Introduction

This paper discusses the systematic approach and method of carrying out experimental research, albeit on a simple and small scale. This is less troublesome than it may seem, but requires attention to detail, consistent thoroughness and discipline.

Is that of important concern? Yes, because 'experimenting' has more meanings than it should have. Carrying out real experiments is a serious and honest task, while mixing chemicals and fooling around has nothing to do with serious endeavour. I have to admit that in the past because of lack of time, reliable apparatus and chemical resources -I also 'sinned' in this regard. This can be a waste of time, money and materials, possibly with bad results and perhaps harmful consequences.

The NSTW-method

Every researcher, as well as the serious amateur, who does any work in the fields of natural sciences and technology, follows certain rational steps during his research. This course of thoughts and actions, in a systematical approach, is what is called *The Natural Scientific Thinking and Working Procedure*.

Every research starts with a certain observation, from which is derived a defining of a problem, which is a well-considered question as a result of the perceived phenomenon in the everyday life, or in experiment. By trying to find an answer to the problem in a systematical way of guessing, a theoretically sound answer is found, the hypothesis.

This is a normal course of action in daily life. Everybody asks questions and these questions are more or less answered. Whether the answer is actually true, often can or will remain unknown. This may happen many times. One asks questions and one guesses for the right answer, but does not go beyond that.

In natural sciences it is very usual to advance to the next phase, the experiment, which is a phase so much neglected in many other sciences, or that one leaves out, for ease, in daily life.

The experiment is the ultimate test of the hypothesis! From the result of the experiment one will find out whether or not the hypothesis was defined correctly, and if the question was fully answered. Sometimes one can find out that the question was incorrect and has to be rephrased. A sufficient number of experiments must be carried out to obviate coincidences, errors or other false factors. Furthermore a separate test series should be carried out in order to establish a reference or datum by which any new set of results can be compared.

For example, consider the baking of bread. One could argue that home made bread is often sticky and heavy. Question: How can
one bake bread that is lighter? In reaction, after observing some details, the answer may be that the bread dough should rise longer. In this case experimentation will provide the answer. If experimentation reveals a relationship between several hypotheses, a theory can be evolved. This is an important final phase in the course of performing experiments and research. As regards the baking of bread, several factors might influence the lightness of the bread: perhaps the (type of) yeast, the baking temperature or the (type of) bran contents of flour might be factors. Experiments may reveal the answers and an overall theory may then be evolved.

The systematic approach

As has been stated, the research begins with a 'simple' observation. It is very important that this observation can be repeated or reproduced, and that the observation is done as objectively as possible, in order to prevent any errors in perception. Therefore, it is desirable that several independent researchers make the observations.

Regrettably, most often, things already go wrong in this beginning phase of research! If an experiment has to be reproducible for someone else, as much clear and important specific detail should be provided. This is a point that is lacking in many common studies. (Remember the first days of the "cold nuclear fusion" some years ago the whole project turned out to be phoney boloney.)

The problems that occur with objective judgment are often caused by prejudice or errors in human perception. One tends to see what one wants to see, or that which one thinks one should see. With judging the quality of pyrotechnical coloured light -besides atmospheric influences and distance to the source -the perceptibility and sensitivity of the human eye and nerve systems are of importance. (Not to mention the influence of experience, habituation and personal taste!) Not everyone can distinguish colours 'as well' as some other people do, so not everybody is 'as objective' in this respect! (Question: What is objective -are there standards?) A pitiful limitation of the human eye and nerve physiology.

After this, the problem is defined. Defining a good unequivocal question is as difficult as observing objectively. To be of scientific value, a question has to be appropriate, and has to be able to be tested. Questions that must lead to the usefulness or sense of something are difficult, if not impossible to answer and are therefore beyond the scope of natural sciences. Why are bananas curved? Because otherwise they would not fit in their skin. Nonsense! Of importance are questions such as 'how', and 'what is it caused by'.

Consequently, as an answer to the defined problem, there is the hypothesis. It takes knowledge, experience, insight, feeling, and creativity to find a supposedly useful solution to the problem, which must be able to be tested by verifying and falsifying. Of course one can answer each question with many answers, but the researcher should strive for a simple solution and to reduce that to only two mutually exclusive possibilities: yes or no! Often this can not be done, and the non-equivocal answer will be: it is possible.

