

Polygraph

VOLUME 28

1999

NUMBER 1

Special Edition

Chart Interpretation

The Utah Numerical Scoring System Brian G. Bell, David C. Raskin, Charles R. Honts, & John C. Kircher	1
Manually Scoring Polygraph Charts Utilizing the Seven-Position Numerical Analysis Scale at the Department of Defense Polygraph Institute Jimmy Swinford	10
Development of Deception Criteria Prior to 1950 Norman Ansley	28
Numerical Evaluation of the Army Zone Comparison Test Gary D. Light	37
Numerical Scoring Systems in the Triad of Matte Polygraph Techniques James A. Matte	46
The Academy for Scientific Investigative Training's Horizontal Scoring System and Examiner's Algorithm for Chart Interpretation Nathan J. Gordon	56
The Control Question Technique: A Search for Improved Decision Rules Eitan Elaad	65
Rank Order Analysis Kathleen Miritello	74
Scoring in a Computer Age James Wygant	77
Short Report: Proposed Method for Scoring Electrodermal Responses Donald Krapohl	82
Chart Interpretation – A Bibliography Norman Ansley	85

Published Quarterly

©American Polygraph Association, 1999

P.O. Box 8037, Chattanooga, Tennessee 37414-0037

The Academy for Scientific Investigative Training's Horizontal Scoring System and Examiner's Algorithm System for Chart Interpretation©

Nathan J. Gordon

Abstract

The Horizontal Scoring System was developed at the Academy for Scientific Investigative Training in 1981 as a means of eliminating subjectivity in scoring and the skewing of test results by the subjective selection of control questions for comparison to the relevant question. It allows for objective numerical chart analysis, while eliminating the subjectivity of the assignment of numbers from the traditional seven-position scale, and the additional subjectivity of the selection of which control question to use for comparison to the relevant question. The Examiner Algorithm System was developed at the Academy in 1997, and utilizes discrete measurements which allow the examiner to accurately and consistently determine what constitutes the greatest reaction.

In 1963, Backster developed a numerical scoring system where values ranging from a +3, to a -3, were assigned to each relevant question's independent physiological parameter after it was compared to those same parameters of a control question selected by the examiner. The decision of whether this "control vs. relevant" comparison yielded no difference (0), a slight difference (± 1), a clear difference (± 2), or a huge vs. no difference (± 3), and the corresponding score for the comparison is left to the subjectivity of the individual examiner. There are also differences in opinion on how the examiner should select the control question for comparison to the relevant question (Weaver, 1980). The Army compares the relevant question to the strongest control question, skewing their test toward truthfulness. Backster compares the relevant question to the weakest control question, skewing the test toward non-truthfulness. The University of Utah compares the relevant question to the control question preceding it, not skewing the test outcome toward any direction. Weaver also pointed out that the matter is further complicated in that there are clear differences among these three groups as to what constitutes a reaction. The Army's position is that any change from the norm constitutes a

reaction, while Backster clearly discriminates between what he defines as reaction and relief, and Utah takes a position somewhere in between.

With so much subjectivity and difference in opinions among major groups in our profession, it is not difficult to imagine that examiners of different schools of thought could have differences of opinion analyzing the same polygraph charts, even though they are all using "numerical analysis." These conditions led to our search to eliminate the subjectivity of what number to assign from the traditional seven-position scale, and which control question to select for comparison to the relevant question and to the development of the Horizontal Scoring System (Gordon & Cochetti, 1987).

In the Horizontal Scoring System the subjectivity of control question selection is totally eliminated by the examiner comparing all of the control and relevant questions in each individual parameter, creating a hierarchy of the greatest to least reaction. For example, in a standard Backster "You Phase" format there is a maximum of three control questions (numbers 46, 47, and 48) and three relevant questions (numbers 33, 35 and 37),

making a total of six reactions being compared in each individual parameter. The examiner employing Horizontal Scoring identifies the greatest reaction in thoracic breathing, and assigns it a six (6), the next greatest a five (5), the next greatest a four (4), the next greatest a three (3), the next greatest a two (2), and the smallest reaction a one (1).

