AN ABSTRACT OF THE THESIS OF

Martha J. Chamberlin for the degree of Master of Science in Computer Science
presented on May 15, 1990.
Title: Relationship of Pauses to Problem Solving Events
in Mechanical Design Protocols

Redacted for Privacy

Abstract approved: _____

Thomas G. Dietterich.

This thesis compares two methods for studying the problem-solving processes of
mechanical design engineers. The first method, verbal protocol analysis, was ap-
plied by L. Stauffer to construct a problem-solving model of mechanical design. The
second method, timing analysis, measures the time intervals separating drawing or
speaking actions during the design process. Timing analyisis was applied by the
author to the verbal/video design data collected by Stauffer. This thesis demon-
strates that the two methods are statistically related, and hence, that employing
two different study techniques enhances the reliability of both methods. The two
methods have complementary strengths: protocol analysis reveals the content of the
design process, while timing analysis is much more complete. Hence, a combination
of protocol and timing analysis provides a stronger measure of the design process
than either method alone.

Relationship of Pauses to Problem Solving Events

in Mechanical Design Protocols

By

Martha J. Chamberlin

A THESIS

submitted to

Oregon State University

in partial fulfillment of

the requirements for the degree of

Master of Science

Completed May 15, 1990

Commencement June 1992

APPROVED:

Redacted for Privacy

<

_____

Professor of Computer Science in charge of major

Redacted for Privacy

_____

Head of Department of Computer Science

Redacted for Privacy

_____   _____

Dean of Graduate School

Date thesis presented May 15, 1990

Typed by Martha J. Chamberlin for Martha J. Chamberlin

Table of Contents

List of Figures

List of Tables

List of Appendix Figures

Relationship of Pauses to Problem Solving Events in Mechanical Design Protocols

Chapter 1

Introduction

Protocol Analysis is a technique that is commonly applied to study human problem solving. The measurement of reaction time (or timing analysis) is another technique commonly used to study human cognitive behavior, including human problem solving. This paper examines the process of applying both of these techniques to the same protocol and then comparing the results, looking for similarities between the two techniques.

Protocol analysis involves asking a subject to think out loud while solving a problem, with the experimenter recording the resulting protocol (a verbatim transcript of the session) for later analysis. This technique was first used extensively by Newell and Simon [19] in their study of the ways in which people solve simple puzzles. Since Newell and Simon's pioneering work, protocol analysis has been used to study human problem solving in many different areas [2, 1, 10, 24, 26].

Before the introduction of protocol analysis as a technique for studying human problem solving, common techniques used in this area included introspection (the experimenter analyzing what the experimenter had done), retrospection (the experimenter asking a subject to analyze what the subject had done), and solution analysis (the experimenter analyzing the solution that the subject came up with) [14]. Another commonly used technique for studying human problem solving is observing external behaviors such as eye movement during the problem solving process [14].

A major advantage of protocol analysis over previously used methods is that it is a technique that studies the human problem solving process *as it is occurring*

[10, 24]. Another major advantage of protocol analysis over the techniques listed in the previous paragraph is that a protocol is a record of what the subject believes is occurring *inside the head*, not just what the experimenter observes externally [10, 30].

Despite its strengths, protocol analysis also has significant weaknesses. Although its supporters emphasize how it is a record of what occurs in the head during the problem solving process, there is serious doubt as to whether or not this is really the case [9, 20, 24]. First of all, there is no proof that what the subject says during a protocol is an accurate inference of the subject's actual cognitive processes. Secondly, it is argued that speaking while solving a problem is a very unnatural process and therefore the vocalization itself could be a confounding variable.

Even if protocol obtained from recording verbalizations is an accurate inference of the subject's cognitive processes, there is a problem with the analysis of this protocol. There is no objective method for analyzing verbal protocols. It is also very difficult to reproduce the results of any protocol analysis [10, 12, 30]. This makes it difficult to have confidence in protocol analysis as a sound scientific method.

Besides the scientific validity problems, protocol analysis is a very long, tedious technique to complete. Just transcribing a protocol from an audio or video tape into a written form can take 10 hours of experimenter time for every 1 hour of subject's protocol time [25]. This restricts the size and complexity of the problems it is practical to analyze using this method. Even given restricted problems, the complexity of the technique leads to consistency problems within a given analysis [12, 30].

Timing analysis — the technique of measuring the time required to accomplish a task — has long been considered a firm, objective measurement of cognitive activity [14, 23, 31]. Timing analyses can include the time to accomplish the actual task, the time to begin the task once a signal has been given (reaction time or RT), or a combination of the two. Timing analysis has been called "chonometry" by some psychologists ([14]).

Different techniques for timing analysis have been employed in a variety of areas to study cognitive complexity [18, 21, 31]. In these studies, increased time to complete a cognitive task has been linked with increased cognitive complexity of the task. The length of the pause preceding a cognitive task is also considered a main indicator of the cognitive complexity of the task [4]. Research looking specifically at the pauses preceding a task in relationship to the complexity of the task shows that there is a strong correlation in increased length of the pause and the increased complexity of the task [3, 13, 17]. Timing analysis also has the strong advantage of being a technique that can be easily automated for many behaviors through the use of such devices as voice-activated recorders [16] and electronic sketching pads [23].

A major weakness in this method is the amount and number of confounding variables. While measuring the time to complete a task, there is always the danger of the subject becoming distracted or bored and therefore doing tasks other than the specified task [23]. When measuring the pause before a task, there is no external evidence about what cognitive processes are occurring during the pause [4]. The difference in lengths of pauses has also been associated with such confounding factors as sex [6], individual style [7], and environmental differences [28].

Another major weakness in timing analysis is its limited scope of application. By definition, it is measuring only the time to complete a task ("when"), not what content led to the completion of the task ("how"). The content of a task is vital to understanding the task, so the loss of content is a major drawback of measuring time alone [23, 24].

From this review, complementary strengths and weaknesses between protocol analysis and timing analysis evidence themselves:

- Protocol analysis is a long, complicated technique, while timing analysis can be easily automated.

- Protocol analysis is very subjective, while timing analysis is an objective technique.

- Timing analysis is limited because it deals with no content, while protocol analysis is primarily a technique for observing content.

If timing and protocol analysis are techniques for studying the same phenomena (i.e., human problem solving), then we would expect to see a correlation between the results of the two. If there is a relationship between the two techniques, then they could be combined to give more objective and complete results than are possible with either technique alone. The purpose of this paper is to study the relationship between protocol analysis and timing analysis to see if a melding of the two techniques is a worthwhile venture.

Chapter 2

Method

This paper discusses the process of comparing the results of a protocol analysis to the results of a timing analysis. The protocol analysis used was an analysis of the design process of mechanical engineers completed by L. Stauffer in the fall of 1987. The timing analysis was performed on the same protocol data as the protocol analysis. This work was completed by the author in the spring of 1990. The remainder of this chapter discusses the protocols and their analyses in more detail.

2.1 Protocol Analysis

Of the five subjects used in the protocol analysis, three were professional mechanical engineers working in industry and two were graduate students in ME with experience in industry. All had at least a BS in mechanical engineering. The least experienced subject had two years of industrial experience; the most experienced subject had 14 years of industrial experience. The average amount of experience in industry for the subjects was nine years. Table 2.1 shows the distribution of the subjects' experience and education. The data from Subject 3 was not included in either of the analyses because the recording was not good enough to consistently understand what was said.

The subjects were allowed to design freely. They were provided with incomplete, high-level design specifications for actual industrial designs. Their work was followed until they produced detailed working drawings. Each problem was designed to take about 10 hours to complete, in sessions over 2 to 4 days, depending on the subject's schedule and the actual time taken to complete the design. The sessions

Table 2.1: Description of Subjects

SUBJECT	DESIGN EXPERIENCE	EDUCATION
1	2 years in industry	ME graduate student
2	12 years in industry	BS in ME
4	4 years in industry	ME graduate student
5	14 years in industry	BS in ME
6	9 years in industry	MS in ME

were held on consecutive days so that the design could stay fresh in the subject's mind, and to mimic actual work conditions in industry.

At the beginning of the design session the subjects were asked to think aloud. During the design session they were reminded to keep talking if they paused for more than two minutes. Throughout the design session, the experimenter would answer the subject's questions. Other than these situations, the subjects were left alone as they designed. Before the actual design session began for each subject, they were given a short test problem to design, so that they could get used to the environment and to thinking aloud as they designed.

The designing session took place in a conference room at each subject's place of employment. The conference room was equipped with a video tape recorder and camera, along with a back-up audio tape recorder. The main physical considerations were to get the microphone close to the subject without getting it in the way and to keep the equipment controls turned away from the subject so he or she was not distracted when tapes were changed. Two audio tape recorders were used to prevent delays during tape changes.

After the completion of the subjects' protocols, the protocol data filled 30 video tapes and 50 audio cassette tapes. Working from these, the protocols were transcribed for later analysis. Figure 2.1 is an example of a transcript taken from the original protocol.

Along with the audio and video tapes of the subject's work, the subjects

S: So, as far as location of the contacts they'll have to be located on the plastic envelope. So umm, it'll be connected between 'em. And the bottom half of the envelope is two parts. So, I had one side and the bottom to figure out some way to make contact with the batteries on that bottom. Or, maybe I can contact the batteries on the side. It looks like I'll have to contact them on the bottom, which means I'll have to put something on the bottom.

S: The types of things that could go on the bottom as contacts... I'd have to look in an electrical catalog to look at that.

Figure 2.1: Original Protocol Transcript

were given paper and pens so that they could sketch solutions or draw or write or calculate as they worked. Figure 2.2 is an example of the "marks-on-paper" (or simply "marks") made by a subject during a design session.

Once the protocol data was collected, the first analysis step was to develop a coarse breakdown. This analysis was performed on all of the protocol data from the subjects. The coarse breakdown was done to identify subjects' global design strategies and to provide a table of contents for the protocols so that sections could be located quickly. The coarse breakdown identified 4 design stages that are described later in this section.

Based on the coarse breakdown, the following four design stages were identified.

- Conceptual Design: where the overall concept of the solution is designed.

- Layout Component Design: where each basic component of the overall concept is designed.

- Detail Component Design: where each component is designed in detail.

- Catalog Selection: where the actual components are selected from a catalog.

Twenty design stages were chosen for a more detailed analysis: each of the five subjects had a section chosen that was representative of each of the four design stages.

.258 ± .006   - env. ht.

,att.   5.40 mm → .213 in.

.04#5 - wall thickness ~~+ sep. thick~~
        contact thickness.

bottom half.   Top half.

UP   DOWN   UP (+ side down)

÷ on board.   - on board
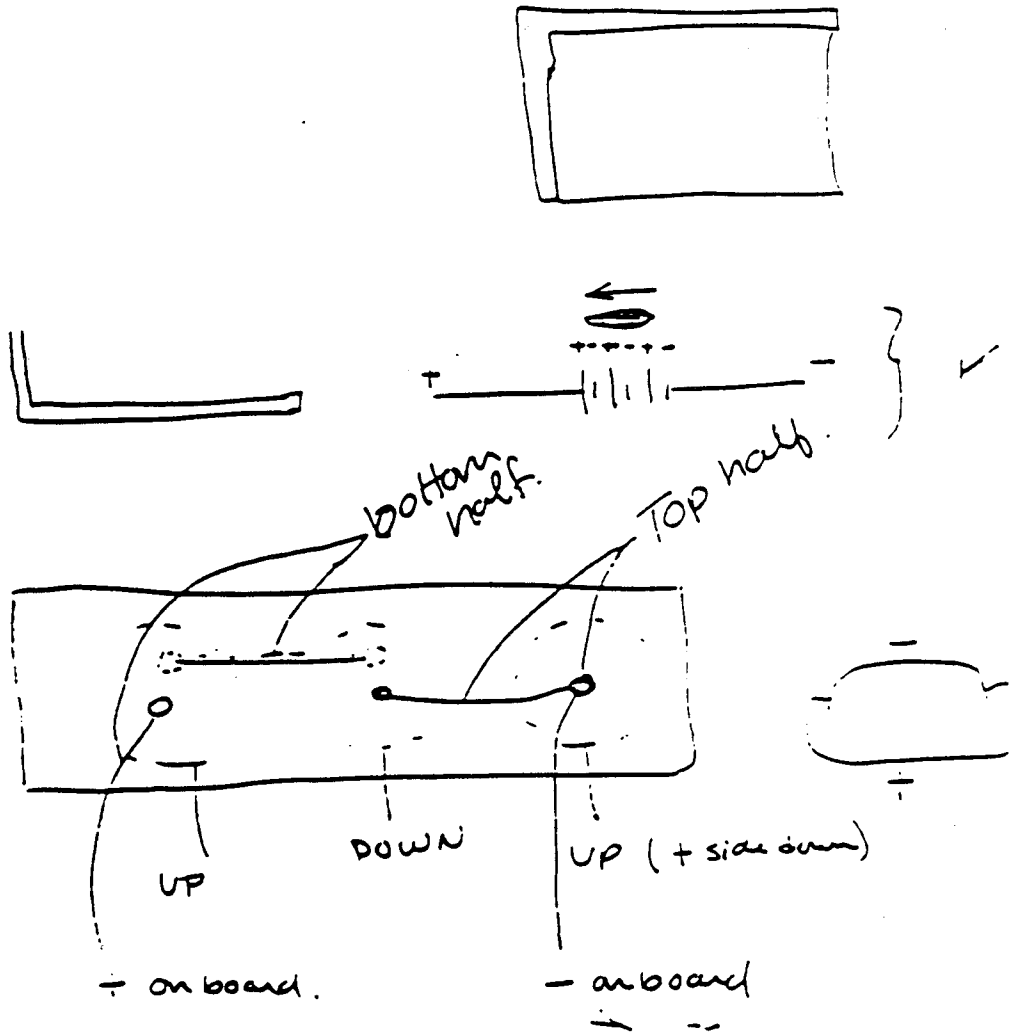
Figure 2.2: Subject's Marks

DESIGN PROCESS – PROTOCOL

TASK 1   TASK 2   ...   TASK m

EPISODE 1   EPISODE 2   ...   EPISODE n

OPERATOR 1   OPERATOR 2   ...   OPERATOR o
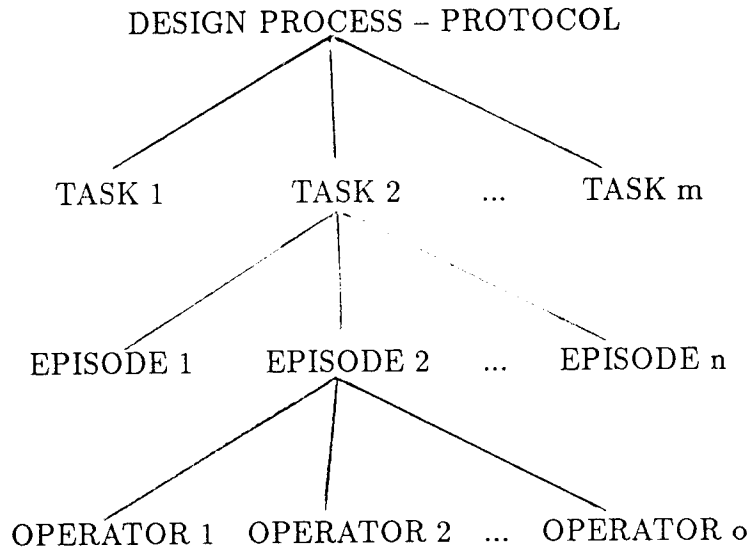
Figure 2.3: The Hierarchy of the TEA Model

These sections, covering 3 1/2 hours, were analyzed to produce a fine breakdown. The purpose of the fine breakdown analysis was to identify the problem-solving processes that a subject performs in order to design. Refer to [25] for more details on how the analysis was performed. The result of the fine breakdown analysis is a hierarchy of performance processes called the TEA Model, as shown in Figure 2.3.

