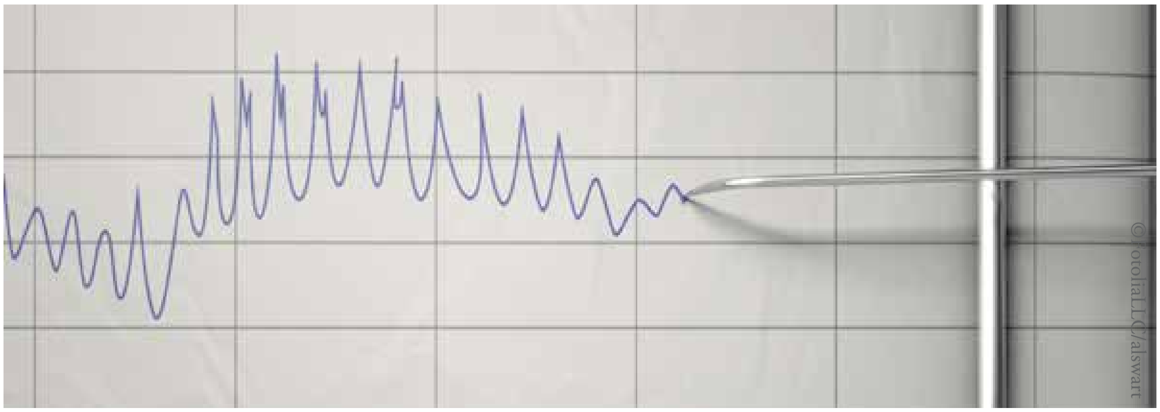
Practical Polygraph: A Recommendation for Combining the Upper and Lower Respiration Data for a Single Respiration Score



by Raymond Nelson and Donald Krapohl

Polygraph testing instruments make use of multiple physiological signals including the cardio and electrodermal activity, respiration activity, physical activity, and vasomotor activity (American Polygraph Association; APA, 2016). Polygraph test data analysis involves the assignment of numerical scores to the data for each of the recording sensors and iteration of the relevant target stimuli. Manual scoring protocols, based on visual inspection and analysis of the recorded data and some automated statistical scoring methods, have relied on nonparametric integer scores using a Likert-type coding system (Likert, 1932), referred to by polygraph field practitioners as seven-position scores, three-position scores and Empirical Scoring System (ESS) scores. An important aspect of numerical and statistical scoring of polygraphic data is that the respiration sensor consists of two transduc-

1 A transducer is a device to convert energy from one form to another. The respiration transducers transform respiration activity into changes into a time series signal that can be digitized and analyzed numerically.

ers[1] and associated recording channels, for the thoracic and abdominal areas, that are evaluated together. In practical terms this means a single respiration score is derived from the two respiration transducers.

A number of publications have described the rubrics and protocols for manual analysis of polygraph data (Bell, Raskin, Honts & Kircher, 1999; Department of Defense, 2006a, 2006b; Dutton, 2000; Krapohl, 2002; Krapohl & McManus, 1999; Handler & Nelson, 2008; Krapohl & Shaw, 2015; Nelson et al., 2011; Raskin, Honts & Kircher, 2014, Weaver, 1980). Krapohl and Russell (2014) described the procedure for dealing with differences in upper and lower pneumograph scores. The present paper expands upon the Krapohl and Russell (2014) description, providing a rationale and procedure for handling scores when a relevant question can be scored against either of two comparison questions, along with a score sheet template that can more easily accommodate the scores arising from this procedure.

Numerical Scoresheet Arrays

A two-dimensional score sheet array can be arranged for each iteration of the sequence of test questions, such that each relevant question is entered as a case row with different sensor scores entered as columnar values (test question x sensor). Table 1 shows an example. Each iteration or repetition of the sequence of test questions is traditionally referred to as a chart, referring to the erstwhile practice of recording polygraph sensor data as a series of tracings or plots onto a time-series paper scroll. Table 2 shows that the cells of the score-sheet array can be rearranged so that the sensor scores are oriented as case rows with the series of relevant questions presented as columnar values (sensor x test question).

Table 1. Score sheet array for a single chart – questions as case rows.