The same process is then repeated in the abdominal breathing patterns, the electrodermal patterns and the cardiograph patterns. After all the hierarchies have been established, each question's pneumograph scores (thoracic and abdominal) are averaged, and then added to the electrodermal score and the cardiograph score, resulting in a total question score for each of the control and relevant questions. This total question score is then assigned a plus (+), in each of the control questions, and a minus (-), in each of the relevant questions. Since the Backster "You Phase" is a single-issue technique, the scores can be combined for a total chart score.

In a general question technique, such as Reid, Arther or the MGQT, there are four relevant questions and two control questions of varying weights. The same ranking of six to one can be performed, since one is still comparing six questions in each parameter. If two pneumographs are being used, we would again average them for a single pneumograph score, and again add them to the electrodermal and cardiograph score for each question for a total question score. We would again assign a plus (+) to the control question scores and a minus (-) to the relevant question scores. However, we cannot combine the scores since the relevant questions of inquiry do not represent a single issue. We must now compare each relevant question's total score with the total score of the control question with which it would traditionally be compared. The difference between these scores will represent that relevant question's final score. For example, if relevant question #3 had a total score of -15, we would compare it to control question #6, which had a total score of a +9, and derive the final score for relevant question #3 as a -6: $(-15) - (+9) = -6$. We would follow the same process to evaluate relevant question #5, and then use the same process to compare relevant questions #8 and 9, to control question #10.

In situations where there are equal reactions between questions in the parameter being scored, we average the scores of the positions they are vying for. For example, we have identified the greatest reaction in a given parameter, and assign it a 6. There are now two equal reactions competing for ranks of 5 and 4. We average those two numbers $(5+4 = 9; 9/2 = 4.5)$ and assign each a 4.5.

The cutoffs we are currently using are ± 1.5 per relevant question, per chart. For three charts, with three "single issue" relevant questions we use a ± 13 , and for three charts with two "single issue" relevant questions a ± 9 . For a single question (i.e., relevant question #3 in MGQT) we use a ± 3 for two charts, and a ± 4.5 for three charts. These numbers need to be reevaluated to determine if adjusting them would result in even more accurate results.

Previous research has shown the Horizontal Scoring System to be highly valid and reliable (Horvath, 1985, and, Driscoll & Honts, 1987). Both studies concluded that Horizontal Scoring was as accurate as traditional scoring, and the latter study stated that Horizontal Scoring was much easier to teach and apply. While Horizontal Scoring succeeded in removing the subjectivity of what number from a seven-position scale to apply, and which control question to select for comparison to the relevant question, it still left the question of what constitutes the "greatest" reaction to the subjectivity of the individual examiner.

In 1997, we devised a mathematical method (algorithm) for measuring changes for each physiological parameter, which in itself reflects the degree and importance of the psychophysiological reactions occurring. The algorithm works independently of any understanding of the psychological or physiological basis of why, or what is actually happening within the parameter. It is designed to simply apply objective mathematical equations to measure what is occurring, and remove examiner subjectivity. Since examiners are applying the same formula to determine the degree of reaction, reliability between examiners' interpretations of the same charts is dramatically increased. Examiner experience in chart interpretation is thereby negated.

Pneumograph

I initially theorized a method for interpreting the pneumograph that was a dual system monitoring changes in pneumograph suppression (PS) and pneumograph duration (PD). I strongly believe these two reactions reflect the major changes occurring in this parameter:

$$(PS + PD) = \text{pneumograph reaction.}$$

To confirm my theory, students at the Academy for Scientific Investigative Training were instructed to measure in millimeters the height of each of the breathing cycles in a specific question. The first four cycles following answering distortion were measured and then totaled. Each control and relevant reaction was then assigned a value for the greatest (defined by the smallest total number, which represented the greatest overall suppression) to the least reaction (Horizontal Scoring System). As previously stated, the highest value assigned was determined by the total of the number of relevant and control questions utilized in the polygraph technique being employed for the exam in question. Each pneumograph was similarly scored based on duration, and given scores from greatest to least. To do this, chart time was measured, in millimeters, from the end of exhalation in the answering distortion cycle to the beginning of inhalation in the fifth respiration cycle.