The basic building block of the TEA Model is the design operator, or simply operator. An operator is the basic unit of problem solving as defined by Newell and Simon ([19]): something that can be applied to certain objects to produce different objects. Objects in the TEA Model environment are:

- *properties* of the design such as "location of the contacts" in Figure 2.1,

- design *constraints* from the problem statement such as "I'll have to contact them on the bottom." in Figure 2.1, and

- design *strategies* such as "I'd have to look in an electrical catalog to look at that." in Figure 2.1.

Operators that are applied to objects in the TEA Model fall into the following categories:

- Acquiring needed information (select, calculate).

- Creating proposed designs (create, patch, refine).

- Evaluating proposed designs (simulate, compare).

- Making decisions as to what to do with a proposed design (accept, reject, suspend).

To accomplish a design, the mechanical engineer applies operators in meaningful sequences that form episodes. An episode is a section of protocol that contains a particular focus of attention for the subject. Whenever the subject's attention shifts to a new idea, a new episode begins. Figure 2.1 contains two episodes, one following each "S:". The focus of the first episode is *specifying* the location of the contacts. The focus of the second episode is *planning* how to specify the types of contacts. The TEA Model has six different types of episodes:

- Assimilate (gather information from external sources).

- Plan (articulate next step(s)).

- Specify (develop design proposals).

- Repair (redesign a failed proposal).

- Verify (double check a design proposal).

- Document (make an external communication).

When an episode is completed, the engineer usually tackles another, closely related episode. A collection of related episodes in the TEA Model is called a task. Generally, a task can be described as an episode of larger scope: "Layout the left battery contact" or "Dimension the flipping frame", for example. Tasks are the designing stages that were identified during the coarse breakdown of the protocol analysis. The designing tasks identified by the TEA Model are:

- Conceptual Design

- Layout Component Design

- Detail Component Design

- Catalog Selection

As depicted in Figure 2.3, operators, episodes, and tasks build up a hierarchical model of the design process of mechanical engineers. According to this model, each more complex design level is made up of multiple components from the lower, less complex design level.

After the TEA Model was developed, it was applied to the 3 1/2 hours of protocol selected for the fine breakdown analysis. This was done by giving the approximate time of each episode and the sequence in which the operators occurred. Figure 2.4 shows the TEA Model applied to the protocol transcript in Figure 2.1. Figure 2.4 indicates that Episode 8 began about 7 minutes and 26 seconds into the task. Episode 8 is a Specify episode, specifying the location of the contacts. The first operator in Episode 8 is a select operator, selecting property 10 which is the location of the contacts.

The preceding is a brief description of the protocol analysis done by Stauffer and its results. To summarize, the major products of the protocol analysis are:

- Video and audio tapes of the design sessions (the protocols).

- Transcripts of the protocols.

- The TEA Model, based on the analysis of the protocols.

- A rough temporal placement of the TEA Model events (i.e., episodes and operators) on the protocols.

The basic limitations of Stauffer's protocol analysis are the following:

- Only experienced mechanical engineers were used as subjects.

- Protocols from only five subjects were used.

EPS8

(7:26-9:00) specify-location of contacts

select[property10: location of contacts]

select[constraint11: ??]

compare[property10 TO constraint11]

refine[property10 $\Rightarrow$ property11: contacts are located on bottom
envelope]

pattern[property11 $\Rightarrow$ property12: contacts are located on bottom
envelope and to the bottom of batteries]

pattern[property11 $\Rightarrow$ proper13: contacts are located on side of
envelope]

select[constraint12: info in problem statement]

accept[property12]

EPS9

(9:00-9:20) plan-how to specify types of contacts

select[property3: four contacts from board to configure in property2]

select[constraint13: use off-the-shelf items if possible]

calculate[strategy4: look in electrical catalog for contacts]

accept[strategy4: look in electrical catalog for contacts]

Figure 2.4: Approximate Placement of TEA Model Events

- Only limited parts of each protocol were used from each subject for the final, detailed analysis. Although each subject produced about a 10-hour protocol, the final analysis was performed on only about 30 to 60 minutes from each subject.

More complete discussions of Stauffer's protocol analysis and its results are contained in [25, 26, 27, 29].

2.2 Timing Analysis

The timing analysis was undertaken after the completion of Stauffer's protocol analysis. The data used in the timing analysis was the conceptual, layout component, and detail component design sections used in the protocol analysis for the fine breakdown analysis. The catalogue selection tasks were not used because not all subjects completed this task and, even when the subjects did complete the catalogue selection, marks-on-paper were not consistently performed. Excluding these tasks allowed for more consistency in the analysis of the timing data. The timing analysis also relied heavily on the TEA Model from the protocol analysis.

The first step of the timing analysis was to manually convert the video and audio tapes to chronometric representation (or timings) of the subject's observable actions. "Observable actions" in this analysis are the subject's speech (when they talk and when they are silent) and the subject's marks (when they make marks on the sketch pads and when they make no marks). So, for example, Figure 2.5 shows the speech timing data for 19.56 seconds of Subject 1's protocol. As the section in Figure 2.5 begins, the subject is not speaking. After 6.26 seconds of silence, the subject begins to speak and continues speaking for the next 0.66 seconds. The data continues in this manner until the end of the section, where Figure 2.5 leaves the subject in mid-speech. Although this example is of the speech data, the marks data was converted in the same manner.

All of the minute and second indications in the timing data are relative to the beginning of the video tapes. For example, the timing shown in Figure 2.5 starts

HOURS	MINUTES	SECONDS	STATE
00	35	12.70	STOP
00	35	18.96	START
00	35	19.62	STOP
00	35	21.93	START
00	35	22.70	STOP
00	35	23.50	START
00	35	23.85	STOP
00	35	28.80	START
00	35	30.07	STOP
00	35	32.26	START

Figure 2.5: Sample Segment of Speech Timing Data

0 hours, 35 minutes, and 12.70 seconds into the subject's video tape. This sample timing concludes at 0 hours, 35 minutes, and 32.26 seconds into the subject's video tape.

Although this timing data was collected manually, it is possible to collect both speech and marks automatically [16]. It was not collected automatically in this study because the data had already been collected without consideration of timing before the decision was made to do a timing analysis.

After the speech timing data and marks timing data were collected, they were rounded to one second units and combined into one data set. The purpose of rounding to the nearest second was to simplify the data while still capturing pauses between non-trivial cognitive transitions [5, 7, 28]. Figure 2.6 shows an example of this rounding and combining as performed on the speech timing data beginning in Figure 2.5 and continuing for the next 27 seconds until 0:35:58. The marks timing data for the same period of time is shown as it would be combined with the speech timing data.

A pause unit, or p unit, is a series of contiguous seconds with neither speech

nor marks in them. In Figure 2.6, a p unit begins at 0:35:20 and continues for two seconds. A non-pause unit, or np unit, is a series of contiguous seconds with either speech, or marks, or both speech and marks in them. In Figure 2.6, an np unit begins after the completion of the p unit that begins at 0:35:20. The np unit also continues for two seconds. A pause/non-pause unit, or p/np unit, is any p unit and its immediately following np unit. In Figure 2.6, a p/np unit begins at 0:35:20 and continues for four seconds, including both the p and the np units.

Each line of data in Figure 2.6 contains one p/np unit. The figure shows that there a p unit at second 0:35:13. This p unit continues for six seconds until the end of the p/np unit (0:35:19) when there is a one second p unit of the subject speaking. Similarly, a p/np unit (and therefore a p unit) begins at 0:35:50. The p unit last for only one second of the p/np unit. The last five seconds of the p/np unit are the np unit. During the first two seconds of the np unit, the subject was making marks-on-paper but not saying anything. The next two seconds of the np unit indicates that the subject was speaking in addition to making marks-on-paper. During the last second of the np unit (and therefore the p/np unit), the subject was speaking but not making any marks-on-paper.

The conceptual and layout component design from Subject 2's protocol were not used because of problems with the accuracy of the data: the video recordings were not clear enough to allow for an accurate timing analysis to be completed on the marks-on-paper. Also, any interactions that the subject had with the experimenter or other distractions (such as discussions off of the design subject and equipment difficulties) were removed to reduce possible confounding effects. After these parts had been removed, the remaining data covered just under two hours.

At this point, the timing analysis was complete and all that remained was to compare the results of the two techniques.

2.3 Comparing Protocol Analysis to Timing Analysis

In order to compare the protocol analysis to the timing analysis, the results of the two methods were combined into one data set. The results used from the

0:35:13 ␣␣␣␣␣␣|

0:35:20 ␣␣||

0:35:24 ␣␣␣␣␣|

0:35:30 ␣␣||

0:35:34 ␣␣␣␣|

0:35:39 ␣␣|

0:35:42 ␣␣␣␣␣--+

0:35:50 ␣--++|

0:35:56 ␣␣|

"␣" – indicates one second of pause:

      when there was no speech and no marks

"-" – indicates one second of marks with no speech

"|" – indicates one second of speech with no marks

"+" – indicates one second of both speech and marks

Figure 2.6: Speech/Marks Timing Data Combined and Rounded to Seconds

protocol analysis were the transcripts from the protocols as shown in Figure 2.4 and the rough placement of the TEA Model events (episodes and operators) on the transcripts as shown in Figure 2.4. The results used from the timing analysis were the timing data as shown in Figure 2.6. The two results were combined by annotating the timing data with the speech contents, marks contents, and TEA Model events.

Before marks-on-paper could be added to the timing data, each mark needed to be uniquely identified. Figure 2.7 is an example of a page of a subject's marks with the marks identified. Each group of marks with at least a one second pause preceding and following its production was given its own identification. For example, although there are many different marks involved in writing ".258", this was all grouped together as mark group 1, since these marks were separated by pauses of less than a second. There was, however, more than a second pause between the marks of ".258" and the marks of "±.006", so "±.006" is identified as mark group 2.

After marks-on-paper were identified, they were added to the timing data along with the transcript data (i.e., the speech content). Figure 2.8 is an example of the timing data with the speech and marks content added. This figure shows the timing data from Figure 2.6, the marks from Figure 2.7, and the applicable part of the speech protocol from Figure 2.1 combined into one data set. In Figure 2.8, the "So umm," indicates that this is what the subject said during the np unit of the p/np unit beginning at 0:35:13. Similarly, the "21a" indicates that mark group 21a happened during the first four seconds of the np unit that is part of the p/np unit beginning at 0:35:50.

The last piece of information that needed to be added was the events of the TEA Model. This was done by working with the combined timing, speech, and marks data as shown in Figure 2.8 and the approximate placements of the events as shown in Figure 2.4. Events were physically placed preceding the p/np unit they were detected in. An example of the combined timing analysis results and protocol analysis results (or simply combined results) is shown in Figure 2.9. In this
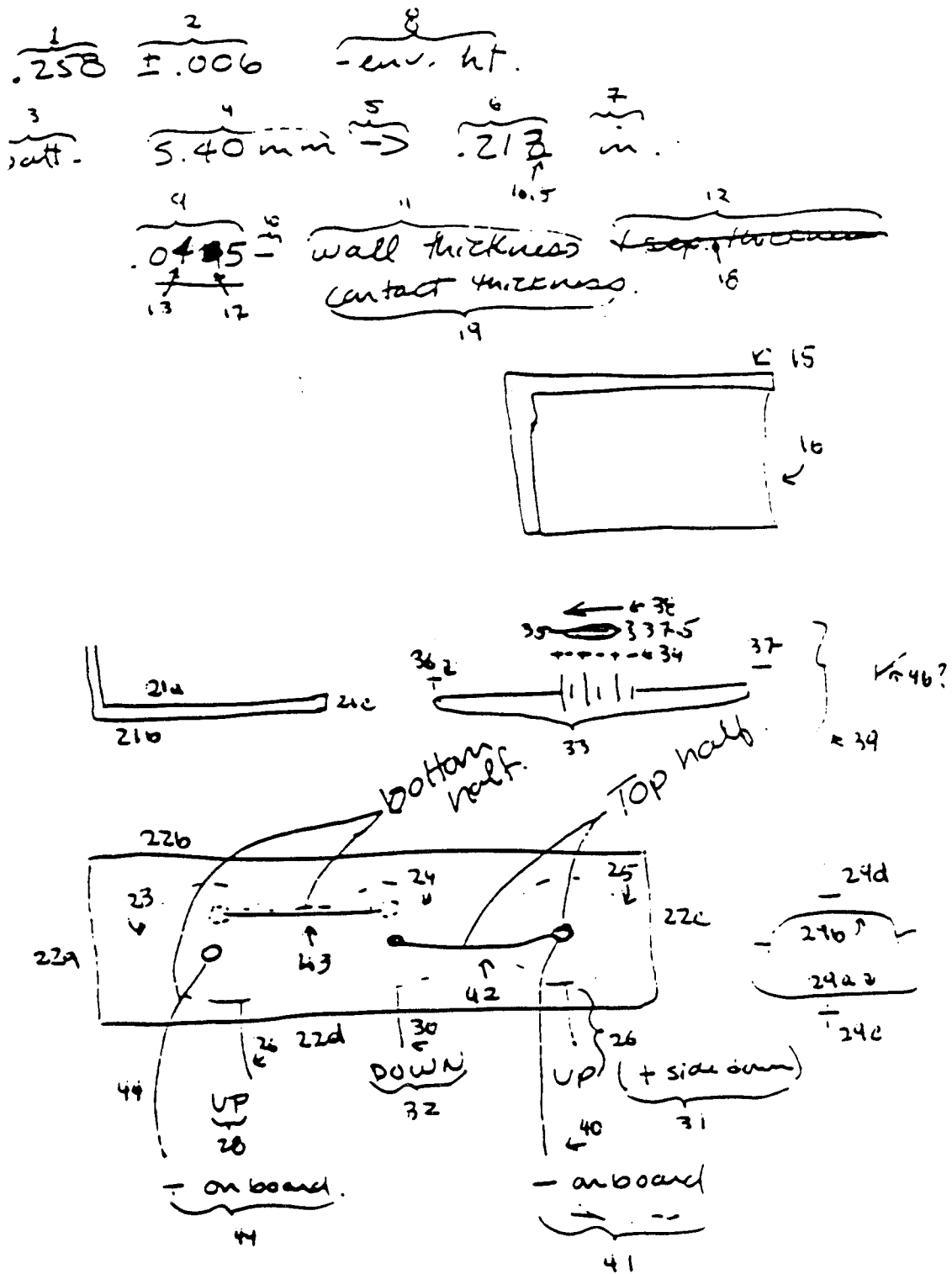
Figure 2.7: Subject's Marks Contents with Labels Added

0:35:13 ⊔⊔⊔⊔⊔⊔| So umm,

0:35:20 ⊔⊔| | it'll be connected between 'em.

0:35:24 ⊔⊔⊔⊔⊔| And the bottom half of the

0:35:30 ⊔⊔| | Envelope is two parts.

0:35:34 ⊔⊔⊔⊔| So,

0:35:39 ⊔⊔| I had

0:35:42 ⊔⊔⊔⊔⊔--+ one side and the bottom

$$21a$$

0:35:50 ⊔--++| to figure out some way to make contact

with the batteries

$$21b, 21c$$

0:35:56 ⊔⊔| on that bottom.