The hypothesis leads to the experiment: empirically we will see whether the answer to the question is correct or not. Only the experiment can prove the hypothesis. Performing good experiments is not easy. Experiments can not always guarantee, or lead to, the right conclusion. Also, in this phase, there are numerous occasions on which one can perform mistakes (or even intentionally commit fraud: such as the AIDS research conducted by professor Buck from Eindhoven, Netherlands, who 'levered up' his experimental data to suit his hypotheses). Often a great insight and ingenious resourcefulness are required to design a good set of experiments.
Of utmost importance for the reliability and value of the experiments used to test or prove the hypothesis, is that parallel to the experiment an equal test series is developed which has only one difference, the variable, compared to the previous.

This zero level datum or 'control' experiment is also called a 'blank'. The purpose of the 'blank' is to create 'a standard' that can be used to put the results into perspective. Furthermore, all details and special circumstances of the experiment must be described extensively.

In other words, in a test to create different flame colours, a basic mix with fixed percentages of potassium perchlorate, chlorine donor and hexamine could form a standard per batch, and the variable would form the addition of a fixed percentage of flame colouring agent, for each sample, per batch. So, the standard or blank(s) here will have no addition at all, the variables will be the different kinds of colouring agent.

If for a certain experiment a good colour is produced, say for blue, and one wants to do further experimentation with that blue colouring agent alone, the next step is varying the percentages -the new variable. The standard or blank here will be the one, per batch, with a fixed percentage of that blue colouring agent. As well as for the blank as for the variables, the percentages for the other compounds stay fixed, of course. Furthermore, it is important to take blanks from each batch (when more batches are needed), because making another batch each time may introduce unwanted variables (errors in weighing, spills, etc.) and so one can also compare blanks per batch.

One important aspect concerning experimentation must be noted: because the work must be done and described under precise and constant conditions and circumstances, and frequently a lot of data has to be processed and validated, it is very laborious work which requires much patience, honesty, some good routine, and discipline. Not every hobbyist can provide this harsh attitude. Besides, one needs a well-equipped laboratory with the right and well-calibrated equipment. Not every hobbyist has access to this. However, with a good dedicated mind and thrift one will make these sacrifices in order to perform a serious hobby.

Finally a theory is formed. One can speak of a theory if relationships can be made between several hypotheses which have been proven by several experiments. Often, one will notice that through the retrieved data from the experiments, new hypotheses must be defined -which will require new experiments. (This will keep one busy for a while and off the streets, or is per definition the street the poor mans’ laboratory?) A theory never has any absolute value, but unlike non-natural scientific predictions (like paragnosti c predictions) natural scientific predictions have a high degree of probability. Sometimes the acknowledged hypothesis is of such high probability that one can call it a natural law. In most cases the results of an experiment will force the researcher to review his hypothesis critically, and the shape in which these experiments have been performed. They can compel him to make adjustments or advance research. A theory will always be an approximation of real life truth, and therefore the real life truth is schematised through models.

**An uncomplicated example**

Unintentionally, I may have discouraged people to seriously consider shaping up their experimental stimulus in the direction of genuine research. This is certainly not my intention, because it really can be done quite simply.

**Observations.** Sometimes chemical reactions take place at a low speed, like the corrosion of an iron nail that can take months, or even years. However, more often oxidation reac-
tions occur with fire, and sometimes even loud explosions will happen in a fraction of a second. The speed at which a reaction proceeds is apparently not always the same and chemical reaction rates may appear to vary enormously.

Definition of the problem. It all appears to be quite normal for the uninterested and uninhibited, who just accept these natural occurrences. However, the serious amateur pyrotechnician will ask himself - what causes these differences and how can he adjust the speed of certain reactions for certain effects. In other words, what causes the differences in reaction rate?

Hypothesis. An explanation or an answer to these questions could be: the reaction rate depends on the level of equal distribution (that is, the particle size of reactants in a certain space).