In 1997, Emanuel Cohen, one of the students attending the Academy for Scientific Investigative Training from Israel, listened to my lecture on interpreting the pneumograph based on my formula of PS + PD. He then suggested a simpler mathematical equation for establishing pneumograph reactions:

1. Measure in millimeters the heights of the first four cycles after answering distortion in the tracing being evaluated.
2. Measure the duration of these four cycles from the end of the exhalation cycle containing answering duration, until the beginning of the fifth inhalation cycle.

3. Divide the duration by the total of the four cycle heights in the reaction being evaluated, which in essence gives the amount of suppression and duration in the tracing.

4. The larger the number, the greater the reaction.

In Figure 1, the thoracic respiration has been scored using the algorithm and then assigned numbers from 6 to 1 in accordance with the Horizontal Scoring System. The greatest reaction was to question R12, and it received a 6. Questions R9 and R6 were tied as the next greatest reaction, and since they were vying for horizontal positions (ranks) 5 and 4, they were each given a 4.5, which represents the average of those two positions ($5+4 = 9$; $9/2 = 4.5$). The three control questions, C5, C8 and C11, were also tied for horizontal positions 3, 2 and 1. Each received the average of those scores ($3+2+1 = 6$; $6/3 = 2$), a 2.

In Figure 2 the abdominal respiration tracing is scored, and the two pneumograph scores for each question will be averaged, so that the pneumograph score only represents one-third of the questions total score.

Electrodermal

Measurement of the electrodermal recording is performed by multiplying the height by the base of the tracing. The height of the tracing is established by a straight line drawn from the highest peak to the base. The duration is established by measuring the distance of a straight line drawn from the beginning of the reaction, straight out until the point it intersects the downward movement of the tracing in automatic mode. Multiplying these two measurements together reflects tracing amplitude plus duration, and the larger the number, the greater the reaction. In Figure 3, the electrodermal responses are scored.

Cardiograph

The cardiograph tracing measurement is established by drawing a twenty-second straight line out from the bottom of the cardiograph tracing at the beginning of the

question. A measurement, in millimeters, is made to determine the height of any changes that occur in the baseline of the cardiograph above that line. In essence, we are measuring increases in blood volume. In Figure 4, the cardiograph is scored and tied Horizontal position scores are averaged (R6 & R12, C8 & R9).

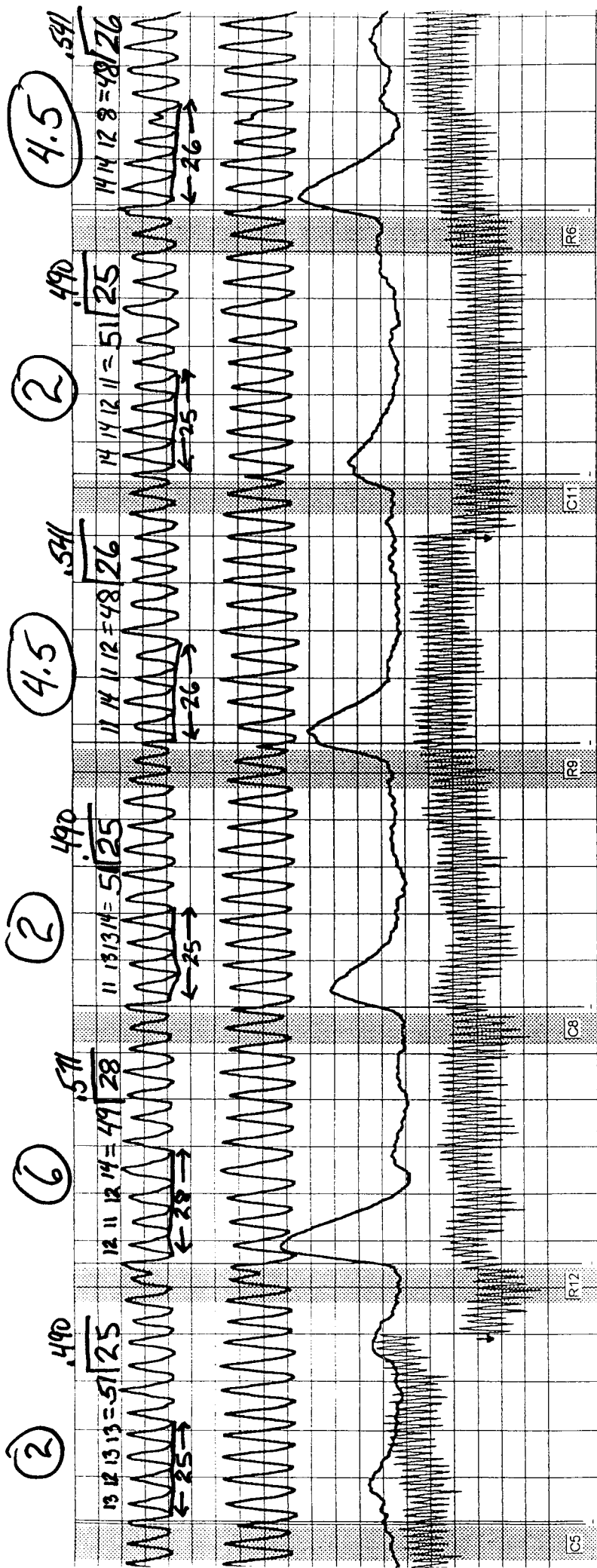
In Figure 5, the results are entered onto a Horizontal Scoring System sheet, and since it was a single-issue examination the scores can be combined, for a total chart score of -22.5. The same procedure would then be followed for the other charts in the examination and a total examination score derived.