"⊔" – indicates one second of pause:

when there was no speech and no marks

"-" – indicates one second of marks with no speech

"|" – indicates one second of speech with no marks

"+" – indicates one second of both speech and marks

Figure 2.8: Example of Combined Timing, Speech, and Marks Data

example, the refine operator was detected in the p/np unit that begins at 0:35:24 and the pattern operator was detected in the p/np unit that begins at 0:35:50. The combined results for all subjects and all tasks is in Appendix A.

Two experimenters independently placed events in order to test for reproducibility. The first experimenter was the author (a graduate student in Computer Science). The second experimenter was an undergraduate student in Mechanical Engineering. Both experimenters were intimately familiar with the protocol analysis data, having spent long hours working with the protocol tapes.

Each experimenter did her or his best to place each event in the p/np unit where there was evidence that the event had occurred. For example, Figure 2.4 indicates that there is a refine operator that defines the location of the contacts (property10) to be on the bottom of the envelope (property11). The first p/np unit in Figure 2.8 that mentions the location of the contacts is at 0:35:24 when the subject says "And the *bottom half* of the..." (emphasis added). Therefore, the refine operator is placed at the beginning of the p/np unit that starts at 0:35:24.

Although in the preceding example it is fairly clear where the event belonged, this was not always the case. For example, consider Figure 2.10. Although the subject may have decided at 0:39:22 (or earlier) that she was going to look up information later (strategy5), there is no spoken evidence of exactly what the strategy is until 0:39:34. Similarly, it might be argued that she accepts strategy5 at 0:39:28 (the first time she speaks with surety about her goals), or perhaps at 0:39:34 when the strategy is vocalized, or at 0:39:44 when she decides that the strategy is a good plan, or even later, when she goes on to the next topic. As shown in Figure 2.11 Experimenter 2 placed the create event at 0:39:22 while Experimenter 1 waited until 0:39:34 to place the event. Both experimenters decided to place the accept event at 0:39:44, when the subject decides that the strategy is a good plan.

There were also events that had little or no evidence of having occurred at all. For example, referring back to Figure 2.4, one of the events is a select operator selecting constraint11 that is described as being "??". Since it isn't clear what constraint is being selected, it is difficult to decide exactly when vocalization of the

0:35:13 ␣␣␣␣␣␣| So umm,

0:35:20 ␣␣| | it'll be connected between 'em.

       refine[property10 $\Rightarrow$ property11: contacts are located on

       bottom envelope]

0:35:24 ␣␣␣␣␣| And the bottom half of the

0:35:30 ␣␣| | Envelope is two parts.

0:35:34 ␣␣␣␣| So,

0:35:39 ␣␣| I had

0:35:42 ␣␣␣␣␣--+ one side and the bottom

            21a

       pattern[property11 $\Rightarrow$ property12: contacts are located on

       bottom envelope and to the bottom of batteries]

0:35:50 ␣--++| to figure out some way to make contact

       with the batteries

           21b, 21c

0:35:56 ␣␣| on that bottom.

"␣" – indicates one second of pause:

       when there was no speech and no marks

"–" – indicates one second of marks with no speech

"|" – indicates one second of speech with no marks

"+" – indicates one second of both speech and marks

Figure 2.9: Combined Timing Analysis Results and Protocol Analysis Results

0:39:22 ⊔⊔⊔⊔⊔| So what I would probably do

0:39:28 ⊔⊔⊔| What I'm going to do

0:39:32 ⊔| I'm going to do,

0:39:34 ⊔⊔⊔| I'm gonna hold on to that and look at that.

0:39:38 ⊔| | | But for now I'm gonna keep on working on what

I have, and when I have more electrical questions, I'll

0:39:42 ⊔| look them up.

0:39:44 ⊔⊔⊔| Good plan.

Events to place:

> create[strategy5: look up information later and continue for now]
>
> accept[strategy5: look up information later and continue for now]

Figure 2.10: Confusing Event Placement

constraint occurred.

To help work around this ambiguity, each experimenter rated the level of certainty with which each event was placed as either SURE (S) or UNSURE (U). Figure 2.11 shows an example of how events were placed using the S/U level of certainty. Following each event are the experimenters' certainty levels, separated by a colon if both experimenters placed the event at the same p/np unit. For example, the create operator was placed on the p/np unit starting at 0:39:22 by Experimenter 2 and on the p/np unit starting at 0:39:34*** Experimenter 1. Neither experimenter was sure of these placements. Both experimenters placed the accept event on 0:39:44. This time experimenter 2 was sure of his placement while experimenter 1 was not. The complete data of this form is found in Appendix A.

Once the timing data was annotated with the speech contents, the marks contents, and the TEA Model events, the protocol analysis and the timing analysis were ready to be compared. The results of this comparison are found in Chapters 3 and 4.

    create[strategy5: look up information later and continue for now]{2U}

0:39:22 ⊔⊔⊔⊔⊔| So what I would probably do

0:39:28 ⊔⊔⊔| What I'm going to do

0:39:32 ⊔| I'm going to do,

    create[strategy5: look up information later and continue for now]{1U}

0:39:34 ⊔⊔⊔| I'm gonna hold on to that and look at that.

0:39:38 ⊔| | | But for now I'm gonna keep on working on what

    I have, and when I have more electrical questions, I'll

0:39:42 ⊔| look them up.

    accept[strategy5: look up information later and continue for now]{1U:2S}

0:39:44 ⊔⊔⊔| Good plan.

Figure 2.11: Certainty Levels in Event Placement

Limitations on the timing analysis beyond those inherited from the protocol data include:

- Possible inaccuracies in the timing since they were done by hand, working from the video tapes of the protocols.

- The results of Stauffer's protocol analysis were used to guide the timing analysis. Therefore, this does not represent an independent experiment, and the results of the timing analysis might be influenced by the results of the protocol analysis.

- Only four of the original six subjects were used due to clarity problems with the video tapes.

Chapter 3

Reliability of the Protocol Analysis

3.1 Issues and Implications

The first issue of comparing the protocol analysis and the timing analysis concerns the reliability of the protocol analysis. A protocol analysis is reliable if the results of the analysis can be reproduced. As discussed in Chapter 1, protocol analyses are often considered unreliable because they cannot be consistently reproduced.

It is important to consider the reliability of the results of Stauffer's protocol (i.e., the TEA Model), because it bears a direct connection with the reliability of the comparison done in this paper. If there is no reliability in the TEA Model, then there is no point in comparing the results of the protocol and timing analyses described in Chapter 2.

3.2 Method

The consideration of the reliability of Stauffer's protocol was done by working with the combined timing, speech, and marks data as discussed in Section 2.3 (example shown in Figure 2.8) and the approximate event placement as discussed in Section 2.1 (example shown in Figure 2.4). Each experimenter worked alone and results were not compared until all of the events had been placed by both experimenters.

Once the placements had been completed, they were compared, and each event was noted as begin either "agreed" (both experimenters placed it on the same

p/np unit) or "non-agreed" (each experimenter placed it on a different p/np unit).

3.3 Results

Table 3.1 shows the allocation of event placements by the two experimenters. 549 of the events (66.63% of the total events) were agreed upon by both experimenters. This gives strong evidence to the reliability of these events. Of the 275 events that were not agreed upon, 108 events (39.27%) were not placed by Experimenter 2 because of how uncertain he was about them. Still, most of the events can be placed by two independent experimenters and are therefore reliable.

The experimenters were sure of the placement of over 90% of the events that they both agreed upon (93.62% for Experimenter 1 and 96.17% for Experimenter 2). This is more evidence of the reliability of most of the events. The experimenters were sure of the placement of only about half of the events they did not agree upon. A lower percentage of sure events would be expected in the non-agreed category since the non-agreed events' reliability is already in question.

Another observation about the data in Table 3.1 is that Experimenter 2 was less unsure of his event placement than Experimenter 1 (3.83% to 6.38% of agreed events and 3.64% to 57.09% of non-agreed events). This is consistent with what would be expected, since Experimenter 2 was working within his field of expertise (Mechanical Engineering) while Experimenter 1 was outside of her area of expertise (Computer Science).

Experimenter 1 was also willing to place all of the events while Experimenter 2 would not willing to place 108 of the events. This might also be explained by the differences in their background: Experimenter 1, having faith that Stauffer had evidence for all of his events, was more willing to see evidence for events that may or may not have been there.

Despite the strong evidence that Stauffer's events exist, it is important to remember that this was not a "blind" placement. The experimenters were working from clearly defined events and approximate placements of them, rather than from the protocol transcripts. So, although it is reasonable to assume that Stauffer had

Table 3.1: Event Placement by the Experimenters

AGREED EVENTS:

	EXPERIMENTER 1	EXPERIMENTER 2
SURE	514	528
	93.62%	96.17%
UNSURE	35	21
	6.38%	3.83%
TOTAL	549	549
	100.00%	100.00%

NON-AGREED EVENTS:

	EXPERIMENTER 1	EXPERIMENTER 2
SURE	118	157
	42.91%	57.09%
UNSURE	157	10
	57.09%	3.64%
NOT	0	108
PLACED	0%	39.27%
TOTAL	275	275
	100.00%	100.00%

TOTAL EVENTS:

AGREED EVENTS	NON-AGREED EVENTS	TOTAL EVENTS
549	275	824
66.63%	33.37%	100.00%

reliable evidence for about two-thirds of his events, it isn't clear that an independent experimenter working from the protocol transcripts would isolate the same events as being significant to the design process.

Still, the evidence in favor of the reliability of Stauffer's protocol is strong enough that it is worthwhile to continue in the comparison between the protocol and timing analyses. Chapter 4 gives the details of this comparison and its results.

Chapter 4

Event Placement and Timing

4.1 Issue and Implications

The main issue considered in this chapter is whether or not there is a pattern to the placement of TEA Model events with respect to the lengths of the pauses and non-pauses in the timing data. To study this issue, the combined timing and protocol results were statistically analyzed. An example of these results are shown in Figure 2.9 of Chapter 2, the complete results are in Appendix A.

If the events in Stauffer's protocol analysis are *not* placed randomly, then this is positive evidence that there is a relationship between protocol analysis and timing analysis and that a melding of the two techniques is a worthwhile venture. Non-randomly placed events would also demonstrate that both techniques are measuring at least some of the same problem solving processes.

If the events in Stauffer's protocol analysis *are* placed randomly with respect to the pause lengths, then there is no evidence from this study to indicate that there is any relationship between protocol analysis and timing analysis. This could be an indication that the two methods are measuring different processes, that Stauffer performed an inaccurate protocol analysis, or that the assumptions underlying timing analysis as specified in Chapter 1 are incorrect.

4.2 Method

There are several limitations to the combined timing and protocol data that restrict how it can be analyzed. Many statistical tests are not applicable to this

data:

- Tests for normally distributed data (i.e., t-tests or simple ANOVA test [23]) cannot be used because the TEA Model events were not normally distributed on the timing data.

- Intersubject tests (i.e., median-test or Krusdall-Wallis ANOVA [23]) cannot be used, because data was analyzed for only 4 subjects.

- Intrasubject tests (i.e., sign test or Friedman ANOVA [23]) might produce skewed results because the design sections were not chosen randomly.

The strongest statistical test that is appropriate for this study is a goodness-of-fit test as defined in [15, 22]. A goodness-of-fit test compares an observed frequency distribution to an expected frequency distribution and indicates whether or not the two distributions are the same. In this study, the observed distribution is the distribution of events across pause/non-pause units (p/np units) of various lengths. The number of events assigned to p/np units of length 1 second, 2 seconds, 3 seconds, and so on, are measured. These are the observed number of events, or observed frequencies. The number of events that would have been expected if the events had been assigned randomly to the p/np units are also computed. These are the expected number of events, or expected frequencies.

Let o_i be the observed number of events assigned to p/np units of length i. Let e_i be the expected number of events assigned to p/np units of length i. Let n be the number of p/np lengths being considered. Then the statistic

$$\chi^2 = \sum_{i=1}^{n} \frac{(o_i - e_i)^2}{e_i} \tag{4.1}$$

is distributed approximately χ^2 with $n - 1$ degrees of freedom. By computing this quantity and inspecting a χ^2 table, a value for alpha can be obtained. Alpha is the probability that there is an error if it is concluded that the observed distribution is different from the expected distribution. In this case, alpha will be the probability that there is an error if it is concluded that events are *not* randomly placed with

Table 4.1: Distribution of P/NP Units

CASE 1: AGREED EVENTS	CASE 2: EITHER EVENTS	CASE 3: NON-AGREED EVENTS	CASE 4: NO EVENTS	TOTAL EVENTS
180	133	177	747	1237
13.41%	10.75%	14.31%	60.39%	100.00%

respect to the pause lengths. Following standard statistical terminology, each value of i is called a "trial".

4.2.1 Observed Frequencies

There were two refinements made to the data in order to improve the accuracy of the values for o_i. The first refinement works to eliminate unreliable p/np units while the second refinement works to eliminate unreliable events.

In the timing data, any particular p/np unit falls into one of four cases:

- CASE 1: The p/np unit had only agreed events assigned to it.

- CASE 2: The p/np unit had one or more agreed events along with one or more non-agreed events assigned to it.

- CASE 3: The p/np unit had only non-agreed events assigned to it.

- CASE 4: The p/np unit had no events assigned to it, agreed or non-agreed.

Table 4.1 shows the distribution of these forms.

Non-agreed events are suspect in their placement in the timing data so it is not clear if p/np units with only non-agreed events on them should be considered as having events on them or not. Since the main concern in this section is to examine the relationship between events and the size of the p/np units that they are placed on, the number and lengths of p/np units that are considered as having events on

them are key to the results of this analysis. To avoid the confounding effects of this, all Case 3 p/np units (i.e., p/np units with non-agreed events on them) were not considered further in this analysis. This means that a small portion of p/np units (14.31% of the total p/np units) were removed from the pool, leaving 85.69% of the total p/np units for the goodness-of-fit test.

The second refinement that was done to the data was eliminating non-agreed events from the event pool. Since the relationship between the TEA Model events and the timing data relies heavily on the accuracy of the annotation described in Section 2.3, only those events whose placement was agreed upon by both experimenters were used for the goodness-of-fit test. All agreed events were used, whether SURE or UNSURE.

4.2.2 Expected Frequencies

The expected frequency e_i for the ith trial was calculated using the following equation:

$$e_i = ut_i * er \tag{4.2}$$

where:

- ut_i is the unit time: the total number of seconds in the timing data for p/np units of length i.

- er is the event rate: the number of TEA Model events that are expected to occur in each second.

The event rate er is calculated as follows:

$$er = \frac{te}{tt} \tag{4.3}$$

where:

- te is the total events: the total number of agreed events observed.

- tt is the total time: the total number of seconds in the timing data for p/np units of all lengths (Cases 1, 2, and 4 only).

Table 4.2: Event Rates for Each Subject and Area

	Conceptual Design	Layout Design	Detail Design
Subject 1	0.0838	0.1669	0.0765
Subject 2	N/A	N/A	0.1084
Subject 5	0.1327	0.1324	0.0557
Subject 6	0.1247	0.1000	0.0633

Ave: 0.1047

Std: 0.0353

Equation 4.2 reflects the assumption that all seconds in the timing data (p/np units), are equally likely to have an event placed on them. In the observed data, all events were placed on whole p/np units rather than individual seconds. In order to be consistent with the observed data, an expected event that is randomly placed on any second within a p/np unit is considered to be placed on the whole p/np unit. This means that a 10-second p/np unit has twice the probability of an event being placed on it as a 5-second p/np unit.