Experiment. If one wants to find out whether or not the reaction rate is dependent of the level of equal distribution, one can conduct a series of experiments to test that. One will need reliable equipment, like a good timing device, balance, marking gauge (vernier callipers), and a good set of standardized quality sieves. In this experiment one can use the reactive metal magnesium, whereby one uses equal mass quantities in unequal level of distribution (different particle size in a certain fixed space, which will be the variable), ranging from coarse little chunks to fine powder (of the same source!), in a combustion test. For the mixing, a standardized quantity (and quality) of oxidizer, for example potassium nitrate (of constant particle size, say 100 mesh), will be added in stoichiometric ratio. The set up for the test, and the circumstances do have to be the same at each individual test so that only the level of distribution (of one chemical or of the complete mix) varies and all other factors (of importance) remain completely constant. The particle size does have to be carefully determined, and the time can be measured by using a stopwatch. All data have to be carefully and clearly written down in a table. The combustion can be performed outside at atmospheric conditions (noting there is no rain, no significant wind, and also noting temperature, ambient pressure and relative humidity), by piling up the mixture and starting the fire in exactly the same manner each time. To compensate for relatively small errors and variations, or to eliminate huge differences caused by mistakes or other influences, all measurements must be performed twice, or even better three times. The results can later, if possible, be averaged - if the range between the values of deviation permits this.

Theory. The hypothesis will be confirmed by the fact that the chemical reaction rate does indeed depend on the level of distribution. The more intimate the level of distribution (the finer the particle size in a certain defined space), the faster the reaction will proceed. The increase of the reaction rate can be explained on a molecular level through the collision model. The experiments can be expanded by trying to find out whether or not this theory is also valid for other chemicals (or the other chemicals in the same mixture). One can also search, through means of experiments, for other factors that influence the reaction rate, such as temperature and pressure, density, concentration, catalysts, purity, coatings, etc. The experiments will become much more complex.

When all these results are combined or related, one will be lead to a generally valid theory of chemistry. The chemical reaction rate is influenced by a number of factors, such as: the nature and properties of the chemicals, the ratios in which their quantities are mixed, the level of distribution, the amount of energy released, the presence of catalysts, the level of confinement and compression, the presence of inert compounds, the thermal conductivity of the mixture, the cylinder (if present) in which the mixture is contained, the velocity and movement of the
burning mass, the atmospheric pressure, etc.

Two practical cases

It can be very educational to see how some people conduct experiments in practice. Closely followed is a study of whistles (because it caught the eye first), by Miss S. Partin; the second article, also concerning whistles by Mr. J. Toker, will be glanced upon. The subject is interesting.

Fussing over trifles is not the intention, but as a matter of fact it will be shown that fundamental shortcomings may be unnoticed by uncritical or uninstructed readers. These studies are nice approaches on a scholar level, but not (yet) mature enough for publication and sale, let alone being printed in a quality paper such as Pyrotechnica. The studies of Partin and Toker recently appeared in Pyrotechnica XVI, without appreciable change or revision by a technical board.

The Partin report was presented several months ago, in early February '95. The 16 page report is sold for $7 and authored by S. Partin, a female member of the PGI. She has devoted herself, as part of a science project at her school (of which the level is unknown, regretfully), to research catalysts in whistle compositions that influence the chemical reaction rate.

To introduce the theory for readers: by adding a catalyst to a whistle composition, the chemical reaction rate is influenced and a characteristic change of frequency and pulsation can occur. Positive catalysts give a high piercing sound, and for negative catalysts a change of sound can occur from high frequency to the low, pulsating, flatulating sound which reminds one of the farting sound produced by people who are specialized in, as it is called here, petomania (the gassing art of playing the national anthem after a meal of beans). It has been known for several years that similar slow, pulsating effects can be created by applying a nozzle to a normally nozzle-less whistle, causing a repression of the pulsation and thus creating a delay (from: private correspondence with L. Jackson, Hull, England, 1986). The nozzle probably represses and disturbs the inward and outward flow of gas streams, but as with adding catalysts, how the chemical reaction rate is being influenced and retarded so simply and unpunished for such potentially high energetic mixes, is still unexplained.