In conclusion, the Horizontal Scoring System has been successfully used in the field since 1981, and the Examiner Algorithm System since 1997. The Academy's Horizontal Scoring System clearly removes examiner subjectivity in the assignment of numerical values to reactions and the skewing of test results by arbitrary control question selection. This system is also easier to analyze and teach, allowing for broader application. The Examiner Algorithm System simplifies and standardizes the process of determining the degree of a reaction. The reliability and validity of the algorithm in a blind evaluation study will be published later this year.

References

- Driscoll, L.N., & Honts, C.R. (1988). An evaluation of the reliability and validity of rank order and standard numerical scoring of polygraph charts. *Polygraph*, 16(4).
- Gordon, N.J., & Cochetti, P.M. (1987). The horizontal scoring system. *Polygraph*, 16(2), 116-125.
- Horvath, F. (1985). Personal communication.
- Weaver, R. S. (1980). The numerical evaluation of polygraph charts: Evolution and comparison of three major systems. *Polygraph*, 9, 94-108.

Figure 1

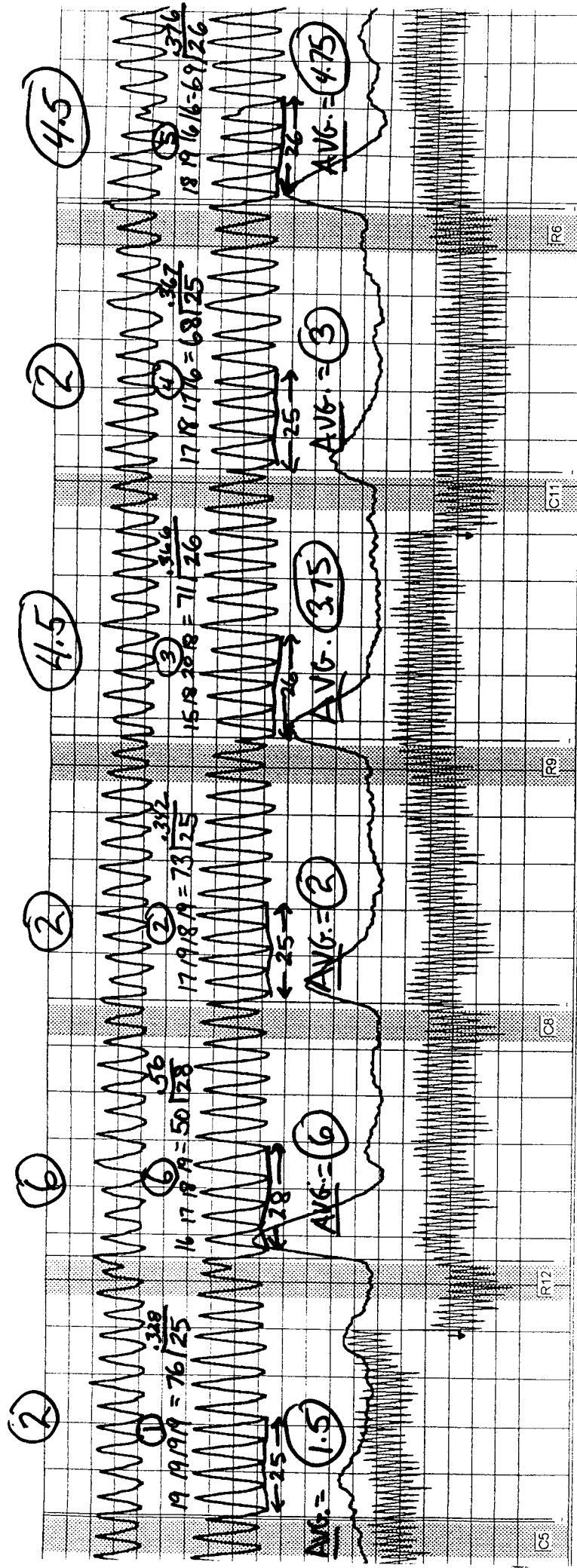


page 1 of 3 99-05-01 Exam 1
 Subject: CA
 Examiner: njg
 Date: 1/13/99 View Size Normal
 Time Start: 2:29:44 PM End: 2:34:09 PM Duration: 04:25
 Cuff Pressure Start: 68 End: 70
 98-05-01R (Zone)

page 2 of 3 99-05-01
 Subject: CA
 Examiner: njg
 Date: 1/13/99 View Size Normal
 Time Start: 2:29:44 PM End: 2:34:09 PM Duration: 04:25
 Cuff Pressure Start: 68 End: 70
 98-05-01R (Zone)

Gain Settings:	P2	P1	GS	CA
Recorded: Start	8.5	4.2	2.0	1.9
Recorded: End	8.5	4.2	2.0	1.9
Printed: Start	8.5	4.2	2.0	1.9
Printed: End	8.5	4.2	2.0	1.9

Figure 2



page 1 of 3 99-05-01
 Subject: [blank]
 Examiner: njg
 Date: 1/13/99 View Size: Normal
 Time Start: 2:29:44 PM End: 2:34:09 PM Duration: 04:25
 Cuff Pressure: Start 68 End 70
 98-05-01R (Zone)