	Respiration	Electrodermal	Cardio	Vasomotor
R1	0	2	0	0
R2	0	0	1	1
R3	-1	2	1	0
R4	0	-2	1	1

Table 2. Score sheet array for a single chart – sensors as case rows.				
	R1	R2	R3	R4
Respiration	0	0	-1	0
Electrodermal	2	0	2	-1
Cardio	0	1	1	1
Vasomotor	0	1	0	1

Table 3 shows a score sheet for three completed repetitions of a question sequence that includes four relevant questions. Taken together the dataset of scores for all iterations or repetitions (charts) of the question sequence are mathematically a three-dimensional array (questions x sensors x repetitions). However, manual scoring protocols have traditionally relied on the simple addition of positive and negative scores for the sensor array, and for this reason it is more convenient for scorers to represent the three-dimensional score-sheet in two dimensions (chart/test question x sensor) as seen in Table 3.

Table 3. Score sheet array for a single chart – sensors as case rows.				
Chart 1	R1	R2	R3	R4
Respiration	0	0	-1	0
Electrodermal	2	0	2	-2
Cardio	0	1	1	1
Vasomotor	0	1	0	1
Chart 2				
Respiration	0	0	-1	0
Electrodermal	2	2	2	2
Cardio	1	1	1	-1
Vasomotor	0	0	0	1
Chart 3				
Respiration	0	1	0	-1
Electrodermal	2	0	0	0
Cardio	0	0	0	1
Vasomotor	0	1	0	1
Chart 4				
Respiration				
Electrodermal				
Cardio				
Vasomotor				
Chart 5				
Respiration				
Electrodermal				
Cardio				
Vasomotor				
Subtotals	7	7	4	3
Grand total	21			

When the three-dimensional score-sheet array is rearranged in two-dimensions, the columns can be easily summed to derive the relevant question subtotals[2]. Question subtotals can then be summed to achieve a grand total score. Neither subtotals for individual charts nor the subtotals for individual sensors are used in any standardized way when manually scoring polygraph data. In the score-sheet arrays shown in Tables 1, 2 and 3 it can be seen that each of the physiological sensors receives a single numerical score. For the Electrodermal, Cardio and Vasomotor sensors it is a simple matter to derive a single numerical score for each sensor for iteration of each relevant question. A single numerical score is also assigned to the respiration sensor, even though the respiration sensor consists of two transducers.

Some field polygraph examiners may use a variation of the score-sheet array that includes separate data-entry cells for the scores of the thoracic and abdominal sensors. A potential problem arises because the column sums will include two respiration scores for each presentation of each relevant question instead of one. The potential problem is that respiration data can become overweighted or overemphasized, relative to the other sensor scores, in ways that are not accounted for in published normative data or polygraph validity studies. Fortunately, this is easily remedied by combining the thoracic and abdominal transducer scores into a single score for the respiration sensor, and by including only the respiration sensor score in the relevant question subtotals. As can be seen in Table 4, it can be helpful to include in the score sheet some additional data-entry cells for the combined respiration sensor score. In this way only the combined respiration sensor score is included when summing the relevant question subtotal scores.

2 This has at times been referred to as "vertical scoring" but the use of the term vertical is equivocal in this context because it refers not to the actual data or recorded physiology but to the organization of the score-sheet array. A score-sheet organized differently, as occurs in some agencies, would give the same mathematical result by summing the scores horizontally. For this reason the term question subtotal is preferred.

Table 4. Score sheet array for a single chart – sensors as case rows, including thoracic and abdominal scores

Chart 1	R1		R2		R3		R4	
Thoracic resp.	0	0	1	0	-1	-1	0	0
Abdominal resp.	0		-1		-1		0	
Electrodermal	2		0		2		-2	
Cardio	0		1		1		1	
Vasomotor	0		1		0		1	
Chart 2								
Thoracic resp.	0	0	1	0	-1	-1	0	0
Abdominal resp.	0		-1		0		0	
Electrodermal	2		2		2		2	
Cardio	1		1		1		-1	
Vasomotor	0		0		0		1	
Chart 3								
Thoracic resp.	-1	0	1	1	0	0	-1	-1
Abdominal resp.	1		1		0		-1	
Electrodermal	2		0		0		0	
Cardio	0		0		0		1	
Vasomotor	0		1		0		1	
Chart 4								
Thoracic resp.								
Abdominal resp.								
Electrodermal								
Cardio								
Vasomotor								
Chart 5								
Thoracic resp.								
Abdominal resp.								
Electrodermal								
Cardio								
Vasomotor								
Subtotals	7		7		4		3	
Grand total	21							