The interesting part of this research is, of course, the probability that some simple chemicals exist that will result in speeding up or delaying the combustion rate, significantly influencing the sound of some whistle compositions. The phenomenon of negative catalysis is reasonably new for whistle compositions. Up to this point Partin's research has to be encouraged, not in the least case because she obviously is one of the few active women in the pyrotechnic 'mans society' -which should be praised and stimulated. But particularly because these kinds of small scale studies are essentially relatively simple -of which a lot more could and should be performed and published by amateurs. The topic is very interesting, possibly for commercial applications.

The major objection, from this writer's point of view, to the study is that the research is absolutely non reproducible -which is a major shortcoming. The final results are likewise. I am not aware of the possible fact that Partin is holding back evidence deliberately because of commercial interests, so I am not proposing that option for real. However, there must be some explanation why the positive catalyst copper oxychloride is mentioned to be effective, while the mineral rutile is mentioned as being the most potential negative catalyst. The report does not make clear that the used rutile is NOT common titanium dioxide. The active material is rather 'tan brown rutile' -thus 'doped' by nature with a relatively high percentage of
iron (II) oxide, ranging from a few to 15%. Probably this 'dope' is of major importance in conjunction with the titanium dioxide. The characteristics of good research, even on a scholar level, are that everybody willing to do so should be able to reproduce the experiments, and obtain the same results. This will, regretfully, not be the case.

**Vibrational whistles: an unclear approach**

The report is titled "Vibrational Whistle Rockets, A Study of Catalyst Selection on Performance". All whistles are vibrational, but that is just a detail.

Two major questions came to Partin's mind: Does an increase in sound as decibels, decrease the energy in thrust? A good question. Furthermore: What impact does (the addition of) iron (III) oxide catalyst have on the composition? This is not a very good question, as we already know the positive influence -more important, 'impact' can not be measured.

As first priority a rocket propellant is searched for that produces the greatest thrust with the loudest whistle sound. She does not use several different oxidizers and fuels, but takes two basic compositions and varies the catalyst. Although the salicylate composition choice is a little odd in percentages, she uses an unexpected variable and from which one might expect surprising results. So, nothing wrong here. However, halfway in research she discovers that the rocket-fuels that are 'quested', quote: 'were not always as pleasing (sound) as other ...' to her. She obviously decides to find "the most pleasing sound" to her, but she forgets that this is absolutely not definable or measurable, and certainly not reproducible. The definition 'pleasing' is rather questionable: what is pleasing to one does not have to be pleasing for someone else. On the contrary, it could be exactly the opposite.

Proposed is that the combustion effect of whistles is similar to that of strobes, but this is not the case. One common feature is that both pyrotechnic mixtures bum intermittently.

Strobes bum by means of a dark and light phase which is a different chemical process more fully understood. Reactions whereby, if magnalium is used, first the (more reactive compound) magnesium and consequently the (flash producing) aluminium is being consumed out of the alloy (Shimizu). Some organic based strobes, however, have other characteristics.

Another difference is that whistle compositions pulse at much higher frequencies and that the tube, and inward and outward flow of gas, plays a very important role. Strobes do not need empty tubes or resonant cavities, nor is significant gas flow to be noticed. Furthermore, in a whistle there is a solid phase at the burning front (Maxwell), and strobes have a liquid phase surface (Shimizu).

**Knowing what is important**

Two decibel meters are mentioned, in the enumeration of conditions, circumstances and apparatus, but the type or manufacturer is not revealed. More important, the report does not specify exactly what the measurement conditions were and at what distance and angle the measurements were taken (with or without any reflecting surfaces nearby, hard or soft ground, wind gusts, etc.). These are all very important issues that have a big influence on the research of sound. A clear drawing of the situation fails to be presented. Also, an attempt was made to specify the quality of the sound, but regretfully without reliable apparatus this is not measurable nor reproducible. The whistle effects are described in terms that can be interpreted in various ways, all rather meaningless, like: disappointing (!), soft, pleasing (!), high pitched, piercing, met my
goals (!), sporadic sound frequency (!?), surprising (!), different than expected (!!) and finally I loved it! (!) As a result one can not use these unquantified terms.