Because of individual differences between human beings, we would expect that individual mechanical engineers might have different event rates. It is also possible that individual mechanical engineers have different rates during different phases of their design activity.

Table 4.2 shows the event rates for each subject and design area. The average event rate and standard deviation of the event rates from Table 4.2 are calculated as follows [23]:

$$Ave = \frac{\sum_{j=1}^{n} er_j}{n} \tag{4.4}$$

$$Std = \frac{\sum_{j=1}^{n} er_j - Ave}{n-1} \tag{4.5}$$

where n is 10 (i.e., the total number of subjects and design areas).

Since the standard deviation from Table 4.2 is 0.035 (about one-third of the average), there is strong evidence that the event rates are significantly different for each subject and each design area, or perhaps some other uncontrolled variable. To protect against confounding factors that result from these variations in event rates, Equation 4.2 was calculated separately for each subject, and each phase of the design process for each subject.

Table 4.3 shows an example of how the expected number of TEA Model events are calculated for each trial. For consistency with future tables, only episode events are considered in this example: operator events were omitted. Expected frequencies are calculated similarly for both events. The trial chosen in this table is p/np units of length 2 seconds. The columns of Table 4.3 are as follows:

- The first column indicates the subject and design area under consideration. For example, the first subject is Subject 1 (S1) and the first design area is Subject 1's Conceptual Design as described in Chapter 2.

- The second column shows the number of seconds of length 2 p/np units that were observed in Cases 1, 2, and 4 of the protocol (ut_2). For example, Subject 1's Conceptual Design had a total of 34 seconds devoted to p/np units of length 2. This means that there were a total of 17 p/np units of length 2 seconds in Subject 1's Conceptual Design.

- The third column gives the total number of agreed TEA Model events observed in each section (te_2). For example, 12 agreed events were observed in Subject 1's Conceptual Design.

- The fourth column gives the total time of Case 1, 2, and 4 p/np units observed in each section (tt). For example, Subject 1's Conceptual Design lasted for a total of 551 seconds, of which 34 seconds were 2 second long p/np units.

- The fifth column gives the events rate for this subject and design area as calculated by Equation 4.3. For example, during Subject 1's Conceptual Design, she had the following event rate for episodes:

Table 4.3: Calculation of Expected Episodes

Subject	Section	ut_2	te_2	tt	er	e_{i2}
S1	Concept	34	12	551	0.02	0.74
	Layout	38	9	432	0.02	0.79
	Detail	24	15	317	0.05	1.14
S2	Detail1	34	12	541	0.02	0.75
	Detail2	18	12	262	0.05	0.82
S5	Concept	32	13	476	0.05	0.87
	Layout	46	9	682	0.01	0.61
	Detail	14	6	245	0.02	0.34
S6	Concept	36	5	270	0.02	0.67
	Layout	90	12	731	0.02	1.48
	Detail	26	4	216	0.02	0.48
	Total:	392	109	4706	n/a	8.70

Subject   - Subject number

Section   - Design section

ut_2 - Total observed time of 2-second long p/np units
for this subject and design section

te_2 - Total observed episodes for this subject and design section

tt - Total observed time for this subject and design section

er - Observed episode event rate for this subject and design section

e_{i2} - Total expected episodes in 2-second long p/np units
for this subject and design section

$$er = \frac{te_2}{tt} = \frac{12}{551} = 0.02$$

This means that she designed at a rate of 0.02 episodes per second. Note that, since each event rate was kept separate during the during the analysis, the total event rate is not applicable.

- The final column gives the expected number of randomly placed TEA Model events in eac'h section (e_{i2}) as calculated by Equation 4.2. For example: each of the 17 p/np units of length 2 seconds in Subject 1's Conceptual Design would expect to have a total of 0.74 episodes randomly placed on them:

$$e_{i2} = ut_{i2} * er = 34 * 0.02 = 0.74$$

Based on Figure 4.3, the total number of episodes expected to be placed randomly on p/np units of length 2 is 8.70 events. The other $109 - 8.70 = 100.30$ episodes would be randomly placed on p/np units of length other than 2.

4.2.3 Goodness-of-Fit Test: chi-square

Once the observed and expected events are calculated for each p/np unit, Equation 4.1 of the goodness-of-fit test can be applied. Table 4.4 shows an example of how this χ^2 is calculated. Only episode events are considered in this table so that it will fit on one page. The complete tables used for calculating χ^2 in this study are contained in Appendix B. The columns in Table 4.4 contain the information described below.

- The first column gives the length of the p/np units being examined, or trial length. For example, the first row is for p/np units of length 2 seconds. Note that several trials (e.g., 1, 20, 21, etc.) are missing. This is because there were no p/np units of those lengths in the observed data.

- The second column is the trial number, or i. For example, p/np units of length 2 seconds are trial number 1. Trials with $e_i = 0$ were omitted from the

χ^2 equation, therefore they were not numbered in this column. See below for more explanation of this.

- The third column shows how many TEA Model episodes were observed in this trial. For example, p/np units of length 2 seconds had a total of 10 episodes observed on them.

- The last two columns show how many TEA Model episodes would be expected if they were placed at random. For example, p/np units of length 2 seconds would expect to have 8.7 real or 9 integer episodes placed on them randomly.

Since χ^2 in Equation 4.1 approaches infinity as e_i approaches zero, it is necessary to avoid near-zero values of e_i. This was accomplished by rounding the expected number of events to the nearest integer, and then using this approximation as the value of e_i. All unit lengths with integer $e_i = 0$ were then dropped from the equation. This leaves 15 trials to be encorporated into Equation 4.1 as follows:

$$\chi^2 = \sum_{i=1}^{n} \frac{(o_i - e_i)^2}{e_i}$$
$$= \frac{(10-9)^2}{9} + \frac{(10-15)^2}{15} + \frac{(18-16)^2}{16} \cdots \frac{(3-3)^2}{3} + \frac{(3-4)^2}{4} + \frac{(1-1)^2}{1} + \frac{(1-1)^2}{1}$$
$$= \frac{1}{9} + \frac{25}{15} + \frac{4}{16} + \frac{1}{17} + \frac{121}{10} + \frac{9}{9} + \frac{0}{7} + \frac{1}{7} + \frac{4}{4} + \frac{4}{5} + \frac{1}{1} + \frac{0}{3} + \frac{1}{4} + \frac{0}{1} + \frac{0}{1}$$
$$= 32.79$$

Since 15 of the rows in this table were used, there are $15 - 1 = 14$ degrees of freedom (dof). From a χ^2 table, alpha = 0.005. In other words, there is a 0.5% chance of begin wrong in concluding that the episodes in Table 4.4 are *not* distributed randomly: there is a non-random pattern to the episodes in Table 4.4.

4.3 General Relationship Results

Table 4.5 shows the alpha values for the distribution of events on the timing data. These results are in three forms as follows:

- The first form is for trials based on the length of the p/np units. For example, consider the p/np unit in Figure 2.9 that begins at 0:35:50. The length of this

Table 4.4: Episodes on P/NP Units

Trial Length	i	o_i	e_i (real)	e_i (int)
2	1	10	8.70	9
3	2	10	14.93	15
4	3	18	16.32	16
5	4	16	17.18	17
6	5	21	9.90	10
7	6	6	8.63	9
8	7	7	7.26	7
9	8	8	6.52	7
10	9	2	4.40	4
11	10	3	5.39	5
12	11	0	1.30	1
13	12	3	2.65	3
14		0	0.49	0
15	13	3	3.87	4
16		0	0.26	0
17	14	1	0.98	1
18		0	0.49	0
19	15	1	0.90	1
24		0	0.44	0
29		0	0.38	0
total:	15	109	109.00	109

chi: 32.79

dof: 14

alpha: 0.005

Table 4.5: Alpha Values for Events on Timing Data

P/NP UNITS	P UNITS	NP UNITS
0.005	0.950	0.050

p/np unit is 6: one second of pause, two seconds of marks alone, two seconds of both speech and marks, and one second of speech alone. This p/np unit would be put into trial length 6, and the pattern operator that is placed on it would be added to the observed events for trial length 6. This is how the example in Table 4.4 was completed.

- The second form is for trials based on the length of the p units. For example, there is a p/np unit in Figure 2.9 that begins at 0:35:50. The length of the p/np unit is six, but the length of the p unit is only 1: the first second of the p/np unit. This p unit would be put into trial length 1, and the pattern operator that is placed on it would be added to the observed events for trial length 1.

- The third form is for trials based on the length of the np units. For example, there is a p/np unit in Figure 2.9 that begins at 0:35:50. The length of the p/np unit is six and the length of the p unit is 5: all seconds except for the first second of the p/np unit. This np unit would be put into trial length 5, and the pattern operator that is placed on it would be added to the observed events for trial length 5.

According to Table 4.5, there is only a 0.5% chance of events being placed randomly with respect to the entire p/np units. This is the strongest pattern evident. Next is the placement of events with respect to np units alone. This shows a 5% chance of being random which, although it is worse than the entire p/np units, is still statistically significant. Finally, there is a 95% chance of the events being placed randomly with respect to the p units alone.

Table 4.6: Distribution of Episodes and Operators

Episodes	Operators	All Events
109	440	549
19.85%	80.15%	100%

The results in Table 4.5 can be seen as support for the theory that preceding pauses are indicators of cognitive activity as discussed in Chapter 1. However, the fact that the entire p/np unit is the best predictor of events suggests that there is also cognitive activity involved in the non-pauses as well.

It is clear from Table 4.5 that there are distinct patterns in the placement of events with respect to the length of the timing data. Now a main issue is: what is the relationship between protocol analysis and timing analysis? Several variations of this analysis were examined to give a better insight into how protocol analysis and timing analysis might be related. The following variations were investigated:

- the placement of episodes versus the placement of operators and

- the placement of events on marks timing data versus speech timing data.

4.3.1 Episodes and Operators

Table 4.6 shows the relative proportions of episodes and operators. Since episodes are comprised of multiple operators, it would be expected that there would be many more operators than episodes. This is, in fact, the case: operators make up over 80% of all events. This means that any pattern evident in the placement of the episodes may be overshadowed by the pattern of the operators. Therefore, separating out the two event types could uncover a prominent pattern in episodes.

Table 4.7 shows the alpha values for the distribution of operators and episodes. The only suggestion of a pattern in these numbers is that the data is getting more and more random. When all events are considered together, both p/np units and np units show non-random event distributions. But operators alone show a

Table 4.7: Alpha Values for Operators and Episodes

	P/NP UNITS	P UNITS	NP UNITS
All Events:	0.005	0.950	0.050
Operators:	0.025	0.950	0.950
Episodes:	0.950	0.950	0.950

non-random pattern only when the entire p/np unit is considered, and episodes alone show no non-random patterns.

If episodes are more complex events than operators and lengths of pauses and/or non-pauses are indicators of event complexity, we would expect to see episodes distributed less randomly with respect to the timing data than operators. According to Table 4.7, however, episodes are more randomly distributed than operators. This might be an indication that operators are actually the more complex events requiring greater "preparation" by the designer. However, this might also be an indication that operators and episodes are fundamentally different events and that operators can be detected by timing analysis whereas episodes cannot. Another possible explanation is that the p/np units used in this experiment were of the wrong granularity to be able to accurately predict events the size of episodes.

There is another possible explanation for the values in Table 4.7. Recall from Table 4.6 that there are 109 episodes, 440 operators, and 549 events. Since episodes are only about 20% of all events, there may not be enough of them in this sample to show a pattern. Statistical analysis of a larger sample might produce a stronger pattern for episodes.

4.3.2 Speech and Marks

The next issue considered is the difference between the placement of events on speech timing data and marks timing data. This was done to try to isolate any confounding effect that "thinking aloud" during a design session might have

on the timing data. If there *is* evidence of confounding effects, then this weakens the results of protocols collected using the technique of "thinking aloud". Another reason to consider speech and marks separately is to look for evidence that the two processes are distinct. If they *are* distinct, then the validity of marks timing data from a protocol is not weakened by the confounding effects of "thinking aloud".

Table 4.8 shows the distribution of event types on p/np units in the four forms discussed earlier in Subsection 4.2.1. The number of p/np units is different for speech alone, marks alone, and the combination of speech and marks since the presence or absence of speech or marks determines the p/np units. For example, there are only 692 p/np units for marks although there are 1340 p/np units for speech. This is what would be expected since speech is much more common in the protocol data than marks.

Table 4.1 first introduced the four cases that describe the p/np units in this paper. Table 4.8 shows the same four cases that were defined in Table 4.1 as they relate to the p/np units for marks alone and the p/np units for speech alone. Since the p/np units defined by speech alone and marks alone are different, the proportion of events in each of the four cases is also different. As before, the p/np units of Case 3 were removed from the data to reduce the risk of questionable p/np units confounding the results. Cases 1, 2, and 4 comprise 1153 of the Speech p/np units (86.04%) and 637 of the Marks p/np units (92.05%). Therefore, as with the combined data, relatively few p/np units are removed from consideration in both Speech and Marks. Again, all non-agreed events were removed from consideration.

Since none of the subjects used in this study were accustomed to "thinking aloud" during a design, we would expect this to affect the distribution of the events with respect to the speech timing data. Furthermore, whenever the subject stopped talking for more than two minutes during the protocol session, the experimenter prompted the subject to "keep talking": another environmental condition that the subjects were not accustomed to. These are two reasons to suspect that events would be more randomly distributed with respect to the length of speech p/np units than marks p/np units. Table 4.9 shows the alpha values for the speech timing data and

Table 4.8: Distribution of Events on Speech and Marks P/NP Units

SPEECH P/NP UNITS:

CASE 1: AGREED EVENTS	CASE 2: EITHER EVENTS	CASE 3: NON-AGREED EVENTS	CASE 4: NO EVENTS	TOTAL EVENTS
186	133	187	834	1340
13.88%	9.93%	13.96%	62.24%	100.00%

MARKS P/NP UNITS:

CASE 1: AGREED EVENTS	CASE 2: EITHER EVENTS	CASE 3: NON-AGREED EVENTS	CASE 4: NO EVENTS	TOTAL EVENTS
60	113	55	464	692
8.67%	16.33%	7.95%	67.05%	100.00%

Table 4.9: Alpha Values for Speech and Marks

	P/NP UNITS	P UNITS	NP UNITS
Combined:	0.005	0.950	0.050
Speech:	0.005	0.005	0.005
Marks:	0.005	0.005	0.005

the marks timing data.

Table 4.9 is a pleasant surprise: when speech and marks are separated, events are non-randomly placed with respect to all forms of p/np units. This evidence suggests that thinking aloud as enforced during Stuaffer's protocols was *not* a confounding variable when considering the p/np units from cases 1, 2, and 4. It is also interesting to note from Table 4.9 that both speech timing data and marks timing data are more accurate alone than they are together. This is evidence that the speech and marks processes are at least partially distinct.

4.3.3   Episodes and Operators, Speech and Marks

The last issue considered is the combined effects of operators and episodes and the marks timing data and the speech timing data. Examining each of these separately might give us more insight into exactly what is happening. Table 4.10 shows the alpha values from these analyses.

Figure 4.10 shows a clearer picture emerging. A significant difference between the speech timing data and the marks timing data is evident: marks p/np units are significantly related to both episodes and operators while speech p/np units are significantly related to operators but not episodes. This is evidence that marks p/np units are more reliable predictors of events, as we would expect if timing data is a reliable predictor of events and thinking aloud is a confounding variable.

