

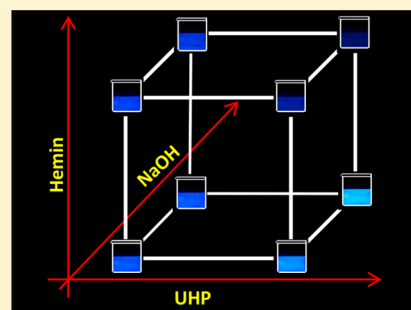
# Applications of Chemiluminescence in the Teaching of Experimental Design

Tomasz Krawczyk,\* Roksana Słupska, and Stefan Baj

Department of Chemical Organic Technology and Petrochemistry, Faculty of Chemistry, Silesian University of Technology, ul. Krzywoustego 4, 44-100 Gliwice, Poland

**S** Supporting Information

**ABSTRACT:** This work describes a single-session laboratory experiment devoted to teaching the principles of factorial experimental design. Students undertook the rational optimization of a luminol oxidation reaction, using a two-level experiment that aimed to create a long-lasting bright emission. During the session students used only simple glassware and registered the impact of the concentrations of a catalyst, a base, and an oxidant on chemiluminescence with their own eyes. The laboratory facilitated the development of hands-on experience with the statistical design of experiments, the foremost of which was the intuitive understanding of this industrially and scientifically relevant research technique.



**KEYWORDS:** Upper-Division Undergraduate, Laboratory Instruction, Interdisciplinary/Multidisciplinary, Chemometrics

## INTRODUCTION

In a conventional experiment each investigated factor is individually varied, while all of the others are held fixed. Although this seems intuitive, it is neither the only nor the most effective way of conducting research. First of all, it ignores possible synergistic effects (interactions) involving two or more factors. In other words, it assumes that those factors behave independently, which leads to an incomplete understanding of the system. Second, each experiment serves only one purpose, that is, to evaluate the impact of a given factor.

A more insightful method of experimentation is to vary all of the factors of interest simultaneously in a systematic fashion and to measure their impact (response) on the investigated system. This approach, called factorial experimentation (FE), offers many advantages, such as the possibility of identifying interactions and a more efficient utilization of data. Moreover, the number of experiments may be easily predicted and adjusted depending on the number of factors, the depth of knowledge desired, and economic constraints. It is worth pointing out here that the reader may encounter a few equivalent terms in literature that mean a variable in an experiment that influences the outcome: factor, independent variable, and parameter. The most often used is “factor”.

The subject of FE has rich literature. Monographs addressed to various audiences have been frequently published, especially in the recent decade. It has also received a considerable coverage in trade journals. Excellent examples of practical applications,<sup>1</sup> as well as broader discussions on modern developments in the field,<sup>2</sup> can be easily found.

Although the principles of FE have been discussed in the chemical literature since the 1940s,<sup>3,4</sup> it is still an overlooked area in scientific education. A short survey of chemical literature

proves that the principles of FE are rarely applied to real research and that the traditional “one factor at a time” method is still prevalent. The main difficulty in the widespread application of FE is probably the reluctance among many chemists to apply statistics during research if not required in a particular area. The other important reason is the lack of adequate courses<sup>5</sup> at the undergraduate or graduate level that would familiarize students with the technique to the point that they are able to use and benefit from it in their future work.

Since 1990, when a valuable review of basic concepts in experimental design appeared in this Journal,<sup>6</sup> a number of articles dealing with the teaching of FE have been regularly published. These were most often devoted to the application of FE in the optimization of analytical procedures, such as HPLC and GC separation,<sup>7,8</sup> atomic absorption spectroscopy (AAS) detection,<sup>9</sup> spectrophotometry,<sup>10,11</sup> and gravimetry.<sup>12</sup> Some of the published sessions were designed for organic synthesis courses.<sup>13–15</sup>