How to select the comparison question for each relevant question

Before the thoracic and abdominal transducer scores can be combined, and before the sensor scores can be summed to obtain the subtotal and grand total scores, it is first necessary to obtain the numerical scores. This must be done using the correct procedure for the selection of relevant and comparison question pairs. Selection of which comparison question to use for evaluation with each relevant question is determined by the polygraph test format (i.e., the design of the sequence and content of test questions) and the intended method for test data analysis. In other words, different test formats and different analysis protocols can sometimes employ different ideas for the selection of which comparison question will be assessed against each relevant question.

Polygraph field practitioners and polygraph trainees have traditionally memorized the various rules for selecting relevant and comparison question pairs for analysis. However, Nelson (2017) described a set of heuristic principles that can be used for manual or automated selection of question pairs for the complete variety of polygraph test formats. To simplify this matter, the selection of which comparison question to evaluate with each relevant question will rely on one of two basic approaches: 1) select the comparison question immediately preceding the relevant question, or 2) choose from the nearest two comparison questions preceding and subsequent to the relevant questions and select the comparison question with the greater phasic response.

Some polygraph formats are structured such that the selection of relevant and comparison question pairing is unambiguous, usually relying on the first of these two basic approaches, with the relevant question normally paired and evaluated with the preceding comparison question. Other test formats make use of the second method, and require the examiner to choose from the comparison questions that precede and follow the relevant question.

As a matter of convention, when choosing between two comparison questions, responses at the relevant question are evaluated against those at the comparison question that produced the greater change in physiological activity. Tables 3 and 4 are virtually identical except that Table 4 has a cell for the combined pneumo score. The important procedural difference is that thoracic and abdominal scores are combined *prior* to entering the data onto the score sheet when using

Table 3, and are combined *after* entering the transducer scores onto the score sheet when using Table 4.

Herein exists a potential area of inconsistency; if, for example, a program or field practitioner were to require that responses to relevant questions, for thoracic and abdominal transducers, be evaluated with the same comparison question. In practice this would require that field examiners evaluate each response to each relevant question to both comparison questions for each respiration transducer and then combine the transducer scores for both comparison questions before selecting the comparison question and score that will be entered onto the score sheet. In addition to increasing the potential for undocumented analytic activity, this could result in a reduction of numerical scores that contribute to truth-telling because it would prevent the field practitioner from selecting the transducer score from the comparison question that produced the greater change in physiological activity. This would be an unfortunate form of bias to impose, and for this reason common practice among polygraph field practitioners and polygraph programs has been to obtain the numerical scores for each recording transducer by selecting the comparison question that produced the greater change in physiological activity – even if this means that the thoracic and abdominal transducer scores are obtained via different comparison questions.

Done correctly, when using a polygraph test format for when responses to relevant stimuli questions are evaluated with two comparison questions, scores that contribute to conclusions of truth-telling can occur if either of two comparison questions produce a greater change in physiology than the relevant stimuli questions. At the same time, scores that contribute to conclusions of deception will occur only if the response to the relevant question produces a greater change in physiological activity than both comparison stimuli.

How to combine the thoracic and abdominal respiration scores

Although summing the thoracic and abdominal respiration score is incorrect, aggregation of the numerical values for two transducer scores into a single value is actually a simple and straightforward matter using the sign values of the numerical scores. The numerical value of zero is neither + nor – and can be thought of as signless or unsigned (sometimes referred to as 0 signed). A simple rubric can be applied to seven-position, three-position and ESS scores.

1. If the sign values of the thoracic and abdominal sensors are opposite (+ and –) then the combined score is 0.

2. If the sign values are not opposite (including, + +, + 0, – –, or – 0) then the combined score is taken from the transducer score with the greater absolute value[3]. In other words, we use the score that is further from zero.