When the data is considered as in Table 4.10, there is almost no difference in the events being placed on p/np units versus p units versus np units. This is

Table 4.10: Alpha Values for Episodes and Operators on Speech and Marks

P/NP UNITS:

	Episodes	Operators
Speech:	0.950	0.005
Marks:	0.005	0.005

P UNITS:

	Episodes	Operators
Speech:	0.950	0.050
Marks:	0.005	0.005

NP UNITS:

	Episodes	Operators
Speech:	0.950	0.005
Marks:	0.005	0.005

further evidence that both pauses and non-pauses contribute to the complexity of cognitive processes as was earlier proposed to explain the results in Figure 4.5.

In summary:

- Marks timing data is a stronger predictor of events than speech timing data.

- Speech timing data predicts operators better than it predicts episodes.

- There is no significant difference in the marks timing data's ability to predict episodes versus operators.

There is enough evidence in this section connecting timing analysis with protocol analysis to warrant further study into this area. Most of the questions raised in this section are deferred until the completion of further research. However, there is one more issue that is easy to examine within this study as it stands. This is the issue of the effect of CASE 4 p/np units on the distribution of events. This is the topic of the next section.

4.4   Non-Blank P/NP Unit Relationship Results

A common criticism of protocol analysis is that it is incomplete. If Stauffer's protocol analysis is incomplete, this could create a signal that correlates with the timing data. The purpose of this section is to test this possibility and demonstrate that the relationships revealed in Section 4.3 are not an artifact of incomplete protocol analysis.

If Stauffer's protocol analysis is incomplete, there must be more events than the 824 that are described in Table 3.1 of Chapter 3. If there are more events, then the distribution of events on p/np units as described in Figure 4.1 could be incorrect: missing events would probably fall on some of the Case 3 and 4 p/np units, moving them into Cases 1 or 2. Since Case 3 has already been disqualified as being suspect, the p/np units in Case 4 are the most likely to change with the addition of more events.

Table 4.11: Alpha Values for Events on Timing Data, Cases 1 & 2

P/NP UNITS	P UNITS	NP UNITS
0.005	0.005	0.005

Removing the Case 4 p/np units leaves only those units with one or more agreed events on them. This means that p/np units that should have had events placed on them but did not because of the incompleteness of the protocol analysis are removed. These are the units that are producing the signal observed in the preceding section if this signal is caused by the incomplete pause analysis. If removing these p/np units still shows the signals, then the results from Section 4.3 are valid regardless of whether the protocol analysis is complete. If the signals are weakened by removing the Case 4 p/np units, then at least part of the signal evident in Section 4.3 is caused by the incompleteness of the protocol analysis.

This section follows the same pattern as Section 4.3. However, all of the results in this section look at the placement of events on just p/np units that have events on them (Case 1 and 2). This means that p/np units with no agreed events on them (Case 3 and 4) are not considered in the data pools.

Table 4.11 shows the new results obtained by removing the Case 4 p/np units from the data originally considered in Table 4.5. This table strengthens the significance of the distribution of events with respect to np units (from 0.050 to 0.005) and causes the distribution of events with respect to p units to become significant (from 0.950 to 0.005). This is evidence that missing Case 4 p/np units are *not* confounding the results and that the results in Table 4.5 are valid regardless of the completeness of Stauffer's protocol analysis. In fact, deleting Case 4 units strengthens the statistical significance of the results and hence suggests that incomplete protocol data could be weakening the analysis in Section 4.3.

This pattern continues when we consider the placement of operator and episode events as shown in Table 4.12. Comparing this to Table 4.7 shows that

Table 4.12: Alpha Values for Operators and Episodes, Cases 1 & 2

	P/NP UNITS	P UNITS	NP UNITS
All Events:	0.005	0.005	0.005
Operators:	0.005	0.005	0.005
Episodes:	0.005	0.950	0.950

Table 4.13: Alpha Values for Speech and Marks, Cases 1 & 2

	P/NP UNITS	P UNITS	NP UNITS
Combined:	0.005	0.005	0.005
Speech:	0.005	0.005	0.050
Marks:	0.005	0.005	0.050

only episodes are still showing up as randomly distributed. But even episodes are non-random when the entire p/np unit is considered. It is interesting to note that episodes are still showing up as more random than operators. Again, this is not what we would expect if episodes are the more complex events and if both events can be equally detected by the timing analysis employed in this paper.

Another observation of Table 4.12 is that all variations of the data are not randomly distributed with respect to the timing data when the entire p/np unit is considered. This is more evidence that processing for Stauffer's events happens during both the pause and non-pause parts of the protocol.

Next we again separate the speech and marks p/np units and explore the effect of removing Case 4 p/np units (see Table 4.13). Comparing this table with Table 4.9, we see that much the same pattern is evident. The results aren't quite as strong as when we consider the np units: but 5% error (alpha = 0.050) is still significant.

The final analysis separates both marks and speech p/np units and episode

and operator events. Table 4.14 shows these results. When comparing this table to Table 4.10, it is clear that this table does not follow the same pattern as the previous tables in this section:

- When considering the entire p/np unit in Table 4.14 there is still only one case that is random. However, instead of being the speech timing data with respect to episodes as we would predict, it is the *marks* timing data with respect to episodes. This might be an indication that Case 4 p/np units are a confounding variable for speech timing data but actually improve the accuracy of marks timing data.

- P units in Table 4.14 follow the same pattern as the previous tables in this section. The only change is a strengthening of the non-random distribution of operators with respect to speech timing data. But the change is only 5% to 0.5%: both are statistically significant.

- The np units show the greatest change. What in Table 4.10 had been almost totally non-random has become almost totally random in Table 4.14. Again, the biggest change was with the marks timing data. This is more evidence to suggest that blank p/np units confound the speech timing data but seem to strengthen the marks timing data.

In summary, the significant relationship between timing analysis and protocol analysis is not affected by any incompleteness in the protocol data or analysis. The slight changes that are observed when Case 4 p/np units (in other words, those most likely to be incomplete in the protocol data) are omitted suggest that the protocol data is indeed incomplete.

Table 4.14: Alpha Values for Eps and Ops on Speech and Marks, Cases 1 & 2

P/NP UNITS:

	Episodes	Operators
Speech:	0.005	0.005
Marks:	0.950	0.005

P UNITS:

	Episodes	Operators
Speech:	0.950	0.005
Marks:	0.005	0.005

NP UNITS:

	Episodes	Operators
Speech:	0.950	0.050
Marks:	0.950	0.950

Chapter 5

Conclusions

There are two primary results from this study. The first primary result has to do with the reliability of Stauffer's protocol analysis. Since the two independent experimenters were able to agree on the existence *and* the exact placement of most of the TEA Model events, there is evidence that the TEA Model produced by Stauffer in his protocol analysis is a reliable model of at least some of the cognitive processes occurring during a mechanical engineer's design.

The second primary result from this study is that there is a strong relationship between the TEA Model events in Stauffer's protocol analysis and timing analysis. Hence, to the extent that either technique is actually measuring the cognitive processes of designers, this suggests that protocol analysis and timing analysis are measuring the *same* processes.

Besides these two primary results from this study, there are three secondary results having to do with the relationship between the TEA Model and timing analysis. First, the relationship between the timing data and TEA Model operators is stronger than the relationship between the timing data and TEA Model episodes. This was an unexpected result, based on the TEA Model hierarchy. There are several hypotheses that can explain this:

- TEA Model operators are more significant cognitive events than TEA Model episodes.

- There are not enough TEA Model episodes in the data to show a significant pattern with respect to the timing data.

- The grain size of the timing analysis was either too large or too small to be able to detect the TEA Model episodes.

- TEA Model episodes are measuring cognitive events in a manner that is not detected well by timing analysis as performed in this study.

There is no strong evidence from this study to support any one hpyothesis over another.

Another secondary result is that TEA Model events have a stronger correlation to marks timing data than to speech timing data. This supports the hypothesis that the experimental demand for the subject to "keep talking" reduces the reliability of the speech timing data. Because there was no experimental requirement for subjects to "keep drawing", it is reasonable to conclude that the marks timing data is more reliable. The data analysed in this paper supports the conclusion that the demand to produce externally visible data is confounding.

The final secondary result is that TEA Model events are placed less randomly when only Case 1 and 2 p/np units are considered. Because Case 4 p/np units are most likely to reflect incompleteness in the protocol data and protocol analysis, this analysis with only Case 1 and 2 units shows that the relationship between timing analysis and protocol analysis is not an artifact of the incompleteness of protocol analysis. In fact, because removing Case 4 p/np units strengthens the relationship, this suggests that incompleteness in the protocol data partly obscures the relationship. It also provides evidence that the protocol data is indeed incomplete.

From these results, we can conclude that timing analysis and protocol analysis are reliable methods for observing the mechanical design process. Furthermore, there is much to be gained by combining them. This is because the two methods have complementary strengths and weaknesses. Protocol data provides valuable information about the content of design activity, but it is incomplete and time-consuming to perform. Timing analysis provides no information about content, but it is complete and easy to perform (in fact, it can be completely automated). Future protocol studies should include a timing analysis as well, since this can help identify

and measure incompletenesses in the protocol data and provide a reliability check on the analysis.

Chapter 6

Future Research

There are several directions of research that are suggested by this initial study. Four main areas of future research are considered in this section.

The most obvious question left by this study is the exact relationship between the TEA Model events and the timing data. All that is noted here is whether or not the distribution of events is random with respect to the length of the timing data. But how are they placed? Based on previous timing analyses, we would expect that there are more events on longer p/np units. Is this the case? This is an area of research that could be completed fairly easily, working with the data and results of this study. It may be that, when the number and kinds of events are taken into consideration, stronger patterns are observable.

The second important research direction is to re-evaluate the TEA Model using a combination of timing and protocol analysis. Since the timing data does not correlate strongly with the episodes, the psychological reality of episodes is questionable. Do they exist as identifiable events? If so, why are they not more strongly detected by timing analysis, as the operators are? A detailed comparison of the pauses near episode boundaries might provide some answers to these questions.

Another question about the TEA Model concerns the reality of particular kinds of operators. In this study, all operators were lumped together as one type of event. Are some operators more easily detected by timing data? What changes to the TEA Model would need to be made if all events without timing data support are removed from the protocol analysis?

A third, more speculative direction for research concerns the basic assump-

tion of timing analysis – namely, that pauses occur during cognitive activity that *prepares for* and precedes an externally observable action. An alternative hypothesis is that pauses mark periods during which the subject is reflecting on previous externally observable actions. By rearranging that data to form non-pause/pause (np/p) units rather than p/np units, this issue could be explored. A further direction concerns timing analysis of larger units of activity such as entire events (rather than just the first externally observable action in the event) or other strings of operators. Verbalizations involving no content (for example, "umm" and "uhh") could be deleted and the data reanalyzed.

Another direction to consider with the timing analysis is the effect of granularity on the relationship to the TEA Model. The data described in this paper was averaged over one second, but there is no indication of what results might have been obtained if the data had been averaged over a smaller granularity (such as one-half second) or a larger granularity (such as two seconds). Does a different granularity increase the relationship between timing and pause analysis? Perhaps the correlation between just some of the events (such as the episodes) is increased. Does a different granularity decrease the relationship? Or is there no significant change? Study in this area could lead to a more complete, accurate technique for automatically tracking design.

Finally, an interesting area of research would be to pursue the difference in designing style between subjects and between design areas using timing analysis. There are good reasons to expect many differences in design style among different engineers and among engineers working in different types of problems or different fields. Generally, protocol analysis is too cumbersome to permit large-scale studies of these differences. However, timing analysis could make such a project feasible.

Bibliography

[1] Akin, O (1986). *Psychology of Architectural Design.* Pion Limited, England.

[2] Akin, O (1978). How do architects design? *Artificial Intelligence and Pattern Recognition in Computer Aided Design,* In Latcomb, (Ed.), North-Holland Publishing Company, 65-103.

[3] Byrne, R. (1981). Mental cookery: An illustration of fact retrieval from plans. *Quarterly Journal of Experimental Psychology, 33A,* 31-37.

[4] Carroll, D.W. (1986). *Psychology of Language.* Brooks and Cole.

[5] Chase, W.G, & Simon, H.A. (1973). Perception in chess. *Cognitive Psychology, 4(1),* 55-81.

[6] Dabbs, J.M., & Ruback, R.B. (1984). Vocal patterns in male and female groups. *Personality and Social Psychology Bulletin, 10(4),* 518-525.

[7] Duez, D. (1982). Silent and non-silent pauses in three speech styles. *Language and Speech, 25(1),* 11-28.

[8] Egan, D. & Greeno, J. (1974). Theory of Inductive Learning: Knowledge Acquired in Concept Identification, Serial Pattern Learning and Problem Solving. In L. Gregg (Ed.), *Knowledge and Cognition.* Hillsdale, N.J.: Erlbaum.

[9] Ericsson, K.A. & Simon, H.A. (1980). Verbal reports as data. *Psychological Review, 87,* 215-251.

[10] Ericsson, K.A., & Simon, H.A. (1984). *Protocol Analysis.* Cambridge, MA: MIT Press.

[11] Ernst, G. W., & Newell, A. (1969). *GPS: A Case Study in Generality and Problem Solving.* New York: Academic Press.

[12] Fisher, C.A. (1973). "Advancing the Study of Programming With Computer-Aided Protocol Analysis"

[13] Giffen, G., & Luszcz, M.A. (1983). Are the spoken durations of rare words longer than those of common words? *Memory and Cognition, 11(1),* 13-15.

[14] Glass, A. L., & Holyoak, K.J. (1986). *Cognition.* New York: Random House.

[15] Hog, R.V., & Tanis, E.A. (1977). *Probability and Statistical Inference.* MacMillan Publishing Co., New York. Collier MacMillan Publishers, London.

[16] Horii, Y (1983). An automatic analysis method of utterance and pause lengths and frequencies. *Behavior Research Methods & Instrumentation, 7,* 449-452.

[17] Miller, K., Perlmutter, M. & Keating, D. (1984). Cognitive arithmetic: Comparison of operations. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 10(1),* 46-60.

[18] Moyer, R., & Landauer, T. (1967). The time required for judgements of numerical inequality. *Nature, 215,* 1519-1520.

[19] Newell, A. & Simon, H.A. (1972). *Human Problem Solving.* Prentice Hall, Englewood Cliffs, NJ.

[20] Payne, J.W., Braunstein, M.L., & Carroll, J.S. (1978). Exploring predecisional behavior: An alternative approach to decision research. *Organizational Behavior and Human Performance, 22,* 17-44.

[21] Restle, F. (1970). Speed of adding and comparing numbers. *Journal of Experimental Psychology, 85,* 274-278.

[22] Ross, S. (1976). *A First Course in Probability.* Macmillan, New York, NY.

[23] Saslow, C.A. (1982). *Basic Research Methods.* Addison-Wesley, Menlow Park, CA.

[24] Schweiger, D.M. (1983). Is the simultaneous verbal protocol a viable method for studying managerial problem solving and decision making? *Academy of Management Journal, 26(1),* 185-192.

[25] Stauffer, L. (1987). "An Empirical Study on the Process of Mechanical Design", Oregon State University Thesis, Sept 1987, Corvallis, OR.

[26] Stauffer, L., Ullman, D.G., & Dietterich, T.G. (1987). "Protocol Analysis of Mechanical Engineering Design." Proceedings of the 1987 International Conference on Engineering Design, WDK 13, Boston, MA, August 1987, 68-73.

