Upper Quartile	38	38
Inter Quartile Range	6	10
Minimum x value	25	22
Maximum x value	40	40

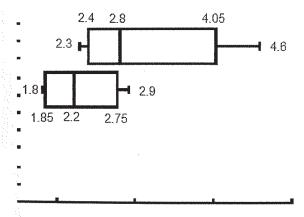


As you can see in this table many of the figures are relatively close to each other, and there isn't a great deal of change between the data for "without stretching" (WoS) and "with stretching" (WS). The Mean score for WOS is slightly higher than that for WS showing that the scores *did* improve after the student went through a stretching routine. Next there is the standard deviation which for WOS is slightly lower than for WS. This signifies that the data for WOS is more closely centered towards the median than it is for the WS data. Although in tennis it is important that your unforced errors remain relatively stable and is therefore a good thing to have a low standard deviation, it is more important

that the average number is low because this is the actual number of errors you are making. In this case the average was decreased by about two, which may not seem like a lot but in tennis every single point counts and even two points can make a significant difference in the outcome of a match. As we can see from the box plot above the majority of the data for WS is below the median and this shows that stretching did in fact lower the number of unforced errors. Also since both maximums both upper quartiles are the same, while the median for WS is 3 lower we can draw the conclusion that stretching may not stop you from having sets where you make a lot of unforced errors, but it will increase the number of sets where you make a low number.

One Variable Statistics (for reaction times)

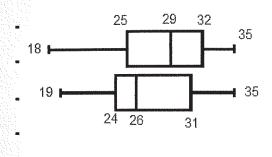
	Without Stretching	With Stretching
Mean	3.14	2.28
Standard Deviation	.836	.417
Lower Quartile	2.4	1.85
Median	2.8	2.2
Upper Quartile	4.05	2.75
Inter Quartile Range	1.65	.9
Minimum x value	2.3	1.8
Maximum x value	4.6	2.9



For reaction times the data is much more conclusive than it was for unforced errors. As we can see from the table the mean, the standard deviation, the median, both quartiles, the inter quartile range and the minimum and maximum values were all decreased from the WOS data to the WS data. We can see that not only is the data more centered around one point but also the average time it took to reach the ball was also decreased. As we can see from the box plot the data for WOS is much more sporadic, the outliers are much more defined. In tennis it is very important to reach drop shots quickly not only to get the shot back in play but also because the earlier you get to the ball the higher up the ball will be. This means that you don't have to hit the ball up into the air as much thus taking away the opportunity for an easy kill shot from your opponent.

Student 2
One Variable Statistics (for unforced errors)

	Without Stretching	With Stretching
Mean	28.6	26.73
Standard Deviation	4.56	4.64
Lower Quartile	25	24
Median	29	26
Upper Quartile	32	31
Inter Quartile Range	7	7
Minimum x value	18	19
Maximum x value	35	35

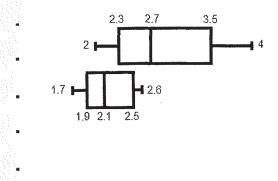


Unfortunately Student 2 did not fare quite as well as Student 1 as far as unforced errors go. Although the mean number of unforced errors was still decreased by about two, which is the most important thing, the overall effect of stretching for this student was not as strong. For instance the set where the lowest number of unforced errors was achieved was without stretching instead of with stretching as would be expected. There could be many reasons for this inconsistency with what was hypothesized to happen. The first reason could be that quite often it is at the intermediate level where a player makes the most drastic changes to his or her game. For instance it is at this level that the service grip is changed from the "flat grip" to the "continental grip" which is a much more advanced grip. Although this change is quite beneficial to the player it can take a great deal of time to get used to, let alone to perfect. Another explanation could be that at the intermediate level the player usually has relatively solid technique as far as strokes go and can hit the ball with topspin as well as speed but has not yet learned when to hit their hardest shots. Therefore the intermediate player will often go for their biggest shot when it is not a good opportunity and will therefore commit an unforced error as a result.

D3

One Variable Statistics (for reaction times)

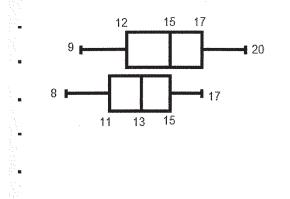
	Without Stretching	With Stretching
Mean	2.86	2.18
Standard Deviation	.656	.306
Lower Quartile	2.3	1.9
Median	2.7	2.1
Upper Quartile	3.5	2.5
Inter Quartile Range	1.2	.6
Minimum x value	2	1.7
Maximum x value	4	2.6