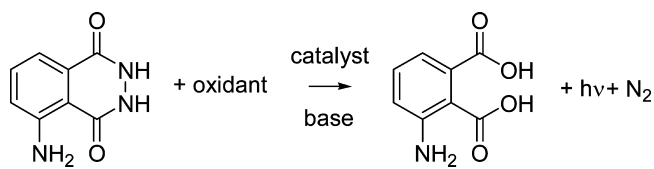
Usually the FE teaching presented in those contributions relied on incorporation of experimental design into an analytical or organic chemistry course. Therefore, with little or no additional time, a laboratory was diversified and students were able to see how FE works in practice. The session described here was developed as a part of a computer-based statistics course that is taught on chemistry's third year of the undergraduate studies. Our aim was to present FE in a simple and attractive way by letting students investigate the impact of key reagents' concentrations on the length of light emission accompanying luminol oxidation. As the outcome of this process can be visually monitored and quantitatively assessed,

each student or a pair of students can perform all of the required runs and obtain their own sets of data without any reliance on analytical instrumentation. This enabled an intuitive understanding of how each factor affects chemiluminescence (CL) and provided a bridge to statistical evaluation of the results. Because of simplicity and low requirements regarding chemicals and glassware, the session can be easily implemented in any setting. It also gives a possibility to use the session as an outreach activity or to teach experimental design at an earlier stage of study.

## THEORY

The light-yielding luminol oxidation is a complex process, which may be summarized in Scheme 1.

**Scheme 1. Simplified Luminol Oxidation Process**



In the first step, a catalyst, usually an iron complex, reacts with an oxidant and its oxidized (active) form reacts with luminol. Then the partially oxidized luminol reacts with another molecule of an oxidant. In the following steps nitrogen is released and blue light is emitted when the 3-aminophthalic acid in an excited state releases its energy.<sup>16</sup> The mechanism indicates that besides temperature and solvent, the final effect of the process may depend on the concentrations of the four main reagents: luminol, a catalyst, a base, and an oxidant. For the purpose of simplicity we decided to fix luminol concentration and investigate only the three remaining factors. According to our experience we selected the most convenient chemicals: hemin, NaOH, and urea hydrogen peroxide (UHP) and their concentrations.

The impact of these on light emission can be investigated either by using a screening design, or a response surface design.<sup>1,2,6</sup> The simplest is a two-level design, which requires from  $k + 1$  to  $2^k$  independent runs, where  $k$  is the number of factors. Each factor is set at only two settings (levels), low and high, usually denoted as  $-1$  and  $+1$  (so-called coded values). This type of design is especially convenient for initial investigations and when a large number of factors are considered. Unfortunately, it cannot detect a pure quadratic

dependence on a factor; therefore, only a simple relationship between the investigated factors given by eq 1 can be found. Such an empirical model is typically the aim of performing FE and allows for the evaluation of the impact of each factor and the prediction of the outcome of any factor's setting.<sup>14</sup>

$$y = a_0 + \sum_{i=1}^k a_i x_i + \sum_{i=1}^{k-1} \sum_{j=i+1}^k a_{ij} x_i x_j \quad (1)$$

In the case of three factors, eight is the number of experiments sufficient to establish a model for the length of the CL emission ( $y$ ) in terms of the three factors  $x_i$ . In total, four coefficients  $a_i$  and three  $a_{ij}$  need to be calculated. The remaining one degree of freedom is necessary for the statistical significance evaluation if no replications of the runs have been made. Alternatively, a three-way interaction  $x_1 x_2 x_3$  can be added to the model.

In FE each run is performed according to an orthogonal array, which could be prepared manually by finding all the possible combinations of factor levels, or more conveniently by a computer program. There are numerous general statistical packages or design-of-experiments-specific software that enable designing experimental matrices and offer intuitive tools to analyze the results. For the purpose of this session we used StatSoft's Statistica as it is readily available at universities under a campus-wide license. A list of alternatives is provided in the Supporting Information. An example plan is presented in Table 1. The interactions were calculated by multiplying the appropriate factor values. Besides the fact that the factors are not correlated (orthogonal) with each other, the order of runs is random.

## EXPERIMENT

Prior to the session, students were expected to understand the basic principles of the regression analysis and the analysis of variance. In our case the laboratory was a final part of the statistics course. It was preceded by a 2 h tutorial, where students were familiarized with the basics of FE. They were solving simple tasks such as the preparation of an experimental matrix, and how the interpretation of data is carried out. Finally, the luminol oxidation and the way of investigating the impact of the concentrations of reagents on the length of light emission were discussed.

At the beginning of the session students were given handouts with concentrations of the reagents and an experimental plan with coded values. To simplify the work, stock solutions had

**Table 1. Experimental Plan and a Set of Results Generated by a Pair of Students<sup>a</sup>**

Run	Coded Values			Interactions			CL Emission Length (s)	
	Hemin	NaOH	UHP	Hemin NaOH	Hemin UHP	NaOH UHP	Experimental	Predicted
1	-1	1	1	-1	-1	1	2880	2941
2	1	1	1	1	1	1	35	-26
3	1	-1	-1	-1	-1	1	282	2201
4	-1	1	-1	-1	1	-1	525	464
5	-1	-1	-1	1	1	1	500	561
6	1	-1	1	-1	1	-1	72	133
7	1	1	-1	1	-1	-1	199	260
8	-1	-1	1	1	-1	-1	3300	3239

<sup>a</sup>Concentrations ( $\text{g dm}^{-3}$ ). Luminol: 0.1. High level (1): hemin, 0.05; NaOH, 20; UHP, 0.2. Low level ( $-1$ ): hemin, 0.005; NaOH, 2; UHP, 0.02. Luminol and hemin solutions contained  $0.01 \text{ g dm}^{-3}$  NaOH. Predicted values were calculated using coefficients  $a_0$ ,  $a_i$ , and  $a_{ij}$  from Table 4 according to eq 1.

been prepared in advance in high and low concentrations (Table 1). The ranges were adjusted to achieve a bright and reasonably long emission for all samples. For this reason suitable concentrations had been predetermined earlier.

Students worked alone or in pairs. Each group was assigned with all of the runs required. This translates into the requirement of eight medium and eight small beakers, four measuring cylinders, and a funnel for each working group as sufficient equipment. First, students combined appropriate stock solutions of luminol, hemin, and NaOH according to the plan. Separately, they measured off aliquots of UHP solutions. This step takes no more than 40 min. When all of the groups were ready, the laboratory was darkened. To initiate the reaction a UHP solution was added to the corresponding mixture of the remaining reagents, and from this point students measured in seconds the time when they could see the emission using their cell phones. It also provided enough light to take notes. This part of the session was completed within an hour as two to three measurements could be done simultaneously. The way to do that was left to students' resourcefulness. For example, after one or two measurements, students were able to anticipate how long the CL would last and either wait until it ceased or start to measure another sample. Notes about the exact time of their mixing the solutions allowed them to calculate the length of the emission afterward.

The visual assessment is slightly subjective and also depends on the intensity of external illumination. Another important factor which contributes to the variance between students' results is the accuracy of the volume measurements. The observed deviations from results given in Table 1 were up to 50%. Nevertheless, the differences between the emissions' lengths within the plan are almost 100-fold (Table 1); therefore, the final results are sufficiently precise to perform a meaningful analysis.

## HAZARDS

All of the reagents used are fairly safe but must be handled carefully. Eye protection and gloves should be used because of the irritant nature of NaOH and UHP solutions. Luminol, hemin, and UHP are not considered to be toxic, but hemin may create stains that are difficult to remove.

## ANALYSIS OF RESULTS

A statistical evaluation of the results can be done by means of the regression analysis (eq 1) and by the analysis of variance<sup>17</sup> (ANOVA) for multiple factors. In the case of two-level designs it is also possible to evaluate the impact of each factor and its interactions simply by comparing the averages for the low and high factors' levels (Table 2).<sup>4</sup> The interpretation of these values (effects) is simple and straightforward. For example, if the level of hemin is changed from low to high, the light emission is shortened by 1654.25 s (on average); therefore the hemin concentration has a negative impact on the emission length. The NaOH concentration had a little negative effect, whereas UHP had a strong positive impact. The physical sense of the interactions is the effect that the simultaneous change of the two factors has on the system. In this case, if hemin and UHP were changed together, it shortened the light emission by 1382.25 s. From Table 2 it is also clear that hemin is more important than UHP in directing the change of the emission length and that the effect of NaOH is approximately 10 times

**Table 2. Impact of the Investigated Factors on Light Emission**

	Average <sup>a</sup> (s)		Effect (s)
	low level $\bar{y}_{\text{low}}$	high level $\bar{y}_{\text{high}}$	$\bar{y}_{\text{high}} - \bar{y}_{\text{low}}$
Hemin	1801.25	147	-1654.25
NaOH	1038.5	909.75	-128.75
UHP	376.5	1571.75	1195.25
Hemin NaOH	939.75	1008.5	68.75
Hemin UHP	1665.25	283	-1382.25
NaOH UHP	1024	924.25	-99.75
Overall average	974.13		

<sup>a</sup>Based on eight runs from Table 1.

lower than that of the remaining factors. It also does not considerably interact with other components.