[27] Tikerpuu, J. & Ullman, D.G. (1988). "General Feature Based Frame Representation For Describing Mechanical Engineering Design Developed From Empirical Data."

[28] Trimboli, C., & Walker, M.B. (1984). Switching Pauses in Cooperative and Competitive Conversations. *Journal of Experimental Social Psychology, 20(4),* 297-311.

[29] Ullman, D. G., Dietterich, T.G. & Stauffer, L.A. (1988). "A Model of the Mechanical Design Process Based on Empirical Data: A Summary". *Artificial Intelligence in Engineering, Design, and Manufacturing, 2(1),* 33-52.

[30] VanLehn, K. & Garlick, S. (1987). "Cirrus: an Automated Protocol Analysis Tool". Proc. of the 4th International Workshop on Machine Learning, UC Irvine, June 1987, 205-217.

[31] vonSzeliski, V. (1924). Relation between the quantity perceived and the time of perception. *Journal of Experimental Psychology, 7,* 135-147.

[32] Walpole, R.E. (1974). *Introduction to Statistics* Second Edition. MacMillan Publishing Co., New York. Collier MacMillan Publishers, London.

Appendices

Appendix A

Combined Timing and Protocol Data

The data for Subject 1 Conceptual Design is listed below. Data that was omitted because of experimenter interaction or topics other than the design is placed between slash/stars: /* data omitted */.

The timing seconds are indicated as follows:

"." - indicates one second of pause

"/" - indicates one second of speech alone

"-" - indicates one second of marks alone

"+" - indicates one second of speech and marks

SUBJECT 1 CONCEPTUAL DESIGN

EPS12S
0:27:21 .../ Okay.
EPS11S
0:27:25 .../ Have to decide
0:27:29 ..../ how to
sel[cons1: connect batteries in series]1S:2S
0:27:34 ...// connect the batteries in series
sel[cons2: connect batteries to pc board]1S:2S
0:27:39 ..// and also how to connect 'em to the
0:27:43 ....../ printed circuitboard contacts.
cre[cons3: batteries must be in bottom half of envelopes]1S:2S
0:27:50 ../ And they'll be on the bottom,
0:27:53 ....// and these have to be connected in series,
0:27:59 .../ so
0:28: 3 .../ I'll have to
sel[prop1: electrical contact]1S:2S
cal[stra1: design prop1 to connect from the bottom to top of batteries]1S:2S

acc[stra1]1U:2S
0:28: 7 ...../ / design a contact to touch
0:28:14 ./ the bottom and the top of the batteries
0:28:16 ..../ And do that
EPS21S:2S
cre[prop2: configure outer batteries right-side-up and middle battery
upside down]1S
0:28:21 ..../ so that they're lined up in series.
cre[prop2: configure outer batteries right-side-up and middle battery
upside down]2S
acc[prop2]2S
0:28:26 ../ Looks like I can do two of them right side up,
0:28:29 .// the two outer ones;
0:28:32 ./ one of them upside down,
acc[prop2]1U
ref[prop1→prop3: four contacts from board to config in prop2]1S:2S
0:28:34 ...//// and then connect from the top of one to the top of the other,
and then from the bottom of one to the bottom of the other.
0:28:41 ../ And from the top to
0:28:44 .../ the contact device
0:28:48 ./ that will connect to the board.
0:28:50 ../ And then the same on
0:28:53 ../// the bottom of... of this will connect to the negative
0:28:58 ../ component.
0:29: 1 ../ The negative is the
0:29: 4 ../ where the battery
0:29: 7 ...// is the smaller part - the top part.
acc[stra3]1S
acc[prop3]2S
EPS31S:2S
0:29:12 ..../ So,
0:29:20 ../ and look at ways
0:29:23 ../ to
cre[stra2: determine envelope space constraint]1S:2S
0:29:23 .../ first of all determine
acc[prop2]1S
0:29:30 .// how much inside room I have for the batteries
acc[prop2]2S
cre[stra3: determine where contacts will go]1S:2S
acc[stra3]1S
0:29:33 ./ and then where it's going to go.
acc[stra3]2S
EPS41S:2S
sel[cons4: outside height must be .258 +/- .006]1S:2S
0:29:35 ..// So the height that I'm given is the outside height,
1
0:29:39 .-+ is .258.
2

0:29:42 .-+/ plus or minus .006,
sel[prop4: battery height is .213]1S:2S
0:29:46 ../ and the height of the battery
3
0:29:49 .-+ is
4
0:29:53 -.+++ given as 5.4 millimeters,
5
0:29:58 -/+ which in inches is
0:30: 2 ..../ 2.58 -.258.
cal[cons5: 0.046 available for wall thickness and separator]1S:2S
0:30: 7 ....// So that's 5.4
0:30:16 ..../ millimeters and
0:30:21 .// is .54 centimeters
0:30:24 ..// times 2.54
0:30:28 ....// divided by 2.54.
0:30:34 ...../ I'll get
6
0:30:40 ....-+ 2.
7
0:30:48 -..+ .2 .2 inches as the battery height.
0:30:52 ——./ So which leaves me,
8
0:30:59 .++ taking the envelope height
0:31: 2 -..../ which is
0:31: 8 ../ .258
0:31:11 ./ subtract that
0:31:13 ./ subtract .212 from that,
9
0:31:15 .-+ I get 0.6
0:31:18 -./ as
11
0:31:22 ....-+ the wall thickness
0:31:28 -./ of the
0:31:31 ./ envelope
0:31:33 -../ and also the seperation
12
0:31:38 ..../+ thickness and the seperator thickness.
13
acc[cons5]1S:2S
0:31:44 ——-...+/ Okay, that's not going to leave me much room
0:31:57 .../ at all in there.
EPS51S:2S
sel[prop5: wall thickness]1S:2S
ref[prop5→prop7: pretty thin wall thickness]1S
sel[prop6: separator]1S:2S
com[prop5,6 TO cons5]1S:2U
0:32: 1 ....// So I have to go for a pretty thin sheet

0:32: 7 ./ of plastic

ref[prop6→prop8: pretty thin separator]1S:2S

0:32: 9 ...// and also a pretty thin separator thickness,

0:32:14 ..../ Or,

sel[cons6: need to hold batteries in place]1S:2S

com[prop5,6 TO cons 5,6]1S:2S

0:32:24 ...../ I could use a separator to

0:32:30 ./ hold the batteries in place

0:32:32 ../// and have the batteries;

ref[prop6→prop9: put holes in separator to hold batteries]1S:2S

0:32:37 ./ and the separator have holes in it and the batteries

0:32:39 ../ fall into the holes

sus(no choice made between both ideas)1S

0:32:42 .../ And

sus(no choice made between both ideas)2S

/*EPS61S:2S

cre[cons7: it is optional whether or not there exist a bottom
envelope]1S:2S

0:32:46 ...../ can I design the circuit board at all?

0:32:51 .../ E: Pardon me? S: Can I design the circuit board at all?

14

0:32:55 ..-+ E: Umm..

0:32:59 -+ S: For instance, if I wanted to add a part onto the circuit board

0:33: 1 -/ that the battery would fall onto

0:33:03 ...//////// E: That's the plastic envelope; you can work with that.
I mean that's up to you to design.

S: Okay. Does the plastic envelope then have a bottom?

Or is that up; again

0:33:14 ./// up to me to decide? E: It's up to you.

cre[cons8: only X1 and X2 dimensions on pc board can be altered]1S:2S

0:33:17 ./// E: But the... S: Okay. E: The P.C. board is fixed, the only
thing that is not fixed is on these two, X1 and X2 branches.

acc[cons7,8]1S:2S

0:33:21 ...../ S: Okay.

0:33:27 ..../////////// E: But like, say the distance from the P.C. board to
the envelope is a fixed distance.

S: All right.

E: So you can't change that.

S: I was wondering if the envelope had a bottom to it

15

0:33:43 .-+ For instance, does it just have

16

0:33:46 -.+ it's like that.

0:33:50 -/// E: That's up to you.

S: That's up to me?

E: Right.

S: Okay.

E: These are all just exterior

0:33:55 ./ dimensions.
0:33:57 ...// S: Alright.
E: So it's up to you to come up with somthing that fits in.*/
EPS71S:2S
sel[prop4: battery height is .213]1S:2S
0:34: 2 ..../ I want to double check that
0:34: 7 ./ battery height.
0:34: 9 ....// It was .54 centimeters times 2.54
/* 0:34:15 ../// divided by .254.
E: This is a drawing of the battery.
sel[cons9: use round numbers]1S:2S
sel[cons11:??]1U
cal[cons10: 0.045 available for wall thickness]1S
0:34:20 ......// Okay. So it gives it .213 inches */
16.5 17
cal[cons10: 0.045 available for wall thickness]2S
0:34:28 ...—-.+/ .045 wall thickness
18
acc[cons10]2S
0:34:38 -.+/ only on that.
19
acc[cons10]1U
0:34:43 ./+++ And umm this can be wall thickness and contact
0:34:48 -+ thickness.
0:34:50 -......./ So,
EPS81S:2S
sel[prop10: location of contacts]1S:2S
sel[cons11:??]1U:2U
com[prop10 TO cons11]1S:2U
0:35:00 ...// as far as location of the contacts
0:35: 5 ..../ they'll have to be
0:35:10 .// located on the plastic envelope
0:35:13 ....../ So umm,
0:35:20 ..// it'll be connected between 'em.
ref[prop10→prop11: contacts are located on bottom envelope]1S
0:35:24 ...../ And the bottom half of the
0:35:30 ..// Envelope is two parts.
0:35:34 ..../ So,
ref[prop10→prop11: contacts are located on bottom envelope]2S
0:35:39 ../ I had
21
0:35:42 .....-+ one side and the bottom
pat[prop11→prop12: contacts are located on bottom envelope and to the bottom of batteries]1S:2S
0:35:50 .-++/ to figure out some way to make contact with the batteries
0:35:56 ../ on that bottom.
pat[prop11→prop13: contacts are located on side of envelope]1S:2S
0:35:59 ....// Or, maybe I can contact the batteries on the side.

sel[cons12: info in problem statement]1S:2S

0:36: 5 ...../ It looks like I'll have to contact them on the bottom.

acc[prop12]2S

0:36:11 ..../ which

acc[prop12]1S

0:36:16 ..../ means I'll have to put something on the bottom.

EPS91S:2S

sel[prop3]1S:2S

0:36:28 ...../ The types of things

0:36:34 ../ that could go on the bottom as contacts

sel[cons13: use off-the-shelf items if possible]1S:2S

cal[stra4: look in electrcal catalog for contacts]1S:2S

acc[stra4]1S

0:36:37 ..../ / I'd have to look in an electrical catalog to look at that.

acc[stra4]2S

EPS101S:2S

sel[prop3]1S:2S

0:36:46 ....../ / So what I need would be a contact running from

22

0:36:54 ....-+ look at the bottom of the board,

0:37: 0 -+ on the envelope

EPS111S:2S

dra(prop2)1U

23

0:37: 2 .—-+ and where my three batteries are gonna be,

24 25

sel[prop2]1S:2S

0:37: 8 .—-/ I'll have

26 27

0:37:15 .-/ these two up

28 29ab

0:37:19 -.+ up being

29cd

0:37:23 —-..+ like so.

sel[cons1]1S:2S

com[prop2 TO cons1]1S:2S

acc[prop2]1U

0:37:30 -+ and this is the

0:37:32 -/ cathode, and that's the anode port.

0:37:35 ./ Those two would be up,

30

0:37:37 -.

31 32

-+ with the positive side down.

0:37:42 —+ Ok, down,

acc[prop2]2S

0:37:46 ...// and I'll need to make a connection

0:37:51 ./ from

sel[consl]1S
0:37:53 ....../ need to connect them in series
EPS121S:2U
sel[cons14: current flows from plus to minus]1S:2S
0:37:60 ..../ from the bottom of – let's see, where are my pluses.
0:38: 5 ..../ Minuses there,
0:38:10 ....// which, as far as my electrical engineering goes
0:38:16 ..// I'm not quite sure if I would connect the negative
0:38:20 ..// to the positive part of the battery
0:38:24 ../ or the negative part of the battery
33
0:38:27 ...+ I want my
0:38:31 –+ current to go
34
0:38:34 -.-+ to the plus.
0:38:39 -+ It's plus/minus, plus/minus.
35
0:38:41 ...//+ The current is gonna go from the plus side to the minus side.
0:38:47 ...// So this... this is gonna be
36
0:38:52 ...+ negative
37
0:38:56 .-/ positive, no it'd
0:39: 0 ./ better be negative
38
0:39: 2 ...-+ Wait now.
sel[cons15: electrons flow from minus to plus]2S
0:39: 7 ...// My electrons are gonna go toward the negative.
sel[cons15: electrons flow from minus to plus]1U
0:39:12 ./ Or the positive part.
0:39:14 ..../ I still don't know if that's correct.
sus[need more info]1U:2S
EPS131S:2S
cre[stra5: look up information later and continue for now]2S
0:39:22 ...../ So what I would probably do
0:39:28 .../ What I'm going to do
0:39:32 ./ I'm going to do,
cre[stra5: look up information later and continue for now]1S
0:39:34 .../ I'm gonna hold on to that and look at that.
0:39:38 ./// But for now I'm gonna keep on working on what I have, and when
I have more electrical questions, I'll
acc[stra5]1S
0:39:42 ./ look them up.
/* 0:39:44 ./ E: Good plan.
0:39:46 ./ S: Good plan. */
acc[stra5]2S
sel[prop2]2S
sel[cons14]2S

0:39:48 ..../ So, assuming that
sel[prop2]1S
com[prop3 TO cons1,14]2S
ref[prop3→prop13: contacts in fig.1, + on pc board to (-) top of
FL battery, (+) bottom FL battery to (-) top of middle battery,
(+) bottom of middle battery to (-) top of FR batter, (+)
bottom of FR battery to (-) on PC board]2S
0:39:53 .../ these will be up, so should this one.
sel[cons14]1U
com[prop3 TO cons1,14]1U
ref[prop3→prop13: contacts in fig.1, + on pc board to (-) top of
FL battery, (+) bottom FL battery to (-) top of middle battery,
(+) bottom of middle battery to (-) top of FR batter, (+)
bottom of FR battery to (-) on PC board]1U
acc[prop13]1U
dra(prop13)1U
40
0:39:57 ..-+/ The contact will go from there to
0:40: 2 ./ the
0:40: 4 ./ negative contact
41
0:40: 6 .-+ on the board.
42
0:40: 9 -...+ And the negative contact from there,
0:40:15 -/ so to the positive side of that one.
43
0:40:18 .....++ And then from the top of this
0:40:25 -++ to the top of that one.
44
0:40:29 -.-/+ Then I'll need to go from the bottom of this
0:40:35 .-/// out to the positive side of the board.
acc[prop13]2S
EPS141S:2S
cre[stra6: cehck if cons14 is correct]1S:2S
sel[stra4]1S
45
0:40:40 —-......../ So I need to
46a
0:40:53 -/ determine electrically
46b
0:40:55 -+// that that is – which way... which way to mount the batteries,
sel[stra4]2S
0:40:59 .../ and also
0:41: 3 ..../ what type of contacts there are
0:41: 8 ../ to mount the batteries
EPS151U
0:41:11 ...1// And the contacts
EPS152S

sel[cons15: contact force is .1 to 1.0 lbs]1S:2S
0:41:17 ....// a minimum force and maximum force
0:41:28 ....../ which is a tenth of a pound
0:41:35 ./ minimum at the batteries
0:41:37 ..../ so,
cre[cons16: contacts may be mounted to envelope]2S
0:41:42 ..../ I'm gonna look at
cre[cons16: contacts may be mounted to envelope]1S
0:41:47 .../ they may mount them directly to the envelope,
cal[cons17: contact force can be achieved by snapping envelope
together]1S:2S
0:41:51 ..// so especially when it's assembled and snapped into place
acc[cons17]1U
0:41:55 ..// the contact will be made and the force will be applied by the
acc[cons17]2S
0:41:59 ./ assembly of the envelope.