This simple rubric can be applied similarly to seven-position scores, three-position scores and ESS scores. Table 5 provides a number of examples using seven-position scores. Table 6 provides examples with three-position scores that can also be thought of as ESS scores.

Table 5. Examples (Ex) with seven-position scores.

	Ex 1	Ex 2	Ex 3	Ex 4	Ex 5	Ex 6	Ex 7	Ex 8	Ex 9
TR	1	-2	-1	2	1	2	-2	1	2
AR	-1	2	2	2	0	1	-2	-2	0
Score	0	0	0	2	1	2	-2	0	2

Table 6. Examples (Ex) with three-position scores (also ESS scores).

	Ex 1	Ex 2	Ex 3	Ex 4	Ex 5	Ex 6	Ex 7	Ex 8	Ex 9
TR	-1	0	1	0	1	1	-1	0	-1
AR	0	-1	-1	1	0	1	-1	0	1
Score	-1	-1	0	1	1	1	-1	0	0

3 An *absolute value* in mathematics is any positive or negative value taken as a positive value (i.e., without the sign). Absolute values are expressed mathematically by bracketing a number between vertical lines. So |-4| and |4| both represent the absolute value 4.

Summary

Scoring the respiration sensor is a process in which each presentation of each relevant question is assigned a single numerical score as a function of the evaluation of thoracic and abdominal respiration transducers. This process includes three important steps: 1) feature extraction – or identification of changes in physiological activity that are correlated with deception and truth-telling in the comparison question testing context, 2) selection of the comparison and relevant question pair and the assignment of a numerical score for each respiration transducer, and 3) reduction of the thoracic and abdominal transducer scores to a single respiration sensor score.

Scoring of polygraphic respiration data is potentially more ambiguous than scoring other sensors because the nature of respiration activity as more easily influenced by a combination of voluntary and involuntary activity. Scoring of polygraphic respiration data is also more ambiguous than scoring other sensor data because the respiration sensor consists of two transducers – one for the thoracic area and another for the abdominal area. Some variations of the score-sheet array include cells for both respiration and abdominal transducers; other score sheet arrangements may capture information only for the combined respiration sensor. There is no inherent basis to suggest that one score sheet is more correct or than another, though there are some advantages in terms of the reproducibility and accountability of the analysis when capturing more information.

As the polygraph profession moves towards increasingly structured and reproducible analytic models it may become increasingly important or useful to continue to increase the degree of standardization among polygraph programs and polygraph field practitioners. It is hoped that documentation of the procedures for combining the respiration transducer scores into a single respiration score will help serve the future needs of the polygraph and the needs of the agencies and communities served by the polygraph profession.

References

American Polygraph Association (2016). *Standards of Practice.* Approved by the Board of Directors on September 1, 2015. [Electronic version] Retrieved August 12, 2017, from http://www .polygraph.org.

Bell, B. G., Raskin, D. C., Honts, C. R. & Kircher, J.C. (1999). The Utah numerical scoring system. *Polygraph, 28(1)*, 1-9.

Dutton, D. (2000). Guide for performing the objective scoring system. *Polygraph, 29,* 177-184.

Handler, M. & Nelson, R. (2008). Utah approach to comparison question polygraph testing. *European Polygraph, 2,* 83-110.

Krapohl, D. J. (2002). Short report: Update for the objective scoring system. *Polygraph, 31,* 298-302.

Krapohl, D. & McManus, B. (1999). An objective method for manually scoring polygraph data. *Polygraph, 28*, 209-222.

Krapohl, D. & Russell, C. (2014). Initial investigation of selected hypotheses regarding the pneumograph in polygraph testing. *Polygraph, 43(3),* 59-70.

Likert, R. (1932). A technique for the measurement of attitudes. *Archives of Psychology, 140*, 5-55.

Nelson, R., Handler, M., Shaw, P., Gougler, M., Blalock, B., Russell, C., Cushman, B. & Oelrich, M. (2011). Using the Empirical Scoring System. *Polygraph, 40,* 67-78.

Raskin, D. C., Honts, C. R. & Kircher, J. (2014). *Credibility Assessment: Scientific Research and Applications.* Elsevier.

Weaver, R. S. (1980). The numerical evaluation of polygraph charts: Evolution and comparison of three major systems. *Polygraph, 9,* 94-108.