Consequently, mention was made of preparing all the composition batches in the same manner, but my conclusion to that is that it is not sufficient, and is no guarantee that this has actually been the case. That errors were made is shown by the wrong sequence in preparing the mixtures - first weighing and then grinding and sieving. This will obviously result in unknown losses that will show up in the results.

Partin actually does not read written information, although a reference list of literature has been given. From experience (Maxwell, Ellern a/o) it is known that there are significant differences from batch to batch, whatever precaution one takes.

Maxwell is very clear in this regard (Maxwell, p. 907). Quote: 'It was found that, to obtain consistent results, measurements designed to investigate any particular effect had to be carried out on the same day under the same conditions; also if the effect of some physical factor such as tube diameter was being investigated for a particular composition then the composition had to be all from the same batch and preferably consolidated at the same time. Similarly the ingredients of the compositions containing varying proportions had to come from the same batch. Unless all these precautions were taken, inconsistencies that might amount to several hundred percent were likely to be experienced'. Unquote. This is the clear and straight language of a sincere scientist. A similar critical view on Partin's own work is not present and missed.

Mentioning the procedures to produce and handle the batches, it is explained that Kraft paper is used to dry the wet pyrotechnic composition on. But everybody knows that this leads to unknown losses of chemicals and flagmatisers. (This is why one should always use a non-absorbing underground, or if needed, by special procedure qualifying and quantifying the materials lost.) There is a list of the most bizarre catalysts, like uranium oxide (!), although it is stated that not all chemicals were used. But why then, this dangerous radioactive uranium oxide? What safety precautions were taken? Why the list of partly unused chemicals?

Some of the given chemical formulas are incorrect, like that of cobalt oxide, copper oxychloride, manganese dioxide, rutile (which, as stated earlier, is not the same as pure titanium dioxide), yellow ochre (which is more commonly based on aluminium oxide and iron oxihydroxide (FeO(OH)), not containing iron (III) oxide). Very predictable oxides that would have been very interesting to see being used in this research are not on the list of used and unused chemicals. When such chemicals are presented in research, as being catalysts, the quality and purity have to be specified, where the chemicals came from (product number), and in case all this data is not available, the source of the chemicals should be given. With the impure chemicals Partin probably used, it is not for sure which compound (or variety) is having catalyst capacity.

Note that potassium benzoate is also characterized here as having the formula KC_7H_5O_2. 3H_2O, as Maxwell and many others have stated. More recent literature (Whelan / Eli-scher) resolved this duality: this compound, as a commercially available chemical, contains no water of crystallization at all.

Another shortcoming in the procedures used to charge the projectiles are that there are no references to the pressures used. Was the same pressure used at all times, and monitored by quality measuring equipment? Quote: 'Small increments of the composition are added and pressed', unquote, but nowhere is it specified how often, how much, and at what pressure. This is quite careless.
During the testing stage the true character and value of the research are revealed: in each test of each catalyst four rocket engines are tested in a static way, measuring the thrust and decibels (and in later experiments the burning time also). The four measurements are averaged. But what is lacking in each test? The ever so necessary zero level 'blank' test, in which there is NO catalyst present. How is it possible to make a good comparison without any reference point, and how can one forget this check? Partin would have noticed significant differences between the blanks per batch. Without blanks one can never say for sure if the catalyst had a positive or negative effect on the combustion rate and other parameters.

Furthermore she averages the results of the four resulting sets of test data, but it would have been very interesting to see how much difference there was between these four sets. These values could be so large that averaging the data would be of no significant use because the differences present are too significant: the resulting averages would be of no value.

There is also a slight problem in producing significant figures: (salicylate rockets) thrust 1 (!) pound, duration 0.65 (!!) seconds, 105 (!!!) decibels for iron (II) oxide; thrust 2.17 (!!!) pounds, duration 3 (!) seconds, 113 (!!!) decibels for uranium oxide; thrust 1.33 (!!!) pounds, duration 1.82 (!!!) seconds, 115.33 (!!!!!) decibels for manganese oxide, to give a few.