Appendix B

Goodness-Of-Fit Data

The following abbreviations are made in this data:

Pau	- Pause Length (Trial Length)
Eps	- Episodes
noblank	- Cases 1 and 2: no Case 4 units
blank	- Cases 1, 2, and 4
Ops	- Operators

The data is listed in the following order:

Speech and Marks P Units
Speech P Units
Marks P Units
Speech and Marks NP Units
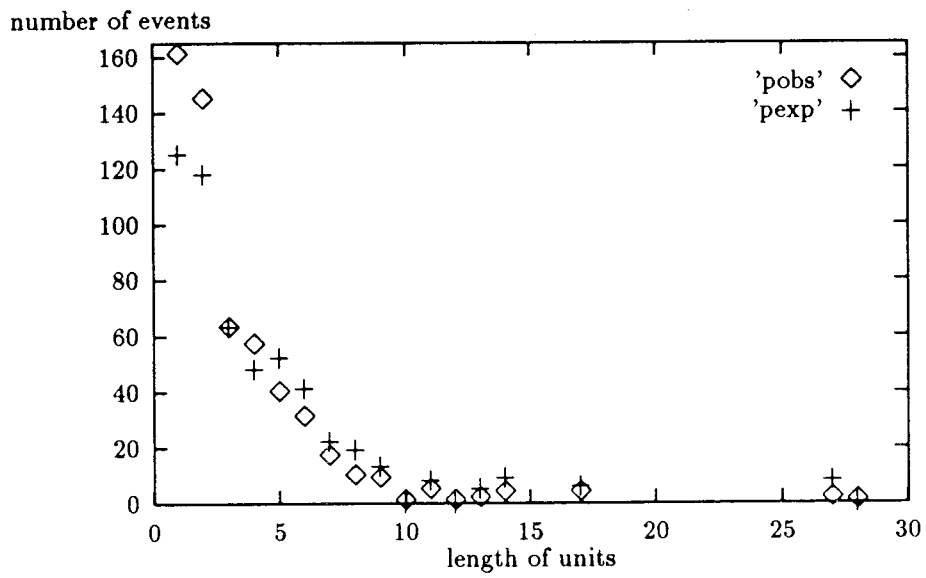Speech NP Units
Marks NP Units

Figure B.1: Observed and Expected Data for Speech and Marks P Units

Figure B.1 is a plot of the observed and expected data as listed in the chart below. The specific plot is of

- all events

- on speech and marks pause units

- cases 1, 2, and 4

The x-axis is the number of events, the y-axis is the length of the pause. The "◇"'s represent the observed data at each point and the "+"'s represent the expected data at each point. For example, for cases 1, 2, and 4 pause units of length 1, there were 160 observed events and 125 expected events.

Speech and Marks P Units

Pau	Eps noblank	Eps blank	Ops noblank	Ops blank	Evs noblank	Evs blank
1	o: 26 e: 15	o: 26 e: 35	o:134 e:103	o:134 e:135	o:160 e:125	o:160 e:167
2	o: 23 e: 18	o: 23 e: 24	o:121 e: 98	o:121 e: 97	o:144 e:118	o:144 e:120
3	o: 12 e: 9	o: 12 e: 14	o: 50 e: 49	o: 50 e: 56	o: 62 e: 63	o: 62 e: 71
4	o: 13 e: 13	o: 13 e: 11	o: 43 e: 38	o: 43 e: 42	o: 56 e: 48	o: 56 e: 52
5	o: 13 e: 14	o: 13 e: 6	o: 26 e: 40	o: 26 e: 24	o: 39 e: 52	o: 39 e: 31
6	o: 9 e: 13	o: 9 e: 7	o: 21 e: 27	o: 21 e: 23	o: 30 e: 41	o: 30 e: 31
7	o: 4 e: 6	o: 4 e: 3	o: 12 e: 16	o: 12 e: 12	o: 16 e: 22	o: 16 e: 16
8	o: 2 e: 4	o: 2 e: 3	o: 7 e: 17	o: 7 e: 12	o: 9 e: 19	o: 9 e: 15
9	o: 2 e: 4	o: 2 e: 1	o: 6 e: 11	o: 6 e: 6	o: 8 e: 13	o: 8 e: 8
10	o: 0 e: 0	o: 0 e: 1	o: 0 e: 0	o: 0 e: 3	o: 0 e: 0	o: 0 e: 3
11	o: 2 e: 4	o: 2 e: 2	o: 2 e: 7	o: 2 e: 6	o: 4 e: 8	o: 4 e: 8
12	o: 0 e: 0	o: 0 e: 1	o: 0 e: 0	o: 0 e: 2	o: 0 e: 0	o: 0 e: 2
13	o: 0 e: 0	o: 0 e: 0	o: 1 e: 4	o: 1 e: 2	o: 1 e: 5	o: 1 e: 3
14	o: 2 e: 5	o: 2 e: 1	o: 1 e: 4	o: 1 e: 1	o: 3 e: 9	o: 3 e: 3
17	o: 1 e: 3	o: 1 e: 0	o: 2 e: 5	o: 2 e: 1	o: 3 e: 6	o: 3 e: 2
27	o: 0 e: 0	o: 0 e: 0	o: 1 e: 7	o: 1 e: 2	o: 1 e: 8	o: 1 e: 3
28	o: 0 e: 0	o: 0 e: 1	o: 0 e: 0	o: 0 e: 2	o: 0 e: 0	o: 0 e: 3
ave:	o:109 e:109	o:109 e:109	o:427 e:427	o:427 e:427	o:536 e:536	o:536 e:536
chi:	0.00	10.00	0.00	6.00	0.00	11.00
dof:	12.00	10.00	14.00	17.00	14.00	17.00
alpha:	0.995	0.950	0.995	0.990	0.995	0.950

Speech P Units

Pau	Eps noblank	Eps blank	Ops noblank	Ops blank	Evs noblank	Evs blank
1	o: 25 e: 5	o: 25 e: 13	o: 69 e: 15	o: 69 e: 38	o: 88 e: 17	o: 88 e: 44
2	o: 8 e: 2	o: 8 e: 12	o: 37 e: 13	o: 37 e: 36	o: 41 e: 15	o: 41 e: 42
3	o: 3 e: 2	o: 3 e: 5	o: 27 e: 15	o: 27 e: 20	o: 30 e: 18	o: 30 e: 23
4	o: 3 e: 2	o: 3 e: 3	o: 7 e: 6	o: 7 e: 9	o: 10 e: 7	o: 10 e: 10
5	o: 1 e: 0	o: 1 e: 3	o: 7 e: 11	o: 7 e: 14	o: 7 e: 13	o: 7 e: 16
6	o: 6 e: 5	o: 6 e: 6	o: 7 e: 8	o: 7 e: 15	o: 9 e: 10	o: 9 e: 18
7	o: 4 e: 2	o: 4 e: 3	o: 10 e: 6	o: 10 e: 10	o: 12 e: 9	o: 12 e: 12
8	o: 6 e: 5	o: 6 e: 5	o: 27 e: 25	o: 27 e: 23	o: 32 e: 29	o: 32 e: 26
9	o: 3 e: 1	o: 3 e: 3	o: 16 e: 8	o: 16 e: 10	o: 18 e: 9	o: 18 e: 12
10	o: 2 e: 2	o: 2 e: 2	o: 11 e: 13	o: 11 e: 10	o: 13 e: 15	o: 13 e: 12
11	o: 1 e: 0	o: 1 e: 3	o: 5 e: 3	o: 5 e: 9	o: 5 e: 4	o: 5 e: 11
12	o: 3 e: 2	o: 3 e: 2	o: 3 e: 6	o: 3 e: 6	o: 5 e: 8	o: 5 e: 8
13	o: 2 e: 2	o: 2 e: 2	o: 10 e: 9	o: 10 e: 7	o: 11 e: 10	o: 11 e: 8
14	o: 1 e: 1	o: 1 e: 1	o: 3 e: 2	o: 3 e: 2	o: 4 e: 4	o: 4 e: 4

```
15 o:  1 e:  1 o:  1 e:  1 o:  2 e:  2 o:  2 e:  2 o:  3 e:  2 o:  3 e:  3
16 o:  1 e:  1 o:  1 e:  1 o: 10 e:  7 o: 10 e:  6 o: 11 e:  8 o: 11 e:  7
17 o:  0 e:  0 o:  0 e:  1 o:  2 e:  4 o:  2 e:  5 o:  2 e:  5 o:  2 e:  6
18 o:  2 e:  2 o:  2 e:  2 o:  0 e:  0 o:  0 e:  1 o:  1 e:  2 o:  1 e:  3
19 o:  1 e:  1 o:  1 e:  1 o:  1 e:  2 o:  1 e:  3 o:  2 e:  3 o:  2 e:  4
20 o:  2 e:  1 o:  2 e:  1 o:  9 e:  7 o:  9 e:  5 o: 11 e:  8 o: 11 e:  6
21 o:  0 e:  0 o:  0 e:  0 o:  1 e:  2 o:  1 e:  1 o:  1 e:  3 o:  1 e:  2
22 o:  4 e:  4 o:  4 e:  2 o:  2 e:  3 o:  2 e:  2 o:  6 e:  9 o:  6 e:  6
23 o:  1 e:  1 o:  1 e:  1 o:  4 e:  3 o:  4 e:  2 o:  4 e:  3 o:  4 e:  2
24 o:  0 e:  0 o:  0 e:  1 o:  5 e:  7 o:  5 e:  8 o:  5 e:  8 o:  5 e:  9
25 o:  0 e:  0 o:  0 e:  1 o:  0 e:  0 o:  0 e:  2 o:  0 e:  0 o:  0 e:  2
26 o:  2 e:  2 o:  2 e:  1 o:  2 e:  3 o:  2 e:  2 o:  4 e:  3 o:  4 e:  3
27 o:  1 e:  2 o:  1 e:  1 o:  2 e:  3 o:  2 e:  2 o:  3 e:  3 o:  3 e:  2
29 o:  0 e:  0 o:  0 e:  0 o:  4 e:  3 o:  4 e:  3 o:  4 e:  4 o:  4 e:  3
30 o:  0 e:  0 o:  0 e:  0 o:  1 e:  4 o:  1 e:  2 o:  1 e:  4 o:  1 e:  3
31 o:  0 e:  0 o:  0 e:  0 o:  1 e:  4 o:  1 e:  2 o:  1 e:  4 o:  1 e:  3
33 o:  0 e:  0 o:  0 e:  1 o:  0 e:  0 o:  0 e:  2 o:  0 e:  0 o:  0 e:  3
34 o:  1 e:  2 o:  1 e:  1 o:  1 e:  4 o:  1 e:  3 o:  1 e:  5 o:  1 e:  3
35 o:  0 e:  0 o:  0 e:  0 o:  1 e:  4 o:  1 e:  2 o:  1 e:  4 o:  1 e:  3
36 o:  1 e:  2 o:  1 e:  2 o:  0 e:  0 o:  0 e:  3 o:  0 e:  0 o:  0 e:  3
37 o:  1 e:  0 o:  1 e:  0 o:  2 e:  0 o:  2 e:  0 o:  3 e:  1 o:  3 e:  0
38 o:  0 e:  0 o:  0 e:  0 o:  3 e:  5 o:  3 e:  3 o:  3 e:  5 o:  3 e:  3
40 o:  0 e:  0 o:  0 e:  0 o:  8 e:  7 o:  8 e:  5 o:  8 e:  8 o:  8 e:  5
44 o:  0 e:  0 o:  0 e:  0 o:  2 e: 11 o:  2 e:  7 o:  2 e: 12 o:  2 e:  8
46 o:  0 e:  0 o:  0 e:  0 o:  7 e: 10 o:  7 e:  7 o:  7 e: 12 o:  7 e:  8
48 o:  0 e:  0 o:  0 e:  0 o:  1 e:  6 o:  1 e:  4 o:  1 e:  6 o:  1 e:  4
51 o:  1 e:  3 o:  1 e:  1 o:  2 e:  6 o:  2 e:  4 o:  2 e:  7 o:  2 e:  4
57 o:  1 e:  4 o:  1 e:  2 o:  2 e:  7 o:  2 e:  4 o:  2 e:  8 o:  2 e:  5
61 o:  3 e:  4 o:  3 e:  2 o: 19 e: 10 o: 19 e:  7 o: 22 e: 12 o: 22 e:  8
63 o:  1 e:  4 o:  1 e:  2 o:  7 e: 14 o:  7 e:  9 o:  8 e: 17 o:  8 e: 11
73 o:  3 e:  5 o:  3 e:  2 o:  7 e:  9 o:  7 e:  6 o:  8 e: 10 o:  8 e:  7
79 o:  4 e:  7 o:  4 e:  3 o: 15 e: 22 o: 15 e: 14 o: 18 e: 26 o: 18 e: 17
87 o:  2 e:  5 o:  2 e:  3 o:  8 e: 10 o:  8 e:  7 o: 10 e: 11 o: 10 e:  8
90 o:  1 e:  0 o:  1 e:  0 o:  4 e: 10 o:  4 e:  7 o:  5 e: 12 o:  5 e:  8
99 o:  1 e:  6 o:  1 e:  3 o:  5 e: 13 o:  5 e:  8 o:  5 e: 15 o:  5 e:  9
116 o:  3 e: 14 o:  3 e:  7 o: 14 e: 27 o: 14 e: 17 o: 17 e: 31 o: 17 e: 20
119 o:  0 e:  0 o:  0 e:  3 o:  0 e:  0 o:  0 e:  9 o:  0 e:  0 o:  0 e: 10
122 o:  0 e:  0 o:  0 e:  0 o:  0 e:  0 o:  0 e:  1 o:  0 e:  0 o:  0 e:  1
139 o:  0 e:  0 o:  0 e:  0 o:  7 e: 16 o:  7 e: 10 o:  7 e: 18 o:  7 e: 12
168 o:  0 e:  0 o:  0 e:  0 o:  6 e: 19 o:  6 e: 12 o:  6 e: 22 o:  6 e: 14
tot:o:106 e:106 o:106 e:106 o:406 e:406 o:406 e:406 o:480 e:480 o:480 e:480
chi:      146.00        24.00       229.00       67.00       334.00      104.00
dof:       31.00        29.00        47.00       52.00        48.00       52.00
alpha:      0.005        0.95         0.005        0.10         0.005        0.005
```

