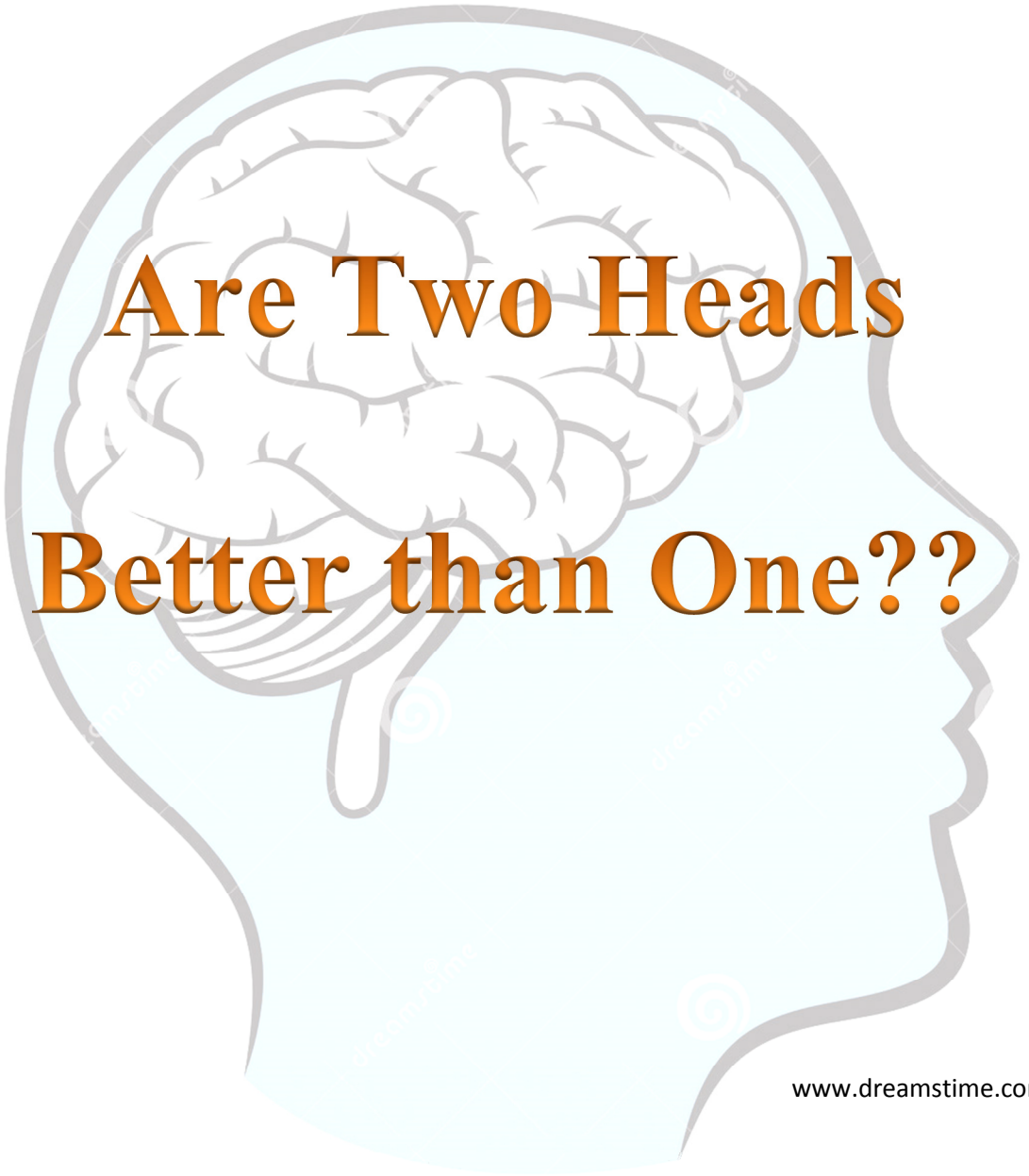


Are Two Heads Better than One??

www.dreamstime.com

Table of Contents

I. Introduction

The general question we hope to answer from this experiment is: Do people learn better in groups than by themselves? If so, how many people in a group is the most productive group size? We plan to answer these questions by finding if a relationship exists between the number of people who are working on a problem in a group and the time it takes to solve a Sudoku problem. We hypothesize that there will be a negative correlation between the students per group and the amount of time it takes to solve the Sudoku puzzle because the more people in a group, the more total contribution there should be.

II. Why did we choose to study this?

Many teachers struggle with the issue of the best way students learn. Obviously, it is difficult for the teacher to please every student, but with the results from this study, the teachers might be able to see if the majority of students work better in groups or by themselves. In addition to this, if the results show which size group is the most productive, then teachers will not have to guess and they can have set table groups of the perfect amount of students in each. According to the CPM Education Program, studies have been conducted to show that students learn ideas more deeply when they talk about ideas with a group of their peers[1]. If the data from this experiment supports our hypothesis then we can support the results of this 2013 study. Our experiment is only testing the speed of students solving a Sudoku puzzle, so we will not have solid evidence that students *learn* better in groups until more research is done, but the results of this study will be a useful reference for teachers at our high school.

III. Experimental Procedure

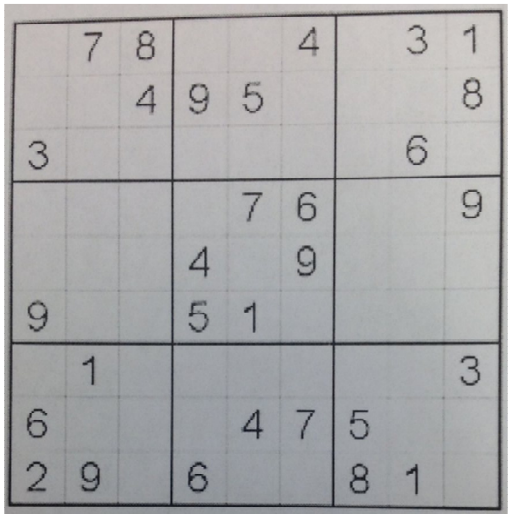
We performed our experiment during 7th period after the AP Statistics Exam on two mixed grade level English classes and one sophomore gym class whose teachers volunteered to let us use

[1]CPM Program. *Executive Summary of 2013 CPM Research Report*. Rep. Print.

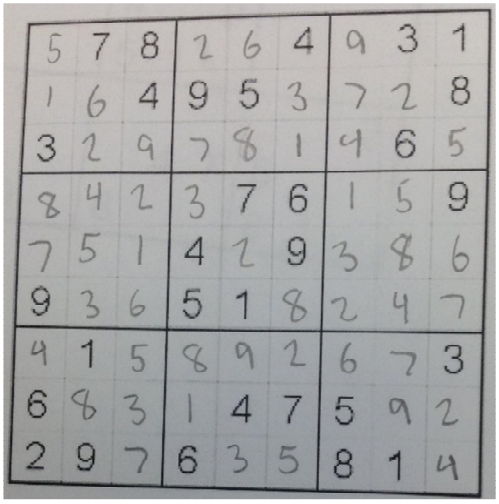
them. When we arrived in these classes, we read to them the following instructions before proceeding to hand out the Sudokus:

"Thank you for participating in this study for our Statistics class. Just as a preliminary to our study, we're going to explain how to do a Sudoku puzzle for those of you who don't know: The objective is to fill a 9x9 grid with digits so that each column, each row and each of the nine 3x3 sub-grids that compose the grid contain all of the digits from 1 to 9. We will be randomly putting you guys into groups and then asking you to complete a Sudoku as a group as quickly as you can. Now we're going to hand out the Sudokus. Please leave them turned over until we say go, and raise your hand when you as a group have all finished. You can talk to each other as much as you want if you have other people in your group. Remember, the group Sudoku puzzle must be entirely filled out before you raise your hands."

Uncompleted Sudoku



Completed Sudoku



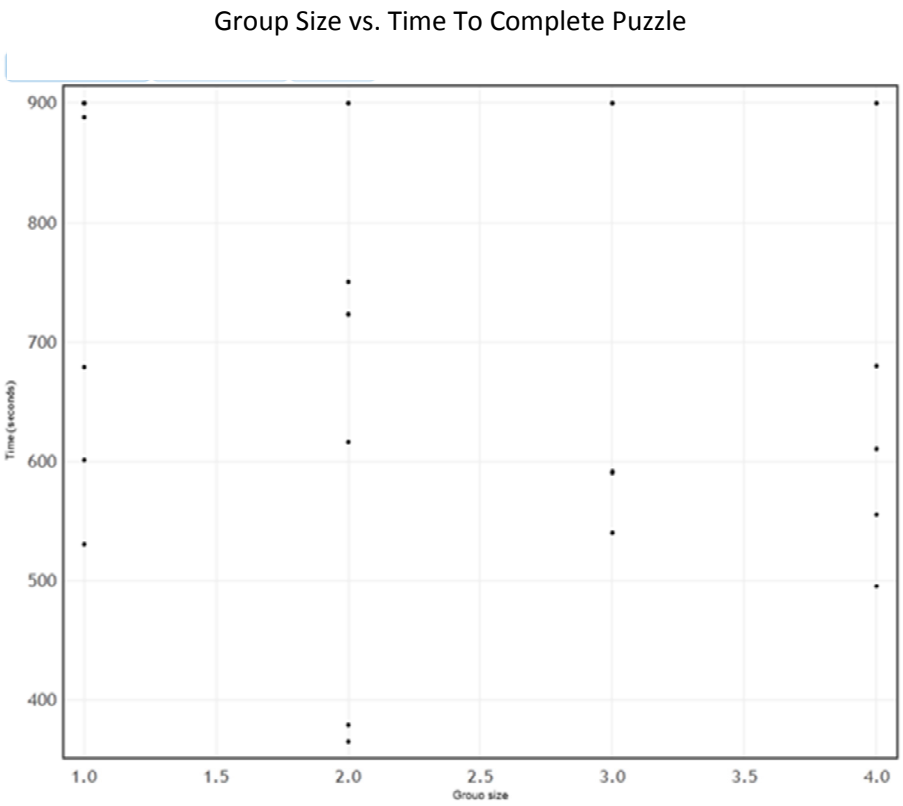
We randomly assigned the students in the classes whose teachers volunteered to groups of different sizes. We obtained class lists ahead of time and numbered the students alphabetically. We then randomly selected four numbers and assigned these students to a group of four, and repeated this process until we had

the desired numbers of groups of four and repeated this process for group sizes three, two and one. This process ensured that our results could imply causation because we randomly assigned subjects to treatment groups, but because we did not take a random sample of the population, we cannot extrapolate any of our results to any population.

We then handed out the Sudokus, one to each group, making sure that they were all of the "easy" level. We said "go," and then started a timer. We watched them work to make sure that there was no interaction between the groups, and we recorded the times when each group finished. We then collected the puzzles, thanked the teacher for letting us use some of his/her valuable class time, and left.

IV. Analysis of Data

For a summary of the results of our study, see Table 1 in the appendix. It displays the times it took groups of different sizes ("4.1" is a group of 4) to complete the puzzle, as well as observations that we took during the process of the study. See the graph at right (and see the appendix for an expanded version) of the scatterplot that represents the data that we collected. This graph has a very slight negative linear association, but the groups whose results were greater than 15 minutes



(which we entered in as 900 seconds, or 15 minutes exactly) were observed in all four groups.

Our first choice of inference procedure was a hypothesis test for slope of the regression line. This inference procedure is appropriate for testing to find whether or not there is a relationship between group size and time it takes to solve a Sudoku puzzle because our alternative hypothesis is that there will be a decrease in time as the group size increases, creating a negative slope.

State:

We will conduct a hypothesis test for slope of the regression line that represents the relationship between group size and time it takes to solve a Sudoku puzzle.

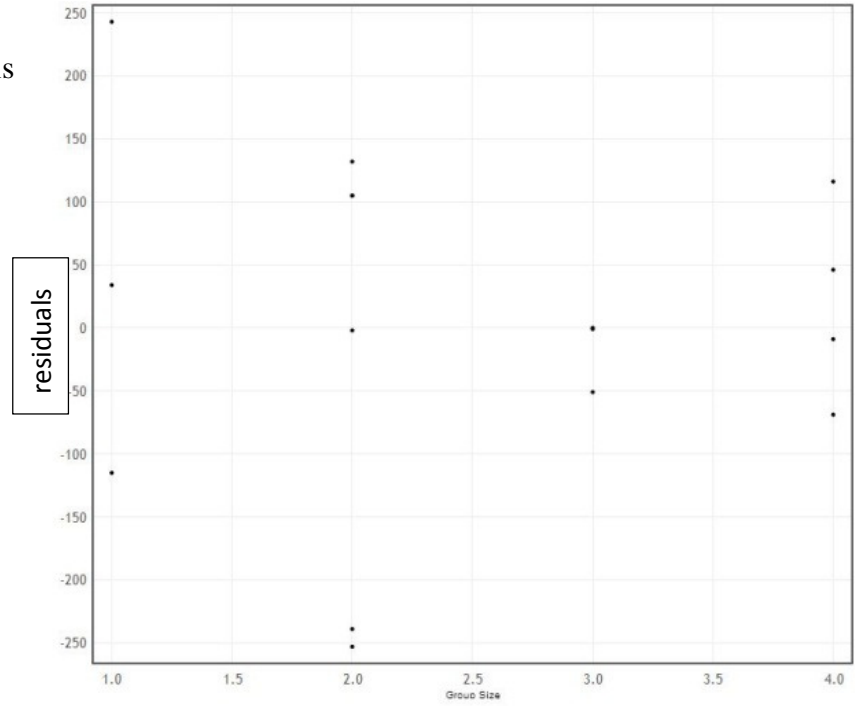
We will perform a linear regression analysis to test the hypotheses:

$\alpha = 0.05$

H_o: $\beta = 0$ where β = the slope of the true regression line comparing

H_A: $\beta < 0$ group size and time to complete the puzzle

(Our alternate hypothesis is one-sided because we believe that there is a *negative* correlation between the size of a group and how quickly it works; larger groups work faster.)



Plan:

Conditions:

~Linear? The residual plot (presented at right) appears to be random.

~Independent? We ensured that the different groups are independent of one another by watching as they complete the puzzle.

~Normal? The distributions for each group size are normally distributed about the mean time for that group size.

~Equal Variance? The residual plot is distributed randomly for each group size and has approximately equal variances.

~Random? The students are randomly assigned to groups of different sizes.

All conditions are met.

Do:

y=803.625+(-30.344)x

t=-1.03495

df=28

p-value=0.1548

If H_0 is true and we perform many samples then there is a 20.45% chance of obtaining results as or more extreme. For every one increase in group size, there is a -25.412 second increase in time it takes to solve the Sudoku puzzle. The y-intercept is irrelevant because we did not test group sizes of zero. The correlation coefficient is -.192, which means there is a weak, negative, non-linear association.

Summary Statistics Switch Variables

Statistic	Group size	Time (seconds)
Mean	2.300	733.833
Standard Deviation	1.119	176.865
Sample Size	30	
Correlation	-0.192	
Slope	-30.344	
Intercept	803.625	

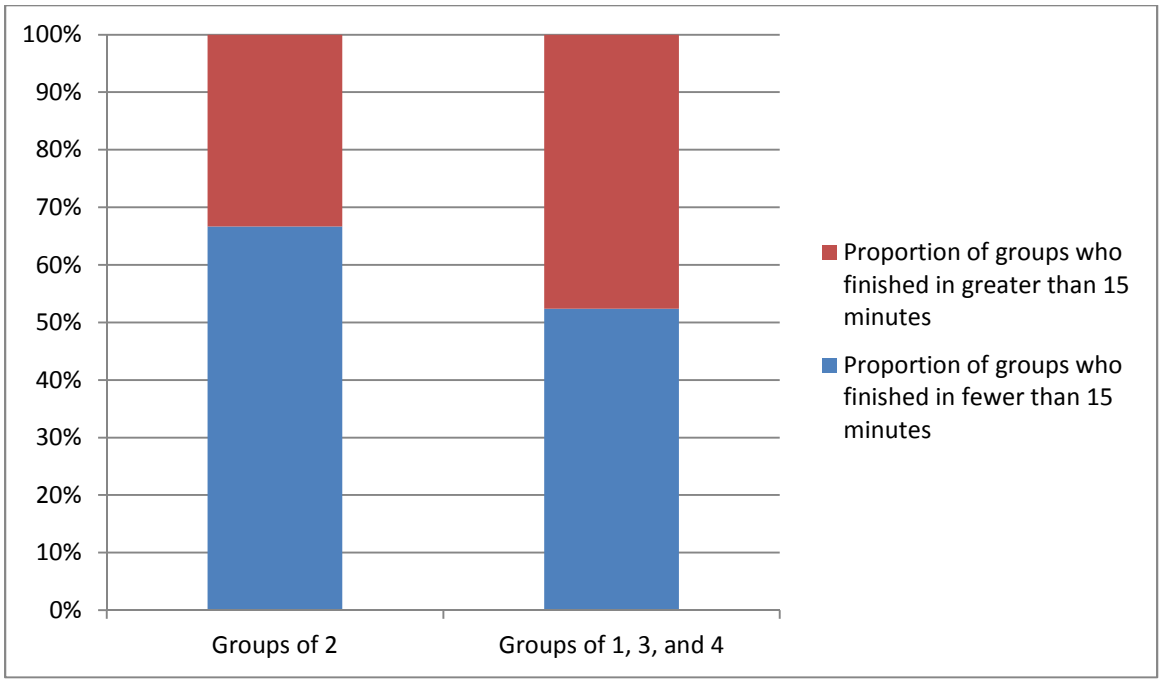
Conclude:

The p-value is greater than our α, so we fail to reject H_o and do not have convincing evidence to conclude that the slope of the true regression line is negative.

Since we did not obtain enough evidence to conclude a negative population slope when we used a hypothesis test for slope, we decided to use a two-proportion z-test to compare the proportion of groups of two that finished the Sudoku puzzle in under fifteen minutes and the proportion of

groups of one, three, and four that finished the Sudoku puzzle in under fifteen minutes. As you can see in the segmented bar graph below, for our study this proportion was much greater for groups of two than it was for groups of 1, 3, and 4. We decided to combine these other three group sizes because we noticed that people seemed to be working less efficiently in all of them, and that the proportion who finished in all three of these group sizes was less than it was for groups of two (See Table 1 in the Appendix).

Proportion That Finished in Fewer than 15 Minutes for Different Group Sizes



State: We will perform a two-proportion z-test to test the hypotheses:

H_o: $p_2 = p_a$

H_a: $p_2 > p_a$

$\alpha = 0.05$

p_2= The true proportion of groups of two people finishing in under fifteen minutes.

p_a= The true proportion of groups of sizes one, three, and four finishing the Sudoku puzzle in under fifteen minutes.

Plan:

Conditions:

~Random? We randomly assigned people to group sizes.

~Normal? $n_1\hat{p}$ (pooled) = (9)(17/30) = 5.1 $n_1\hat{q}$ (pooled) = (9)(1-(17/30)) = 3.9

$n_2\hat{p}$ (pooled) = (21)(17/30) = 11.9 $n_2\hat{q}$ (pooled) = (21)(1-(17-30)) = 9.1

Three of these are greater than 5, but $n_1\hat{q}$ is not. This concerns us (if we had been planning to perform a two-proportion z-test, we would have ensured that the size of each group was large enough), but for the time being, we will proceed with caution. We will carry out the test as if the distribution was approximately Normal.

~Independent? The students selected were not related to one another; the time of one subject did not have any effect on the time of another.

All conditions are met.

Do:

$\hat{p}_2$= 6/9 = .667

$\hat{p}_a$= 11/21= .524

z=.7236

p-value= .2347

If H_0 is true and we perform many samples then there is a 23.47% chance of obtaining results as or more extreme.

Conclude:

The p-value is greater than the α, so we fail to reject H_o and conclude that there is not convincing evidence in support of the hypothesis that the proportion of groups of two people finishing in under fifteen minutes is greater than the proportion of groups of sizes one, three, and four finishing the Sudoku puzzle in under fifteen minutes.

V. Things we would do differently

About halfway through our project, we decided to split up so that we could more accurately observe the groups of different sizes. One of us took several groups out into a common area, while the other stayed with the rest of the class in the classroom. We hoped that this would decrease the distractions that our subjects experienced. Another thing that we learned after experimenting on our first class was that we should check for absences at the beginning of the class, and then reassign people who were in groups of one to fill in the places of the missing students. This didn't, however, affect our data in the first class; we didn't reassign students, so our group sizes were smaller, but the data wasn't affected. In the future, we would also try to find a more generic and less well-known puzzle, since we discovered that a huge factor in the time to complete the puzzle was whether or not the subject(s) were familiar with the strategies that aid in the completion of Sudokus. We would find a task that more accurately assesses subject's problem-solving abilities.

VI. Areas for further study

Our data had a dishearteningly low correlation. We attributed this to the many lurking variables that could confound our results. Even though we randomly assigned subjects to group sizes, it was difficult to ensure that groups would be representative. Some possible confounding variables included:

- Intelligence of the subjects (smarter students could complete the Sudokus more quickly)
- Gender of the subjects (we noticed that girls tend to communicate more than boys)
- Age of students (older students were more likely to have come in contact with Sudoku puzzles before)
- Prior experience with Sudokus (there are strategies that can be learned with practice)

- Ethnicity of subjects (international students worked well alone, but became nearly silent when working with one or more other people)

We now realize that many of these variables could be interesting topics for further study. Some experimental questions could include:

- Do more intelligent students work more effectively in groups? (block by math level)
- Does gender affect the ability to work effectively in groups? (block by gender)
- Does grade affect the ability to work effectively in groups? (block by grade)
- Does language/culture affect the ability to work effectively in groups? (block by primary language, and allow international students to communicate in their native tongues)

VII. Conclusion

Based on the results from the hypothesis test on the slope of the regression line and the two proportion z-test, there is not significant evidence to conclude a relationship between group size and time it takes students to solve a Sudoku puzzle. Our results do not support our original hypothesis that there is a negative correlation between the students per group and the amount of time it takes to solve the Sudoku puzzle. The lack of significant evidence found in this study may be due to confounding variables, such as the ones listed above. These confounding variables may have a bigger influence on the ability to solve a Sudoku puzzle than group size. Further studies should be done to give solid evidence on this topic, but based on this experiment only, teachers may want to assume that students work best in different ways and there is no “perfect” group size.

VIII. Appendix of Tables and Graphs

Table 1

Does group size affect a group's efficiency?		
Group Number	Time (seconds)	Observations
1.1	>900	came in late, but focused quietly; persistent, but almost no progress
1.2	530	left with middle 2 boxes near end
1.3	679	worked quickly
1.4	>900	didn't listen to instructions, but picking it up quickly; very systematic--top to bottom;
		confused--"gave up" at 10 minutes
1.5	601	very focused
1.6	888	focused; getting distracted/frustrated at about 11 minutes
1.7	>900	talking to herself; depressed and frustrated--might not know how to do a sudoku
1.8	>900	didn't understand instructions (international); trying to cheat...?; writing all over
		paper, except in boxes; finally picked it up at about 10 minutes
1.9	>900	slowly but surely; stuck with middle two boxes; "gave up" at 13 minutes
2.1	>900	one person didn't know how to do sudoku-->other explained; new person tried to
		understand and help; old person let new use pencil and try it for himself; new
		wanted to give up, but kept going
2.2	616	one explained it to the other, but then mostly did it by herself; made a mistake and
		sort of gave up
2.3	365	silent concentration; quick; one person wrote numbers, the other gave suggestions
2.4	723	discussion; made mistake-->discussing the problem
2.5	379	talking only when necessary
2.6	723	discussion; girl wrote
2.7	>900	discussion; both had pencils
2.8	750	worked together; one person did most of the work-->had to explain it to the other,
		but the other tried to learn and contribute
2.9	>900	tried even though goofed off a little; discussed together; got stuck on two middle
		boxes
3.1	>900	2 people doing all the work-->third just distracting; unsure of how to complete it;
		didn't know that 2 of the same number can't be in the same box, even though I told
		them-->started over at 4:30; correcting each other-->erasing a lot; talked the
		whole time; finally gave up and third person tried--"took turns" working; filling them
		in randomly until there is a conflict
3.2	>900	having trouble getting started; all just staring at it
3.4	591	talking, giggling; discussing possibilities; one member is lost-->gave up, stopped
		trying but still has some input
3.5	540	some discussion; girls did most of the work; one girl writer; one girl gave suggestions; third
		member left to get a pencil
3.6	>900	started ahead, girls did most work; both girls had pencils; after other groups finished, they
		got bored; one member looked at another group's paper
3.7	590	one member looked distracted; discussed and worked together; one member told another to
		look for a specific number
4.1	495	all girls are into it, but one boy is watching from the side but some input; stuck on middle
		row--very common; got quieter as time went on
4.2	>900	discussing possibilities, but not making much progress; talking a lot--all about numbers; one
		person explaining to others; working slower at 12 minutes; confusing each other, but focused?
4.3	610	one member talked the most and did most of the work; another helped; one person writing;
		briefly stuck with 2 middle boxes, but figured it out quickly
4.4	555	2 members talking, others not; one member has some silent input; stuck with middle 2 boxes
4.5	>900	laughing; two people did most of the work; two boys did not do any work
4.6	680	staring at it blankly for awhile; concentrated-->discussed numbers

Graph 1: Group Size vs. Time to Complete the Sudoku