Practical Polgyraph: That One Weird Trick to Reduce Non-specific Physiological Activity (and Inconclusives)

Raymond Nelson [1]
Eduardo Lebrija [2]

Non-specific physiological activity

Non-specific physiological activity (NSPA, also referred to as non-specific activity, or non-specific physiology), refers to phasic changes in recorded time-series polygraph data that are not timely with the test stimuli. NSPAs are a concern to field polygraph examiners because it can interfere with feature extraction tasks and the assignment of numerical scores. Reduction of NSPAs may reduce the potential for inconclusive test results.

Tonic and phasic activity and polygraph theory

The analytic theory of the polygraph test is that greater changes in physiological activity are loaded at different types of test stimuli as a function of deception or truth-telling in response to relevant target stimuli. Changes in physiological activity can be thought of as either tonic (referring to baseline or long term activity) or phasic (short term changes in physiology in response to changes in the immediate

[1] Raymond Nelson is a psychotherapist, statistician and polygraph examiners, working as a research specialist with Lafayette Instrument Company. He is a past President of the APA, and an elected member of the board of directors. Mr. Nelson is a developer of the OSS-3 and ESS-M algorithms for polygraph test data analysis, and published numerous research and instructional papers on virtually every aspect of the polygraph test.

[2] Eduardo Lebrija is a polygraph examiner in Panama City, Panama, working for International Polygraph Training Center Panama. He has a background in engineering, and is fluent in Spanish and English, and is currently the elected President of the International Association of Polygraph Professionals (AIPP), a professional society for spanish-speaking polygraph examiners throughout the world.

environment). Tonic activity is not used when numerically scoring polygraph data. Phasic responses are of interest when evaluating polygraph time-series data.

Phasic activity can be extracted from the time series data using either visual methods (beginning during the pre-computer era) or via automated feature extraction algorithm. Automation provides the advantage of objectivity, structure and consistency, often relying on a carefully defined set of rules. Visual feature extraction methods also involve a defined set of rules, but may also involve more complex integration of different aspects of the recorded data, and for this reason can be more subjective than automated methods. Extracted features are transformed into numerical values and then aggregated or reduced for further analysis. Information can then be subject to bayesian analysis to render a probabilistic classification of deception or truth-telling. Some analysis methods provide only a statistical classifier (a statistical value that can support a categorical conclusion, though not intended to represent a pragmatic or observable probability).[3]Ideally, all phasic responses to polygraph test stimuli will occur in the context of a convenient and obvious tonic baseline.

Stimulus and response

An inherent complication to any scientific detection of deception test will be that all physiological activity is correlated with multiple human activities. In practical terms this means that all of the physiological changes of interest to the polygraph test may be induced by causes other than deception or truth-telling, and may occur during testing at times other than during the test stimuli[4]. The purpose of all data analysis is to quantify the likelihood that observed activity can be attributed to a particular hypothesis (i.e., deception or truth-telling). In the practical polygraph context, test data analysis serves to quantify the probabilistic strength of the information that is interpreted as indicative of deception or truth-telling, but does not detect or quantify detection *per se*. This does not imply that scientific polygraph results are not objective or are acceptably subjective. Instead, scientific results are ideally based on objective assumptions and reproducible statistical methods.

Any objective attribution or inference that observed changes in physiological activity have been caused by a test stimulus event will be premised on some basic and simple assumptions or requirements. First, observed responses should be time-

[3] Scientific tests are often used to quantify phenomena of interest that cannot be subject to direct physical measurement, and for this reason, scientific test results are often probabilistic. Because deception and truth are amorphous constructs, and are not, of themselves, a physical substance or physical phenomena, all forms of scientific credibility assessment testing are fundamentally probabilistic – relying on statistical correlations for proxy signals and statistical reference distributions (ideally subject to mathematical and logical proof) that have been found to be consistent with empirical data.