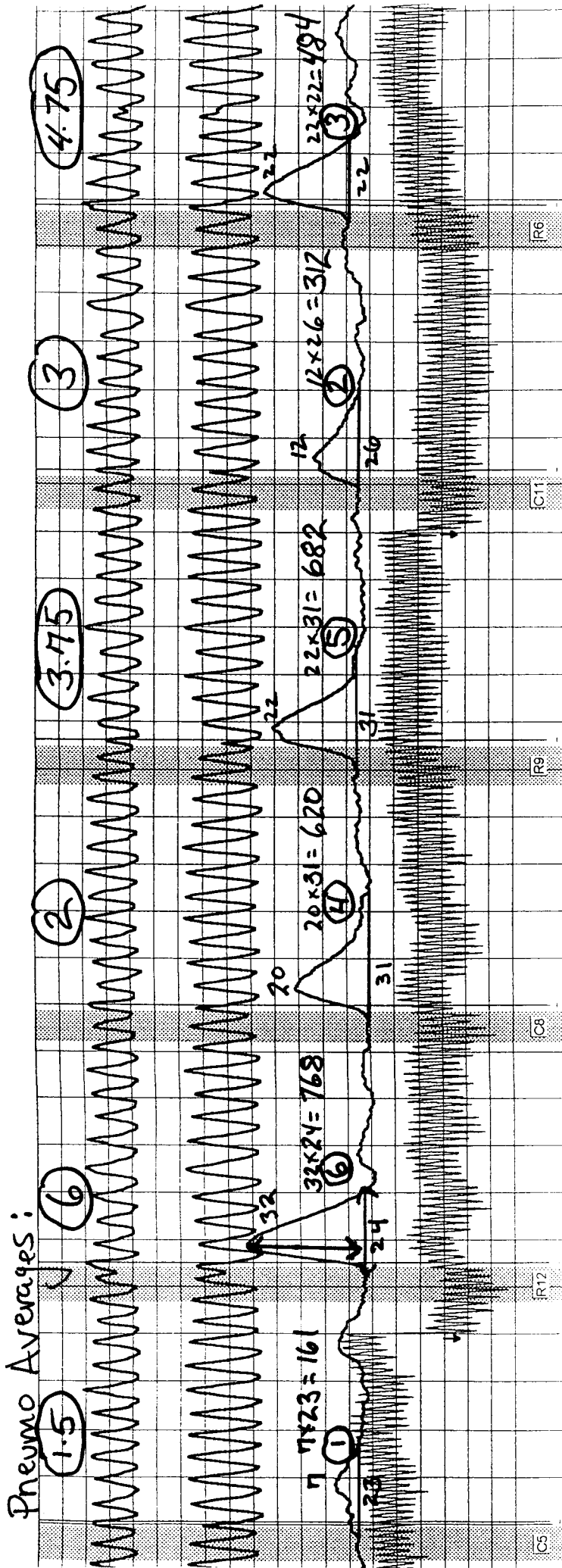
page 2 of 3 99-05-01
 Subject: [blank]
 Examiner: njg
 Date: 1/13/99 View Size: Normal
 Time Start: 2:29:44 PM End: 2:29:44 PM
 Cuff Pressure: Start 68 End 70
 98-05-01R (Zone)

CA	P1	GS	CA
19	8.5	4.2	19
19	8.5	4.2	19
19	8.5	4.2	19
19	8.5	4.2	19

Exam 1

Gain Settings:
 Recorded: Start 8.5 End 4.2
 Recorded: End 8.5 Start 4.2
 Printed: Start 8.5 End 4.2
 Printed: End 8.5 Start 4.2