As we saw before the results for reaction times are much more convincing than those for unforced errors. Although the reaction time wasn't decreased by as much as it was for Student 1 it was still decreased and the other data shows a dramatic difference between WOS and WS. For instance the standard deviation was decreased by over 50% which is an impressive change. This shows how much closer together the data is for WS in comparison to WOS and we can again see this using the box plots. Also the minimum score for WS is lower than for WOS. Despite the fact that it is not a great deal lower does change that it did decrease. Therefore once again the stretching benefited the tennis player. It may not seem like it did significantly because the mean was only decreased by .68 seconds, it actually did. The overall affect of the stretching benefited the player in numerous small ways, and tennis is all about the details. Every little bit helps. Furthermore it is possible that Student 2 just isn't in as good of physical shape as Student 1 was which would justify the not so significant change in average time. Taking this into account further highlights the dramatic decrease in standard deviation, and we can conclude that many more values were

Student 3
One Variable Statistics (for unforced errors)

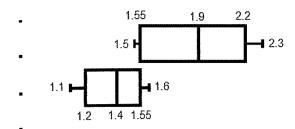
	Without Stretching	With Stretching
Mean	14.53	12.87
Standard Deviation	3.22	2.45
Lower Quartile	12	11
Median	15	13
Upper Quartile	17	15
Inter Quartile Range	5	4
Minimum x value	9	8
Maximum x value	20	17



Once again we see that the calculations done for unforced errors show that the stretching did benefit the player, just not to a great extent. The mean and median numbers of unforced errors were both decreased as was the case with both proceeding students. Also the standard deviation along with the IQR was decreased. Here we can see that as a result of stretching the student had fewer sets that had a very high number of unforced errors. This becomes increasingly more important as you become a better tennis player and you can afford less and less to have those fluke sets that are mentioned in the statement of Plan section at the beginning of the project. Because of this we can conclude that stretching benefited this player. We can also guess as to why there wasn't a more significant change. This could be because an advanced player such as Student 3 will not be affected as much by outside factors such as wind and sunlight, and also stretching. They will hopefully play at about the same level (in the short term) all the time.

One Variable Statistics (for reaction times)

	Without Stretching	With Stretching
Mean	1.88	1.38
Standard Deviation	.299	.172
Lower Quartile	1.55	1.2
Median	1.9	1.4
Upper Quartile	2.2	1.55
Inter Quartile Range	.65	.35
Minimum x value	1.5	1.1
Maximum x value	2.3	1.6



Here we see that there was a dramatic change between WOS and WS. All of the figures were decreased although some, such as standard deviation, were decreased more than others. The reason for the median and mean not being decreased as much as with some of the other results is almost the opposite of Student 2. That is, Student 3 is in quite good physical shape and therefore stretching is only going to make him so much faster. Once you reach a certain level, an asymptote if you will, it is not possible to go any lower. But since Student 3 is in better physical shape and is a faster runner, it becomes much more important from a safety standpoint from him or her to stretch since the chance for injury is heightened.

Chi - Square Test

From the calculations of the data and box plots above I have come to the conclusion that the unforced errors and reaction times were decreased as a result of stretching. The question now is: to what extent were they decreased. In order to answer this question I have turned to the Chi – Square Test.

According to the

math pre-calculus textbook: the basis for the chi-square test (X² test) is that if the observed values are very different to their expected

G1

D3

In order to find X² values it is necessary to make contingency tables of the observed values, and then use these to find the expected values and finally use all of these values together to find the X² value. To find the final value the equation used is:

$$\chi^2 = \sum_{\substack{(Obs - Exp)^2 \\ Exp}}$$

Student 1

Observed Values

	(24 – 32] unforced errors	(32 - 40] unforced errors	Total:
Without Stretching	4	11	15
With Stretching	7	8	15
Total:	11	19	30

**************************************	(24 – 32] unforced errors	(32 - 40] unforced errors	Total:
Without Stretching	$\frac{11 \times 15}{30} = 5.5$	$\frac{19 \times 15}{30} = 9.5$	15
With Stretching	$\frac{11 \times 15}{30} = 5.5$	$\frac{19 \times 15}{30} = 9.5$	15
Total:	11	19	30

$$X^{2} = (\underbrace{4 - 5.5)^{2}}_{5.5} + \underbrace{(11 - 9.5)^{2}}_{9.5} + \underbrace{(7 - 5.5)^{2}}_{5.5} + \underbrace{(8 - 9.5)^{2}}_{9.5} = 1.292$$

Critical Value = 3.841 Degrees of freedom = 1 **C4**

From these results we can conclude that for Student 1 the number of unforced errors a player makes is independent of whether or not he stretches before a match. This means that despite the fact that the unforced errors were decreased it was a significant enough change to prove that they are not independent.

Student 2

Expected Values

	(17 - 26] unforced	(26 – 35] unforced	Total:
	errors	errors	
Without Stretching	6	9	15
With Stretching	6	9	15
Total:	12	18	30

 X^2 Value = 2.22 Critical Value = 3.841 Degrees of freedom = 1

Here once again we find out that the change was not significant enough to prove that they are not independent. However, this time the X^2 value was much closer to the critical value showing that the data was closer to rejecting the null hypothesis.