[4] Polygraph testing methods are designed to reduce the occurrence of phasic activity due to other causes, increase the likelihood that observed activity occurs in response to test stimuli, and provide data analysts with methods of observing and quantifying the likelihood that observed activity are a result of causes such as non-cooperation or faking.

ly with the test stimuli. That is, responses that begin before a stimulus cannot be reasonably and objectively attributed to the stimulus[5]. Changes in physiological activity that are not temporally associated with the test stimuli – those that begin in the absence of a test stimulus event, or are temporally not associated with the test stimuli – cannot be objectively attributed to the test stimuli because it cannot be known exactly what an examinee is thinking or feeling during test (mind-reading is still not possible). Also, it may be hazardous to attribute a response to a test stimulus if it begins long after a stimulus. Due to limitations and variation in human attention, responses that are not timely with the test stimuli may be induced by distraction or attention to topical material unrelated to the test stimuli.[6]

A second requirement is that the environmental context, in which a response or change in physiological activity is observed, should be absent of other observable events or activity during or immediately preceding a test stimulus event. When some unexpected activity or event is observed in the environment contemporaneous with a test stimulus event, it is objectively unknown whether an observed change in physiological activity is the result of the test stimuli or due to the unexpected event or activity.[7]

Ideally, all observed changes in physiological activity are located temporally with the test stimuli, in the absence of any unexpected contemporaneous event or activity that could provide an alternative causal explanation for the observed changes. Inclusion of such changes into the feature extraction and data analysis will inject unintended subjectivity into the analytic results. It is also ideal when there are no phasic changes in physiological activity that are independent of, or in-between, the test stimuli that are of greater response magnitude than responses to test stimuli. NSPAs that are of greater magnitude than responses to test stimuli may begin to cast inevitable doubt on the types of meanings that might be attributed to responses to test stimuli. In reality, it may not be uncommon to observe some NPSAs that exceed the magnitude of responses to test stimuli. Quite obviously, reliance on intuition, divination or guesswork to determine the cause of these changes will introduce subjective noise

[5] Some polygraph examiners sometimes refer to changes in physiological activity as "anticipated" if they occur in the threw seconds prior to a stimulus onset. However, the term anticipated is potentially problematic in that it may encourage an assumption that an examinee is thinking about the forthcoming test stimulus – not the cat or coffee pot at home. Such an assumption is not objectively possible with mind-reading capabilities, which would obviate the need for deception of detection testing.

[6] Determination of the procedural concepts and rules that define what does or does not constitute a timely response are an important topic in feature extraction research. In general, parameters and rules should be based on evidence, and supported by analytic results such as the optimization or maximization of parameter coefficients.

[7] Another, somewhat more nuanced, requirement is that physiological activity is expected to return to tonic levels in between the test stimuli, and that physiological activity in between test stimuli should be generally less than responses to the test stimuli. In other words, changes in physiology that occur in response to test stimulus events are expected to be generally greater than changes in physiology that occur independent from the test stimuli.

and potential bias into the analytic process. As with many nuanced problems, it remains a question for optimization research to provide evidence to support any practical guidance as to acceptable tolerances on this matter. Because scientific tests strive to be objective, with reproducible results, changes in physiological activity that cannot be objectively attribued to the test stimuli might be excluded from analysis. Reduction of NSPAs will be a useful strategy to limit or reduce the potential that changes in physiology are not due to test stimuli.

NSPAs should be excluded from data analysis

If NSPAs are not excluded from data analysis an expected consequence might be to increase the degree of subjectivity and subjective bias in the analytic result. This could contribute to decreased test reliability, with some corresponding increase in the likelihood of obtaining different results from different analysts. In practical terms, there is an expected decrease in test validity[8], because a test cannot be valid if it is not reliable, along with some potential increase in the likelihood of misclassification error.

An expected consequence if NSPAs are numerous within an examination may be that there is insufficient information to achieve a statistically significant result. When this occurs, field polygraph examiners may classify a test result as no opinion. Polygraph examiners use this classification to reduce decision errors. However, some others may regard inconclusive results as an error on the part of the examiner, and for this reason some field examiners may be reluctant to exclude any data from analysis, including when the cause of an observed change in physiology is ambiguous.

The occurrence of some small proportion of inconclusive results is a probabilistic and statistical reality (programs and examiners who report an inconclusive rate of zero should be viewed with caution). Nevertheless, reduction of inconclusive results is an ongoing topic of interest, and for this reason it is useful for field polygraph examiners to learn to identify, and reduce the occurrence of NSPAs to the extent possible. Systematic reduction of NSPAs will require an ability to identify and reduce the different potential causes for their occurrence. Related to this will be an ability to identify different types of NSPAs if such distinctions exist.