Figure 3

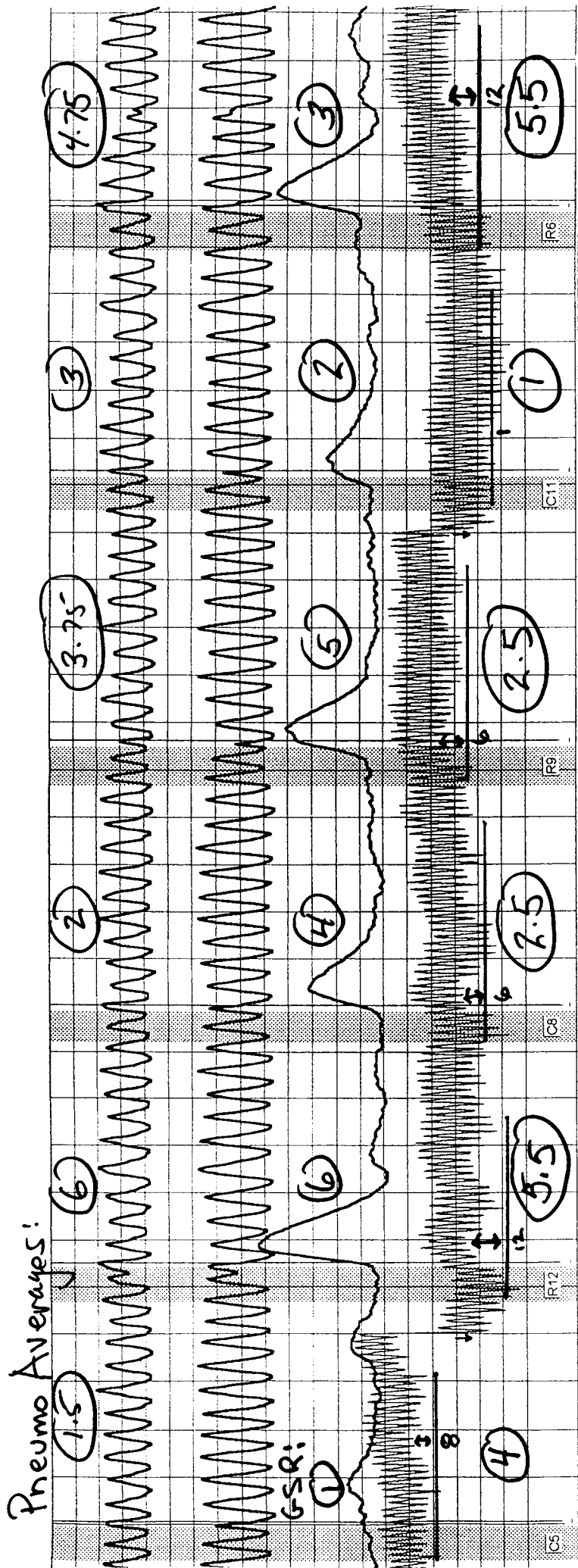


CA	P1	P2	Gain Settings:	Exam 1	CA
3	4.2	8.5	Recorded: Start	Exam 1	19
3	4.2	8.5	Recorded: End		19
3	4.2	8.5	Printed: Start		19
3	4.2	8.5	Printed: End		19

page 1 of 3 99-05-01
 Subject: Exam 1
 Examiner: nig
 Date: 1/13/99 View Size Normal
 Time Start: 2:29:44 PM End: 2:34:09 PM Duration: 04:25
 Cuff Pressure Start: 68 End: 70
 98-05-01R (Zone)

page 2 of 3 99-05-01
 Subject: Exam 1
 Examiner: nig
 Date: 1/13/99 View Size Normal
 Time Start: 2:29:44 PM End: 2:34:09 PM Duration: 04:25
 Cuff Pressure Start: 68 End: 70
 98-05-01R (Zone)

Figure 4



Pneumo Averages:

(1.5)

(6)

(2)

(3.75)

(3)

(4.75)

GSR: (1)

(6)

(4)

(5)

(2)

(3)

(4)

(5.5)

(2.5)

(2.5)

(1)

(5.5)

CA
2
3
3
3

TOTALS
+6.5

page 1 of 3 99-05-01
Subject: [redacted]
Examiner: njg
Date: 11/3/99 View Size Normal
Time Start: 2:29:44 PM End: 2:34:09 PM
Printed: 11/3/99
+8.5

Exam 1