          Marks P Units
Pau Eps noblank Eps blank   Ops noblank Ops blank   Evs noblank Evs blank

```
 1 o: 22 e: 15 o: 22 e: 21 o: 95 e: 67 o: 95 e: 81 o:108 e: 75 o:108 e: 93
 2 o: 21 e: 16 o: 21 e: 20 o:119 e: 91 o:119 e: 83 o:132 e:104 o:132 e: 97
 3 o: 17 e: 13 o: 17 e: 14 o: 70 e: 68 o: 70 e: 60 o: 82 e: 79 o: 82 e: 70
 4 o: 12 e: 11 o: 12 e: 12 o: 50 e: 41 o: 50 e: 46 o: 60 e: 47 o: 60 e: 53
 5 o: 12 e: 14 o: 12 e:  8 o: 25 e: 36 o: 25 e: 31 o: 35 e: 44 o: 35 e: 36
 6 o: 12 e: 12 o: 12 e:  6 o: 27 e: 34 o: 27 e: 24 o: 32 e: 39 o: 32 e: 28
```

```
 7 o:   3 e:   3 o:   3 e:   4 o:  13 e:  15 o:  13 e:  16 o:  14 e:  17 o:  14 e:  18
 8 o:   1 e:   2 o:   1 e:   2 o:   5 e:  11 o:   5 e:   7 o:   6 e:  15 o:   6 e:   9
 9 o:   2 e:   4 o:   2 e:   2 o:   8 e:  13 o:   8 e:   9 o:   9 e:  15 o:   9 e:  10
10 o:   0 e:   0 o:   0 e:   2 o:   0 e:   0 o:   0 e:   7 o:   0 e:   0 o:   0 e:   8
11 o:   1 e:   2 o:   1 e:   2 o:   3 e:   3 o:   3 e:   5 o:   4 e:   4 o:   4 e:   6
12 o:   0 e:   0 o:   0 e:   2 o:   0 e:   0 o:   0 e:   5 o:   0 e:   0 o:   0 e:   6
13 o:   2 e:   3 o:   2 e:   2 o:   3 e:   9 o:   3 e:   8 o:   4 e:  10 o:   4 e:  10
14 o:   0 e:   0 o:   0 e:   0 o:   1 e:   4 o:   1 e:   2 o:   1 e:   5 o:   1 e:   3
15 o:   0 e:   0 o:   0 e:   0 o:   0 e:   0 o:   0 e:   1 o:   0 e:   0 o:   0 e:   1
16 o:   0 e:   0 o:   0 e:   1 o:   0 e:   0 o:   0 e:   3 o:   0 e:   0 o:   0 e:   3
17 o:   2 e:   7 o:   2 e:   1 o:   3 e:  10 o:   3 e:   5 o:   4 e:  11 o:   4 e:   5
18 o:   0 e:   0 o:   0 e:   0 o:   0 e:   0 o:   0 e:   2 o:   0 e:   0 o:   0 e:   2
19 o:   0 e:   0 o:   0 e:   0 o:   0 e:   0 o:   0 e:   1 o:   0 e:   0 o:   0 e:   2
20 o:   1 e:   4 o:   1 e:   0 o:   0 e:   0 o:   0 e:   0 o:   1 e:   6 o:   1 e:   2
21 o:   0 e:   0 o:   0 e:   1 o:   0 e:   0 o:   0 e:   3 o:   0 e:   0 o:   0 e:   4
22 o:   1 e:   4 o:   1 e:   2 o:   3 e:   6 o:   3 e:   6 o:   4 e:   7 o:   4 e:   7
23 o:   0 e:   0 o:   0 e:   1 o:   0 e:   0 o:   0 e:   2 o:   0 e:   0 o:   0 e:   2
24 o:   0 e:   0 o:   0 e:   1 o:   0 e:   0 o:   0 e:   2 o:   0 e:   0 o:   0 e:   2
25 o:   0 e:   0 o:   0 e:   1 o:   0 e:   0 o:   0 e:   4 o:   0 e:   0 o:   0 e:   4
27 o:   0 e:   0 o:   0 e:   0 o:   1 e:   8 o:   1 e:   2 o:   1 e:   9 o:   1 e:   2
28 o:   0 e:   0 o:   0 e:   1 o:   1 e:   8 o:   1 e:   4 o:   1 e:   9 o:   1 e:   5
29 o:   0 e:   0 o:   0 e:   1 o:   0 e:   0 o:   0 e:   2 o:   0 e:   0 o:   0 e:   3
30 o:   0 e:   0 o:   0 e:   0 o:   1 e:   3 o:   1 e:   1 o:   1 e:   3 o:   1 e:   1
64 o:   0 e:   0 o:   0 e:   1 o:   0 e:   0 o:   0 e:   5 o:   0 e:   0 o:   0 e:   6
tot:o:109 e:109 o:109 e:109 o:428 e:428 o:428 e:428 o:499 e:499 o:499 e:499
chi:     0.00      13.00       0.00      30.00       0.00      36.00
dof:    14.00      18.00      17.00      28.00      18.00      29.00
alpha:   0.995      0.95       0.995      0.95       0.995      0.95
```
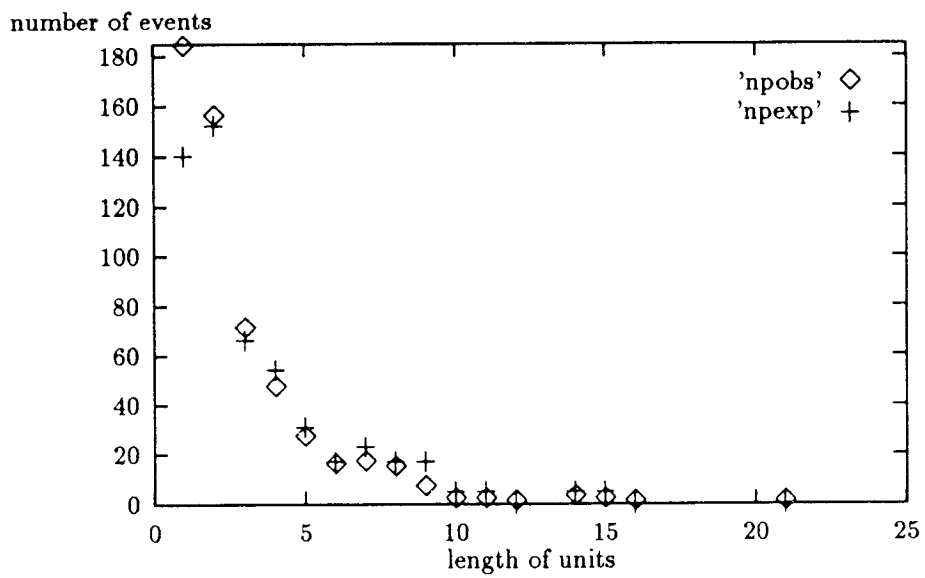
Figure B.2: Observed and Expected Data for Speech and Marks NP Units

Figure B.2 is a plot of the observed and expected data as listed in the chart below. The specific plot is of

- all events

- on speech and marks non-pause units

- cases 1, 2, and 4

The x-axis is the number of events, the y-axis is the length of the pause. The "◇"'s represent the observed data at each point and the "+"'s represent the expected data at each point. For example, for cases 1, 2, and 4 non-pause units of length 1, there were 183 observed events and 140 expected events.

Speech and Marks NP Units

Pau	Eps noblank	Eps blank	Ops noblank	Ops blank	Evs noblank	Evs blank
1	o: 46 e: 35	o: 46 e: 40	o:137 e:102	o:137 e:145	o:183 e:140	o:183 e:184
2	o: 30 e: 34	o: 30 e: 23	o:125 e:123	o:125 e: 95	o:155 e:152	o:155 e:119
3	o: 10 e: 9	o: 10 e: 14	o: 60 e: 57	o: 60 e: 57	o: 70 e: 66	o: 70 e: 71
4	o: 11 e: 12	o: 11 e: 10	o: 35 e: 45	o: 35 e: 40	o: 46 e: 54	o: 46 e: 49
5	o: 4 e: 4	o: 4 e: 6	o: 22 e: 25	o: 22 e: 26	o: 26 e: 31	o: 26 e: 32
6	o: 1 e: 2	o: 1 e: 3	o: 14 e: 15	o: 14 e: 14	o: 15 e: 17	o: 15 e: 17
7	o: 3 e: 6	o: 3 e: 4	o: 13 e: 20	o: 13 e: 15	o: 16 e: 23	o: 16 e: 19
8	o: 2 e: 3	o: 2 e: 2	o: 12 e: 15	o: 12 e: 10	o: 14 e: 17	o: 14 e: 13
9	o: 1 e: 2	o: 1 e: 1	o: 5 e: 15	o: 5 e: 7	o: 6 e: 17	o: 6 e: 9
10	o: 0 e: 0	o: 0 e: 1	o: 1 e: 4	o: 1 e: 5	o: 1 e: 5	o: 1 e: 6
11	o: 1 e: 3	o: 1 e: 1	o: 0 e: 0	o: 0 e: 3	o: 1 e: 5	o: 1 e: 5
12	o: 0 e: 0	o: 0 e: 0	o: 0 e: 0	o: 0 e: 1	o: 0 e: 0	o: 0 e: 1
14	o: 0 e: 0	o: 0 e: 1	o: 2 e: 4	o: 2 e: 5	o: 2 e: 5	o: 2 e: 6
15	o: 0 e: 0	o: 0 e: 0	o: 1 e: 4	o: 1 e: 1	o: 1 e: 5	o: 1 e: 2
16	o: 0 e: 0	o: 0 e: 0	o: 0 e: 0	o: 0 e: 1	o: 0 e: 0	o: 0 e: 2
21	o: 0 e: 0	o: 0 e: 1	o: 0 e: 0	o: 0 e: 2	o: 0 e: 0	o: 0 e: 2
tot:	o:109 e:109	o:109 e:109	o:427 e:427	o:427 e:427	o:536 e:536	o:536 e:536
chi:	0.00	2.00	0.00	5.00	0.00	4.00
dof:	10.00	12.00	12.00	16.00	13.00	16.00
alpha:	0.995	0.995	0.995	0.995	0.995	0.995

Speech NP Units

Pau	Eps noblank	Eps blank	Ops noblank	Ops blank	Evs noblank	Evs blank
1	o: 42 e: 37	o: 42 e: 33	o:132 e:111	o:132 e:111	o:164 e:133	o:164 e:133
2	o: 22 e: 23	o: 22 e: 21	o: 98 e:114	o: 98 e: 98	o:116 e:138	o:116 e:118
3	o: 13 e: 14	o: 13 e: 18	o: 70 e: 54	o: 70 e: 64	o: 76 e: 62	o: 76 e: 74
4	o: 14 e: 23	o: 14 e: 17	o: 37 e: 57	o: 37 e: 53	o: 44 e: 66	o: 44 e: 62
5	o: 7 e: 3	o: 7 e: 4	o: 32 e: 29	o: 32 e: 25	o: 37 e: 33	o: 37 e: 29
6	o: 3 e: 4	o: 3 e: 3	o: 4 e: 8	o: 4 e: 9	o: 6 e: 9	o: 6 e: 11
7	o: 2 e: 1	o: 2 e: 2	o: 5 e: 2	o: 5 e: 6	o: 6 e: 3	o: 6 e: 6
8	o: 1 e: 1	o: 1 e: 2	o: 3 e: 2	o: 3 e: 6	o: 4 e: 3	o: 4 e: 7
9	o: 2 e: 1	o: 2 e: 2	o: 6 e: 4	o: 6 e: 5	o: 8 e: 4	o: 8 e: 6
10	o: 0 e: 0	o: 0 e: 0	o: 1 e: 0	o: 0 e: 3	o: 0 e: 0	o: 0 e: 3
11	o: 0 e: 0	o: 0 e: 0	o: 5 e: 2	o: 5 e: 2	o: 5 e: 2	o: 5 e: 3
12	o: 0 e: 0	o: 0 e: 0	o: 0 e: 0	o: 0 e: 1	o: 0 e: 0	o: 0 e: 1
13	o: 0 e: 0	o: 0 e: 0	o: 0 e: 0	o: 0 e: 1	o: 0 e: 0	o: 0 e: 1
14	o: 0 e: 0	o: 0 e: 0	o: 5 e: 3	o: 5 e: 3	o: 5 e: 3	o: 5 e: 3

Pau	Eps noblank o	e	Eps blank o	e	Ops noblank o	e	Ops blank o	e	Evs noblank o	e	Evs blank o	e
15	0	0	0	1	2	1	2	3	2	2	2	3
16	0	0	0	0	0	0	0	1	0	0	0	1
17	0	0	0	0	0	0	0	1	0	0	0	1
18	0	0	0	0	2	1	2	1	2	1	2	1
19	0	0	0	1	0	0	0	2	0	0	0	3
21	0	0	0	0	4	4	4	3	4	5	4	3
23	0	0	0	0	1	2	1	3	1	3	1	3
24	0	0	0	0	1	3	1	2	1	3	1	2
25	1	1	1	1	3	2	3	1	4	2	4	2
26	0	0	0	0	0	0	0	0	0	0	0	0
27	2	2	2	1	2	3	2	2	4	3	4	2
28	0	0	0	0	1	3	1	2	1	3	1	2
29	0	0	0	0	0	0	0	0	0	0	0	0
32	0	0	0	0	1	3	1	2	1	4	1	2
41	0	0	0	0	2	4	2	3	2	5	2	3
43	0	0	0	0	0	0	0	1	0	0	0	1
46	0	0	0	0	0	0	0	0	0	0	0	0
66	0	0	0	0	2	7	2	5	2	8	2	5
tot	109	109	109	109	418	418	418	418	495	495	495	495
chi:	6.00		3.00		27.00		19.00		24.00		21.00	
dof:	10.00		10.00		22.00		25.00		22.00		25.00	
alpha:	0.95		0.99		0.95		0.95		0.95		0.95	

Marks NP Units

Pau	Eps noblank o	e	Eps blank o	e	Ops noblank o	e	Ops blank o	e	Evs noblank o	e	Evs blank o	e
1	63	63	63	60	167	143	167	209	207	180	207	247
2	24	23	24	21	129	134	129	94	145	151	145	109
3	7	6	7	9	48	42	48	39	54	47	54	45
4	7	6	7	6	27	37	27	28	32	41	32	32
5	2	2	2	3	14	15	14	13	15	17	15	15
6	1	1	1	2	11	12	11	8	11	14	11	10
7	1	2	1	2	6	7	6	7	7	8	7	8
8	1	2	1	2	7	9	7	7	8	9	8	8
9	0	0	0	1	0	0	0	3	0	0	0	4
10	1	0	1	1	2	1	2	2	2	1	2	2
11	0	0	0	0	1	2	1	1	1	2	1	1
12	0	0	0	0	4	6	4	3	4	7	4	3
13	0	0	0	0	0	0	0	0	0	0	0	0
14	0	0	0	0	0	0	0	1	0	0	0	2
15	1	3	1	1	3	4	3	3	3	4	3	4
16	0	0	0	0	4	2	4	1	4	3	4	1
17	0	0	0	0	0	0	0	0	0	0	0	1
19	0	0	0	0	2	6	2	2	2	7	2	3
20	0	0	0	0	1	5	1	1	1	6	1	2
21	1	1	1	0	1	2	1	0	2	2	2	1
24	0	0	0	0	1	1	1	0	1	2	1	0
25	0	0	0	0	0	0	0	1	0	0	0	1
29	0	0	0	0	0	0	0	0	0	0	0	0
30	0	0	0	1	0	0	0	2	0	0	0	2
31	0	0	0	0	0	0	0	0	0	0	0	0
33	0	0	0	0	0	0	0	0	0	0	0	0
tot	109	109	109	109	428	428	428	428	499	499	499	499
chi:	0.00		1.00		2.00		8.00		3.00		8.00	

dof:	9.00	10.00	16.00	16.00	16.00	16.00
alpha:	0.995	0.995	0.995	0.95	0.995	0.95