There is a list of non relevant observations, such as a collection of different shades of coloured flame, like: apricot with yellow, red to brown, cold looking flame. Observations are stated which can not be backed up with facts, like: "a possible reason could be the high oxygen content (for the high decibel output using uranium oxide)". Burnt umber (ferro manganese oxide) would, because of the low oxygen content, give a low output.

If this is the case and some chemicals release extra oxygen, the added chemicals are likely not active as catalysts. However, the low percentage added can not be of that great importance for this release.

Although the reader learns that the research was conducted over a longer term period, it is suddenly stated: "weather was clear, dry and 74 degrees F". But how does such a sudden detailed description have to be considered, in comparison with all these other insecurities? In the case of the addition of cobalt oxide, the colour of the pyrotechnic composition granules changes after several days. Concluded: This is possibly why (the) performance was poor. But, is she referring to a chemical reaction (if so, which one(s)?) and why is the test not repeated to check this phenomenon?

Conclusion and missing retrospect

The conclusion at the end of the report states, quote: 'The right catalyst depends on the application....the choice is a matter of personal choice, since the whistle is different for each catalyst'. Unquote.

I am sorry to say it, but this conclusion is, of course, of no use at all. In such disappointing research one will always obtain totally different results, and this is not the result of the self chosen unknown factor (the catalyst) only, but more likely as a result of introducing unknown variables by making coincidental and systematic errors and through unscientific carelessness, from almost all viewpoints of study. This is research that can not be called research, regardless of the scholar level or how small the project was. Much missed is a review, in which Partin would follow her own procedures and final results, showing some criticism of her own labours. The total lack of this is quite distinctive.
There is something strange here: in the Netherlands there is NO school where students could perform such experiments with quite dangerous capabilities (if big whistles detonate, a lot of high energy is released), nor is there any school here where one could perform combustion experiments with radioactive substances such as uranium oxides! I think it is very unwise and irresponsible of teachers to let inexperienced youngsters practice such potentially dangerous (and partly unnecessary) pyrotechnic experiments.

However, when considered, Partin is not to be blamed. She did not get the right support from her teachers, with the science project, that she should have -warning her not to make fundamental misconceptions. For publishing in a journal of good standing, a technical board should provide the support and supervision. Such a nice subject as the catalysis of whistle compounds deserved more serious attention. A missed chance.

A glance upon the second resort

The study of J. Toker, "Experimental Evaluation of Pyrotechnic Whistle Composition: The Effect of Density on Maximum Thrust and Gas Velocity", was published in Pyrotechnica XVI.

This will be a short glance on the study, because the research seems to be systematically rather well organized and looks quite interesting. However, there are also basic shortcomings to be noticed. While the article is written at a reasonably technical level, some simple scholarly errors are made, like making wrong chemical calculations. The stoichiometric ratio between potassium perchlorate and sodium benzoate must be 78.3: 21.7 instead of the wrongly stated 72.3: 27.7, and for potassium perchlorate and sodium salicylate it must be 75.2: 24.8, instead of the wrongly proposed 67.0: 33.0. It would not have been of that great concern, but Toker is making further calculations and assumptions with this incorrect data. Terminology like 'oxygen rich' and 'fuel rich' are frequently mixed up because of this and graphs wrongly interpreted. He also tries to derive regression trends on only a few questionable points in the graph, of which the individual values vary too much.

Furthermore, Toker is apparently neglecting sifting and the phenomenon of particle size as being of great importance -a little careless, to say the least. The graphs shown in his research look surprising at first, but Toker chooses rather strange and uncommon packing densities like 0.5, 1.0 and 1.5 g/cm$^3$ to experiment with, while normally with these compositions a density of 1.5 to 2.0 g/cm$^3$ is more appropriate. The low packing density of 0.5 g/cm$^3$ is not very interesting, since these are almost loose powders that can behave quite unpredictably and unreliable. In spite of the large data tables given, there are still too few sets of experimental data, and some data differs too significantly to be of great value. As stated earlier, again it is shown that large and difficult explainable differences, also per batch, are quite remarkable for whistle compositions.

Literature cited:


