Introduction:

Hypothesis:

Fitt's Law dictates interline spacing must play some role in the reading speed and comprehension. That being said, I hypothesize that it plays a minimal role. The main factors against speed and comprehension are how much the user is distracted by outside influences. If the user is just reading straight through without having to find their place, line spacing will play less of a roles since after the first few lines they will just get into the rhythm of where to expect the next.

Now, if the lines are so far apart that it is hard to remember what happened on the previous line, or if the lines are so close together it becomes hard to read the words themselves, then interline spacing will play a larger role at least in comprehension, but most likely only a minimal difference in speed.

Therefore, at the extremes interline spacing will have noticeable measurable effect, but only when it reaches the point when the lines no longer seem to be related to each other due to how far apart they are or so close they are illegible. In other words, I believe Fitt's Law will have minimal effects when it comes to interline.

If you were to graph the relation of interline spacing against speed, it would most likely be fairly flat until the extremes are reached, but then legibility becomes a factor. If anything, the graph would form a shallow bell curve, becoming nearly constant between standard single spacing up until at least double spacing if not further. I feel that comprehension may be a bit less shallow because the user is fighting against focusing on reading versus the material, but it should be nearly constant in the same areas. (see fig1.)

Types of Users:

I will test only college students due to their habit of quickly reading, comprehending, and answering questions on a given material. The other reason is that almost every student at this point in time has been raised with computers. They therefore are comfortable reading off the medium.

They will also be required to know little to nothing about the material before the test begins so I can truly focus in on what they comprehend and retain from the test.

After the final design…

Design of Experiment:

The basic setup of the experiment is that a user will be asked to read through some text straight through, avoiding rushing as well as rereading as

much as possible, and then answering a series of questions followed by a small questionnaire. Each user will be read the same set of instructions and will be reading off of the same laptop screen at 1024x768 for consistency in font, general size, and preventing user from having the text zoomed. The only obvious variable not taken into account is user's distance from the screen, but I felt this plays such a minimal role and helps the user feel more comfortable that where the laptop is placed is left up to the user.
Execution:

Text:
The text I have chosen involves a general overview of various aspects of threading as related to computer science. Therefore, the users have also been limited further to people who have limited to no knowledge about threads and specifically multithreading. The text is broken up into 5 paragraphs on the following topics: general definition, time splicing, the difference between a thread and a process, concurrency issues, how concurrency issues are solved, and a brief conclusion.

Each paragraph is broken up by an extra line break instead of indenting the first line. The only other thing of notice is that there is an example block/code block. I felt this was a necessary graphical choice since many technical documents such as the one written use this format to show which part is information versus example. I also felt this would not create any problems since all users would be reading the same block so it would not change the outcomes.

Font wise I went with a serif for no particular reason since most books and texts on technical subjects are written in some form of a serif. The code block was made monospace for contrast purposes. Again, since every user had to read the same text, these variables are constant.

There were two major choices I made with the text. The first of which was to limit the width of the page. I do not remember which study I read or where I learned it, but I have found that how far a user has to read horizontally can slow them down significantly. I therefore set the width to a constant and comfortable 800 pixels which is a web standard.[1] This standard was mainly due to monitor resolution, but I have found through personal anecdote and observing users over the years that no one wants to read text that spans the width of the screen.

The other choice I made which I feel strongly about is a change in the experiment spec. *Each user reads only a single page of text and then answers questions.* The user reads one 500 word essay which when spaced

1 http://www.247webpages.com/01/pages/page-measures-01.html

reasonably (single and 1.5) follows the rule that you should “limit page length to one or two screenfuls.”[2] This should shorten how long it takes the user to read the text, make the test less intimidating, and provide a more accurate reading on how many words per second the user reads.

I chose to do this to attempt to eliminate what I dub reading fatigue where the user has reached the end only to realize that they have to keep going. I feared that this would skew the results due to creating the false impression that certain interline spacings would cause the user to slow when in fact it was just due to the length of the test.

The other issue I saw with providing the user with multiple interline spacings was it formed no consistency in design of the experiment. The experiment would be actively working against the user in addition to providing many pages, the user would be subjected to multiple layouts. Also, by presenting the user with multiple layouts it may be easy for them to guess the purpose of the experiment of the test and that may skew the data as well.

The tradeoff of course is that more users would be required to provide accurate statistics. This would require 5 times the users minimum (one for each test) to provide meaningful results. A single text also prevents an order bias. By order bias, I mean that if the test was arranged in such a way that each user read double space last, fatigue may cause that reading to take longer than all the others. In order to prevent an order bias from occurring, multiple variations of the same test would have to be created, shuffling the formats and/or topics. This would not only require more users to test, but the decision between shuffling topics versus only the format would arise and possibly introduce more variables.

I felt that the shortened time, more standard and less intimidating design, and simpler test was a fair trade off against requiring more users. It should allow me to get people who have already taken one test more easily since it appears shorter even if it is not. In the end, since the data is only meaningful on a large number of users anyways, I felt that by eliminating as many variables as possible would be vital. And in this case, I believe it means having a single text.

The text can be viewed at http://users.wpi.edu/~agoebel/hcip2/testc.html (replace testc with a-e to see different variations.)

Quiz:
The quiz itself took the form of 15 questions: 5 multiple choice, 5 true/false, and 5 short answer questions. Each question was worth 1 point and could only be correct or incorrect. Some leeway was given on the short answer if

2 http://websitehelpers.com/design/

they demonstrated enough of a grasp of the concept based on predetermined criteria.

I felt it was necessary to quiz the user in a variety of methods in order to see whether or not multiple choice would influence retention and comprehension. This gave me a better sense of who actually understood the material. I also allowed the users to opt out of the short answer and multiple choice by specifying that “don’t know” was a legitimate option and they should avoid guessing. I assumed people would guess at the true or false even if there was an option to opt out.

As for the material covered, I attempted to cover the minutia such as definitions as well as comprehension through asking questions about the given example (Short answer 1 and multiple choice 4 respectively.) There were also some general comprehension questions based around understanding what was going on in the text versus extrapolating fact from what was provided (Short answer 5 and true/false 1 respectively.)

I tested the questions on two laymans and one experienced programmer to make sure I had covered most of the material provided.

Questionnaire:

There were two questionnaires in this experiment. The first took the form of three screening questions. The first was I did not take any CS majors unless they were freshman. I then asked if they had taken any CS courses here or elsewhere where they would learn a fair amount of theory. Finally I asked if they were familiar with threads or multithreading. If they had not more than a simple concept of what threads are which, they were allowed to take the test.

The second questionnaire asked a couple personal questions and how they felt the test went, as well as allowing them to provide further comments. The two personal questions asked were in regards to their major/profession as well as how often they used a computer. These were asked to determine how familiar the subject was with computers in general. Since it was impossible to see how these factors would apply until the final study was done, I did not use them as part of my screening process.

The next portion asked them to rate on a scale of 1-5 the difficulty they had with the material in question, the quiz itself, as well as the interline spacing and font choice. I added font choice to determine if they felt that their performance was not due to the variable being tested.

Reference sidebar

At no point did I record age, race, gender, grade, or anything non-related to computer usage. I did not feel that these should or would have any effect on my study. Especially since all users should be between the ages of 18-22.

Tools Required:

In the design of this experiment, I found I simply needed a web browser to read the text as well as for the quizzes. I also needed to find a decent remailer so I could manually grade the quizzes. As far as data analysis goes, as of right now Excel 2007 appears to suit my needs well enough except to calculate linear regression, for that I'm using a TI83+.

¬ *Determine the design of your experiment;*
A series of non fiction articles, probably excerpts from Wikipedia, and a 20 question quiz at the end. Each reading will be no longer than a single mouse scroll, and should be displayed on the page. I will keep the span of the page within a legible 700 px width as well. Font size will be constant, as will coloring.

Possibly more than one reading to make sure the material in question is not too hard. Everyone will read and answer two quizzes of the same font size in alternating order?

Reading Task Info

Test Situation
80%, 100%, 150%, 200%, 300%

User Selection:

Execution:

¬ *Do an initial analysis of your data by providing descriptive statistics (i.e., mean and standard deviation) for each subject and also each spacing;*

¬ *Using all your data, produce graphs for your performance metrics;*

¬ *Calculate the Correlation between each metric and your variable, and determine whether it is "significant" at the .05 level (alpha=.05). (i.e., the odds that the correlation is a chance occurrence is no more than 5 out of 100);*

¬ *Do a Simple Linear Regression Analysis for your data and plot the results (also see this description); {Note 6}*

- *Evaluate all the results, producing a conclusion;*

Results

Discussion

References

Figures/Tables

Correlation alpha = 0.5, degrees of freedom = 23, crit = .462

http://www.gifted.uconn.edu/siegle/research/Correlation/corrchrt.htm

There is a correlation, but there is not enough data to determine what it is.