Gain Settings:
Recorded: Start 8.5 End 4.2
Recorded: End 8.5
Printed: Start 8.5 End 4.2
Printed: End 8.5

GS CA
2.0 1.9
2.0 1.9
2.0 1.9
2.0 1.9

Exam 2

page 2 of 3 99-05-01
Subject: [redacted]
Examiner: njg
Date: 11/3/99 View Size Normal
Time Start: 2:29:44 PM End: 2:34:09 PM
Printed: 11/3/99
+6

TOTALS
-16.5
-11.25

Figure 5

Academy for Scientific Investigative Training's Horizontal Scoring System©

CHART 1

	Q# 5	Q# 12	Q# 8	Q# 9	Q# 11	Q# 6
P1	2	6	2	4.5	2	4.5
P2	1	6	2	3	4	5
Avg. Pneumo	1.5	6	2	3.75	3	4.75
E	1	6	4	5	2	3
C	4	5.5	2.5	2.5	1	5.5
Total	+6.5	-17.5	+8.5	-11.25	+6	-13.25

CHART 2

	Q#	Q#	Q#	Q#	Q#	Q#
P1						
P2						
Avg. Pneumo						
E						
C						
Total						

CHART 3

	Q#	Q#	Q#	Q#	Q#	Q#
P1						
P2						
Avg. Pneumo						
E						
C						
Total						

OVERALL DETERMINATION

Q#	CHART 1	CHART 2	CHART 3	TOTAL SCORE
5	+5			
12	-17.5			
8	+8.5			
9	-11.25			
11	+6			
6	-13.25			-22.5

Determination: _____