Student 3

Expected Values

	(8 – 14] unforced	(14 - 20] unforced	Total:
is.	errors	errors	
Without Stretching	8.5	6.5	15
With Stretching	8.5	6.5	15
Total:	17	13	30

 X^2 Value = 1.22 Critical Value = 3.841 Degrees of freedom = 1

Here for the third and final time we see that we must accept the null hypothesis that the factors are independent because the X^2 value did not exceed the critical value. Also the difference between the X^2 and the critical value increased from Student 2 to Student 3. This could be taken to mean (as was suggested above) that Student 3 as the advanced player will not be affected as much by things like stretching and will play relatively the same no matter what.

D3

Interpretation of Results

Through many mathematical processes and calculations I have come to the conclusion that stretching benefits tennis players in many different ways. It slightly decreases a player's unforced errors and more than slightly decreases their reaction time, as can be seen within the One Variable Statistics section of the Mathematical Processes. Also it causes the unforced errors and reaction times to become more consistent and fall within a closer range, as can be seen through the reduction of most of the standard deviations. However, despite these observations according to the X² test the unforced errors were not decreased enough to reject the null hypothesis that the factors are independent. Also I would like to take this opportunity to note why there were no X² calculations for Reaction Times. This was because not enough data was collected to fulfill the requirements for the contingency tables. That is each box in the table should have a frequency of no less than 4 or the results are no acceptable. Unfortunately this would have occurred if the X² test had been done with Reaction Times. I would also like to note that this is no fault of the procedure. This is because in junior tennis matches there are far fewer drop shots than in matches of more advanced college and adult players. Even in the 15 sets that each student played, recording five drop shots was at times a challenge.

Due to all of this I have come to the final conclusion that going through an accepted stretching routine will benefit a tennis player in many ways. First it will slightly decrease unforced errors. Next it will to a medium extent decrease reaction times. Thirdly it will make the data for both of these things fall closer to the median. But finally that stretching does not decrease unforced errors enough to cause us to reject the null hypothesis, and come to the conclusion that the two factors are not independent. Unfortunately it is unknown whether the X^2 Test would have led me to reject the null hypothesis for reaction times since I did not have enough data to run the X^2 Test with respect to Reaction Times.

Validity

Despite the fact that I used all relevant mathematical processes to their utmost potential I was unable to reach the conclusion that I hoped to reach that stretching significantly decreases unforced errors and reaction times. I used One Variable Statistics to find the mean, median, standard deviation, lower and upper quartiles, inter quartile range, and minimum and maximum points of data. This was appropriate because all together these figures give a great representation of the data as a whole and can show the central tendencies of the data and how they are dispersed. Along with this the box plots of both WOS and WS were graphed one over the other in order to visually represent the data and to better enable analysis of the data. This also was highly appropriate because with the two sets of data one over the other comparison of the two became quite easy. Finally I used the X² Test to show whether or not the two factors of: number of unforced errors and whether or not a player stretches before playing tennis are independent. This was appropriate because if the X² values had been above the critical values this would

E1

have meant that they two factors are not independent and that stretching has a significant effect on the number of unforced errors.

There could be many reasons why I did not reach my hypothesized conclusion even with the use of these sophisticated mathematical processes, since as an experienced tennis player I believe my hypothesis to be true. The first could simply be that just because I believe it to be true for me does not mean it is true for others. Obviously I was not one of the students who was observed and for the students that were observed stretching might not have as large of an affect as it does for me. Also I could just be wrong that stretching affects the performance of an athlete. Maybe it's only use is to prevent injury. Personally I think the biggest reason I did not reach the conclusions I was hoping for was because I did not collect enough data, especially with regards to Reaction Times. More data would have enabled much more accurate calculations, especially for the X² test. I believe that more data for the unforced errors would have resulted in lower average numbers of unforced errors after stretching and therefore higher X² values. This in turn would have led me to reject the null hypothesis and come to the conclusion that the two factors are not independent. If I had had more data for the Reaction Times the same thing would have occurred as with unforced errors. The average reaction times would have decreased and the X² values would have gotten bigger.

As a wise teacher or mine once said "an assignment is never finished, it is only due" I think this applies here because although I will be turning this project in I will always continue to think about it. Every time I play tennis I will be constantly watching to see how players who stretch beforehand perform in comparison with how they perform when no stretching is done. I am still confident in my belief that stretching does greatly benefit the tennis player, I just haven't quite come to my final decision of how.

Sources

http://diet.ivillage.com/workouts/wstretch/0,,7b6,00.html

http://sportsline.com/tennis/story/6145343

http://www.tennis.com/yourgame/asktheeditors/asktheeditors.aspx?id=1432