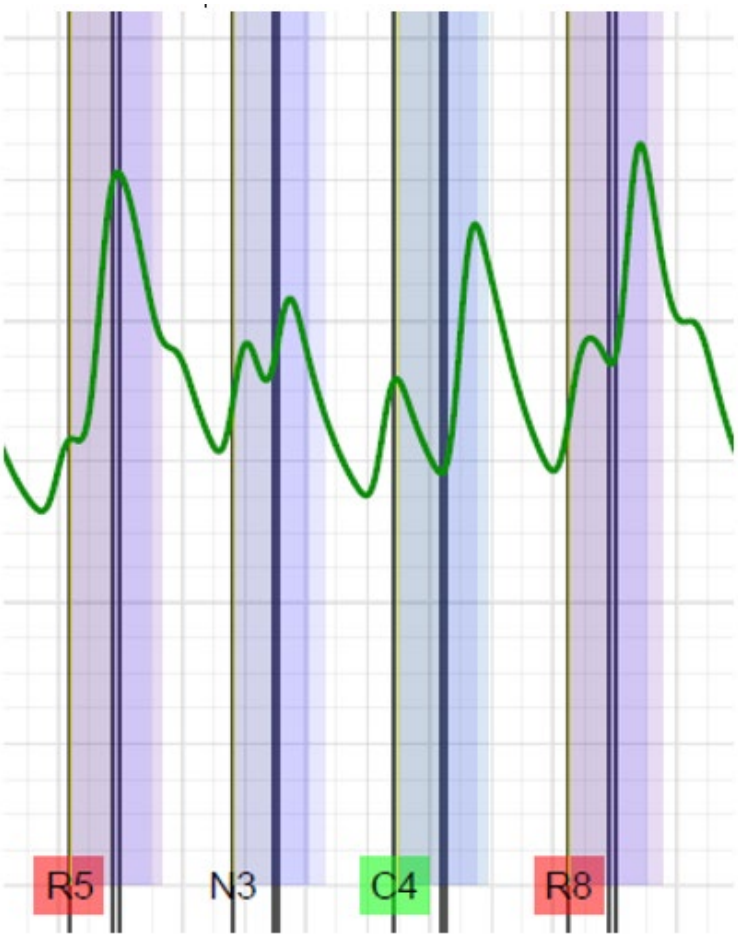
Two basic types of NSPAs

Some NSPAs can be observed in the form of phasic changes that begin during the few seconds preceding the onset of a test stimulus event. Figure 1 shows an example of several NPSAs that occur prior to stimulus onset. This example shows a sequence of four stimulus questions (R5, N3, C4, and R8) for which the onset of a positive slope segment occurs in the electrodermal (EDA) data during the few seconds prior to stimulus onset. It

[8] Scientific test results cannot be valid if they are not reliable and reproducible

is fortunate, in this example that there is an easily identifiable change to negative slope and subsequent positive slope onset during each of these stimulus events. Absence of a positive slope segment that begins following the onset of a stimulus event would result in no feature extraction and a potential increase in the likelihood of a test result that is not statistically significant.

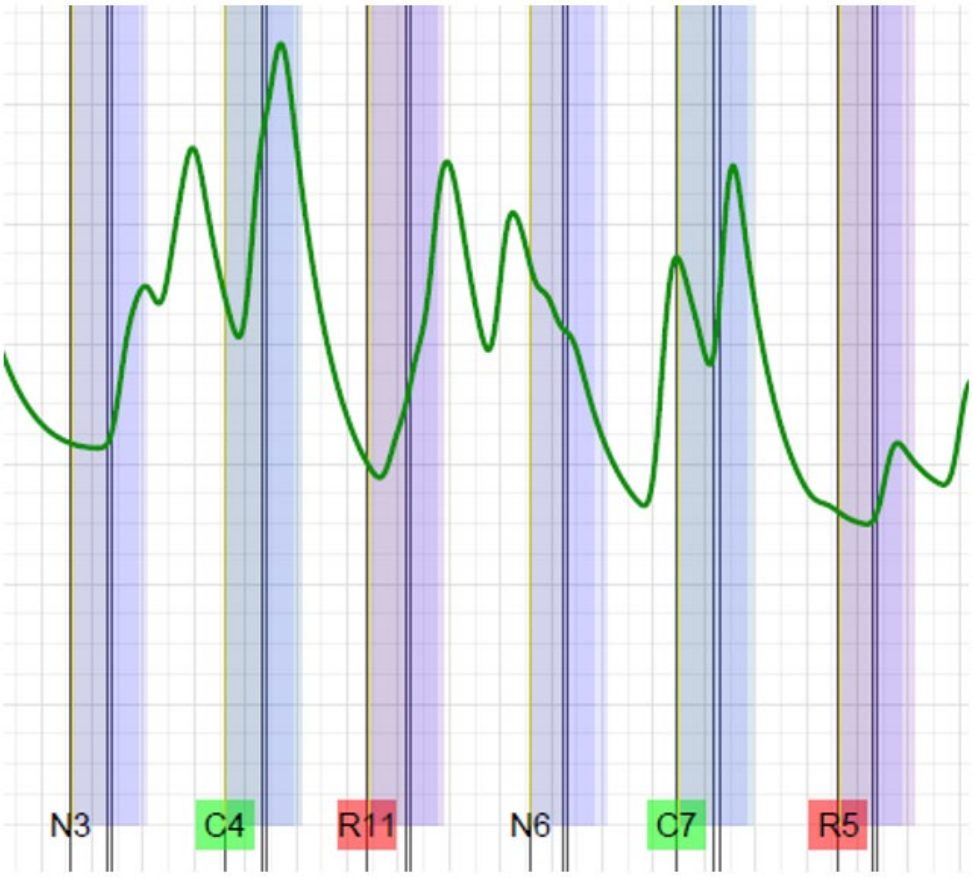
Figure 1. NPSAs prior to stimulus onset.



Other NSPAs can be observed in the form of phasic changes that occur in between the test questions. Figure 2 shows an example of this. In this example is can be seen that some NSPAs can occur in-between or independent of the test stimuli. Also shown in this example, some non-specific changes in physiology can be equal to or greater than the recorded and observed responses to the test stimuli. NSPAs of this type can often be ignored. However, when the NSPAs exhibit generally greater changes in physiological activity than responses to the test stimuli it may begin to cast doubt or uncertainty around the actual meaning or value of

physiological changes that occur in response to the test questions. The most important consideration here may be the frequency of occurrence. How often or how many times does it occur that the NSPAs are of greater magnitude than responses to test stimuli? The difficulty of interest is that without optimization studies it is likely that different examiners will have different subjective levels of tolerance for these events.

Figure 2. NSPAs in-between test stimuli.



Causes of NSPAs

NSPAs can be the result of a variety of possible causes. Not surprisingly, NSPAs become abundant and disruptive when an examination is attempted while others are present and observing during testing. Human interactions are often complex, and some people may often notice and react to a variety of subtle interactions between others and between themselves and others. A variety of observer effects are also possible, in which people may conduct themselves differently or perform differently when they are aware that they are being watched carefully by an

observer other than the examiner[9]. Limiting the testing context to the examinee and examiner is an important strategy for adequately controlling the amount of environmental stimulation and distraction. However, NSPAs can and do sometimes still occur as a result of other causes, including when others are not present in the room during testing. An obvious potential cause of some NSPAs can include any unexpected events in the environment, such as unexpected persons entering the examination room, or unexpected noises in or near the examination room.

Another possible cause for NSPAs is unexpected examinee behavior during testing. For example, an examinee may take a deep breath, or may engage in other physical activity. Unexpected events and unexpected behaviors are often observable, and can often be annotated in the recorded test data. Annotations provide information that is intended to be useful when analyzing and interpreting the test data and test result. However, some NSPAs can occur when the cause is not easily observable, leaving only the recorded physiological data to indicate that an examinee may be distracted or incorrectly engaged in the testing process.

When the cause is not observable it can be difficult to determine with complete certainty whether NSPAs are a result of deliberate and strategic covert behavior during testing, or are simply a result of random or uncontrolled causes. Statistical methods can be employed to calculate reproducible and theoretically sound estimates for the likely cause of these NSPAs. Whether annotated or not, and whether strategic or random, NSPAs remain an important consideration during test data analysis, and can sometimes result in the exclusion of a stimulus segment from data analysis.

Another possible cause that might be overseen as well is small noises coming from old or worn out keyboards and mice. Many examiners might be accustomed to their keyboard or mouse being noisy, as one becomes used to a non-well-oiled hinge. But that might not be the case for the examinee, who may find those peculiar noises to be distracting.

One other possible cause has been observed for some NSPAs – excessive movement on the part of the examiner, visible to the examinee, prior to stimulus onset. For example, an examinee my notice and begin to react when an examiner exhibits a behavioral pattern of gazing intently at the data on a computer display and only looking up at the examinee during the few seconds prior to the next stimulus question. Or, an examinee may notice that an examiner moves the position of their hands, towards the keyboard or computer mouse, in preparation for each subsequent stimulus event. In other words, some NSPAs may be inadvertently

[8] Scientific test results cannot be valid if they are not reliable and reproducible

caused by excessive movement of the examiner's head or hands during testing – when an examiner moves into position in the few seconds prior to the presentation of a stimulus question.

One weird thing to reduce NSPAs

Some NSPAs may be easily reduced or avoided – those caused by excessive examiner movement – by examiners who learn and practice good habits and good skills. Following is a list of suggestions.

• Position the examinee, whenever possible, so that they cannot easily observe the examiner without engaging in overtly disruptive physical movement. (Any position facing the examiner may be problematic.) A position directly opposite and transverse to the examiner, and slightly forward, so that the examiner is slightly behind the coronal plane of the examinee, is often suitable. The arrangement permits the examiner to easily observe the examinee's seated position and facial behavior from the side, while reducing the ability of the examinee to observe the examiner without physically turning their head. This can be thought of as the examinee default position (EDP). A variety of solutions may exist. The goal of the EDP is that it will be easily observable if an examinee breaks from the EDP to actively observe the examiner during testing.

• Examiners should determine their own optimal seated posture that will enable them to observe both the examinee and computer screen while moving only their eyes – refraining from physical movement, to the extent possible, of their overall posture and their head. This can be thought of as the testing default position (TDP). The goal of the TPD is that it will not be obvious or apparent to the examinee whether the examiner is watching the computer screen or the examinee (both should be easily visible from the TDP).

• Examiners should determine the optimal position for their hands during stimulus presentation. An optimal position will allow easy access to the necessary computer keys or mouse input device. For example, an examiner who uses keyboard input should identify the resting hand location that easily situates the fingers on the required keys. Fingers can then access these keys as needed without searching and without delay. Examiners who use mouse input can position their hand on the mouse outside of the examinee's range of view. Examiner hand position should be thought of as part of the TDP. Examiners should strive to conduct the entire examination from the TDP to the extent possible.

• If it necessary to break from the TDP (such as when adjusting the data or moving the computer mouse), the examiner should attempt to return to TDP early, long before it is time for the next stimulus event. Returning to TDP late, just before stimulus presentation, may increase the like-

lihood of computer input error, and may increase the likelihood that an examinee reacts to examiner movement in a way that interferes with recorded responses to test stimuli. By identifying and using a well-developed TDP, and returning early to the TDP after any necessary movement, an examiner NSPAs that might result from excessive and unnecessary movement of their head and hands during testing.

• And last but not least, get newer keyboards or mice if you need to.

Conclusion

Polygraph data is ideally recorded with a physically healthy examinee who remains completely cooperative throughout the interviewing and data acquisition phases of testing. Under optimal conditions, changes in physiology are easily observable in response to test stimulus events, with obvious return to tonic levels in between the test stimuli, and with no distinct changes in physiological activity not associated with the test questions. In reality, NSPAs can be observed in many polygraph examinations, regardless of whether examinees have been deceptive or truthful.

Prototypical NSPAs are non-timely changes in physiology, recorded in the time series data. Regardless of whether or not they are associated with any observable behavior or event, untimely changes in physiological activity cannot be objectively attributed to the test stimuli. Changes in physiology that occur in the absence of any unexpected behavior or unexpected event can be objectively attributed to a test stimulus event if they are timely with the stimulus[10].

The likely cause of NSPAs can often be observed by a watchful examiner. However, it can sometimes occur that NSPAs may occur when an examinee exhibits no observable unexpected behavior, and when no unexpected event is observed in the testing environment. The term nonspecific is intended to convey only that NSPAs are not attributable to the test stimuli – with no required assumptions as to the actual cause. Regardless of whether the cause is observed, and known, or not, NSPAs are a potential inconvenience to data analysis. Understanding and reducing the occurrence of NSPAs is a useful area of training and skill development for field polygraph examiners.

NSPAs are observed as changes in physiological activity that can be qualitatively similar to responses to test stimuli, but

[10] Responses are considered timely if they begin after the stimulus onset or required latency period and before the end of a standardized response onset widow, commonly extending to a few seconds after the end of a stimulus event or verbal answer. Another, more general, solution to the determination of timeliness is that changes in physiological activity are extracted from a fixed time window that is indexed to the stimulus in a standardized way.

with the important distinction that they are not timely with the stimulus events. NSPAs begin within a few seconds prior to a test stimulus, or completely independent of the test stimuli. It can sometimes be difficult to fully differentiate NSPAs from normal responses to test stimuli – such as when a response begins prior to stimulus onset and continues after stimulus onset. Some NSPAs may interfere with the production of a normal physiological response because the physiological systems being already engaged in the NSPA. Most importantly, the occurrence of unexpected behaviors or events during or immediately prior to a test stimuli, or the occurrence of substantial changes in physiology in the absence of any observable cause, can make it difficult to make objective or precise attributions about the meaning or value of observed responses to test stimuli.

An important consideration with NSPAs is that, although they are sometimes characterized as random, they are, in reality, the result of causal factors – regardless of whether or not those causes are easily observed or identified. It is also possible that NSPAs may be indicative of strategic faking (but that is a topic for a different paper). Statistical methods are the preferred solution to formulate reproducible conclusions about the likely cause of observed NSPAs when their cause is not readily observable. When we can statistically align the occurrence of NSPAs as loaded mainly onto certain test stimuli it may become a basis for a conclusion that an examinee has engaged in strategic faking activities. In contrast, NSPAs that are distributed among the test questions in a manner consistent with random activity are less likely to systematically influence a test result, and can often be regarded as a minor inconvenience. Importantly, when we can align NSPAs with inadvertent activity or stimulation from the examiner we can conclude that some improvement in examiner skill and/or self-awareness may lead to an improvement in the effectiveness of testing.

Developing an awareness of the correct EDP and TDP will enable examiners to adapt quickly and easily to a variety of testing environments while pursuing the goal of engaging satisfactory behavior from the examinee, and while reducing their own contribution to NSPAs. Having a template or default position will contribute to expert-level skill development, including the early return to TDP, and the reduction of unintended stimulation prior to stimulus onset. Additionally, an improved ability to understand and manage NSPAs may contribute to improved ability to identify NSPAs that occur due to faking. It is also important to note that some NSPAs can occur as a result of neither systematic faking nor inadvertent stimulation.

Changes in physiology can be reasonably and objectively attributed to the test stimuli when they are timely with the test stimuli and when no unexpected event or activity is observed. These changes can be subject to feature extraction, aggregation and data analysis to render a probabilistic classification of deception or truth-telling. When some unexpected activity or event is observed contemporaneous with a test stimulus event, it is objectively

unknown whether an observed change in physiological activity is the result of the test stimuli or due to the unexpected event or behavior.

Recorded data is always a combination of diagnostic variation and non-diagnostic variation (often characterized as random variation or noise), and some loss of test data can be reasonably expected for many, if not most, polygraph examinations. For this reason, polygraph test data are recorded using multiple sensors, multiple relevant/target questions, and multiple repetitions of the array of test questions. Most often, despite the loss of some information, there remains a satisfactory quantity of usable or interpretable response segments so that data analysis can be completed in a standardized manner. However, if the occurrence of NSPAs becomes too numerous they may begin to cause difficulty with an ability to obtain a satisfactory quantity of usable information. The most likely effect of excessive NSPAs may be an increase in inconclusive outcomes. If NSPAs are misinterpreted as responses to test stimuli, they may contribute to misclassification errors. For this reason, NSPAs represent a potential complication to the analysis, interpretation, and effectiveness of polygraph test data.

An objective attribution or inference that an observed change in physiological activity is caused by a test stimulus event will require that some basic conditions are satisfied. First, the change in physiology must be timely with the stimulus event. A second requirement is that there be an absence of any other observable behavior or event during or immediately preceding a response to a test stimulus event. A systematic and strategic approach to managing the environment and testing context may help to reduce unexpected or undesired chaos during testing and increase the standardization and reliability of the polygraph, thereby advancing the goals of the polygraph profession. Reduction of unnecessary examiner movement, noise, and unintended stimulation, is one weird trick that may help to reduce NSPAs and their potential contribution to misclassification errors and inconclusive testing outcomes.