Comparison of effects is often the last step of analysis in screening experiments with  $k + 1$  runs or when the number of calculated effects is  $k - 1$ . Nevertheless, FE offers the possibility of assessing the statistical significance of results by ANOVA. The use of such a capability ensures that the real effects can be distinguished from those arising from the random error. The analysis can be summarized in Table 3.

**Table 3. Analysis of Variance<sup>a</sup>**

Effect	SS	df	MS	F	p
Hemin	5,473,086	1	5,473,086	181.6182	0.047
NaOH	33,153	1	33,153	1.1001	0.485
UHP	2,857,245	1	2,857,245	94.8144	0.065
Hemin NaOH	9,453	1	9,453	0.3137	0.675
Hemin UHP	3,821,230	1	3,821,230	126.8032	0.056
NaOH UHP	19,900	1	19,900	0.6604	0.566
error	30,135	1	30,135		
TSS	12,244,203	7			

<sup>a</sup>Based on eight runs from Table 1.

The total sum of squares (TSS)<sup>17</sup> reflects the variance in the entire sample. The subtraction of all of the sums of the squares of the individual effects from TSS yields the residual mean square (error). Since each sum of the squares (SS) is calculated as a single comparison between two sets of results (averages for a high and low level), it has one degree of freedom. Finally, the *F*-criterion must be compared with critical values of the Fisher–Snedecor distribution at 1 (numerator) and 1 (denominator) degrees of freedom. In our example the effect of hemin is significant, while the effects of UHP and hemin×UHP are only marginally significant ( $0.05 < p < 0.1$ ). The other three effects are not significant.

In our example we excluded the possibility of the three-way interaction  $x_1x_2x_3$ , which allowed for the performance of the significance test. The other way would be performing replication runs, which would give enough measurements to investigate whether an  $x_1x_2x_3$  interaction is real and subsequently allow for a calculation of the *F*-factors. In many cases a scientist has at least a rough idea of the uncertainty in a particular experiment, and this information could be used as an error estimate.

The other complementary approach of the effects' estimation is to use the least-squares multiple regression and eq 1.<sup>6</sup> The results are summarized in Table 4. If a regression analysis is performed using normalized values, then the coefficients are

Table 4. Regression Analysis<sup>a</sup>

Source	Coefficient	Standard Error	<i>t</i> (1)	<i>p</i>
Intercept ( <i>a</i> <sub>0</sub> )	974.125	61.375	15.8717	0.040
Hemin ( <i>a</i> <sub>1</sub> )	−827.125	61.375	−13.4766	0.047
NaOH ( <i>a</i> <sub>2</sub> )	−64.375	61.375	−1.0489	0.485
UHP ( <i>a</i> <sub>3</sub> )	597.625	61.375	9.7373	0.065
Hemin NaOH ( <i>a</i> <sub>12</sub> )	34.375	61.375	0.5601	0.675
Hemin UHP ( <i>a</i> <sub>13</sub> )	−691.125	61.375	−11.2607	0.056
NaOH UHP ( <i>a</i> <sub>23</sub> )	−49.875	61.375	−0.8126	0.566

<sup>a</sup>Based on eight runs from Table 1.

exactly half the effects from Table 2, while the intercept is equal to the overall average of responses. The standard errors of the estimate are equal for all the coefficients. The significance of the coefficients is determined according to *t*-student distribution. The results from Table 4 are often presented graphically using a Pareto chart.<sup>18</sup>

In order to verify the quality of the model, it is possible to predict the responses using eq 1 and the data from Table 4 and compare them with the actual results (Table 1). It is usually done by plotting the difference between the empirical and the predicted values (residuals) against either the empirical or the predicted values. Residuals of a well fitted model are randomly distributed around zero.<sup>19</sup> In some cases a nonlinear transformation of responses may solve the problem of correlation of residuals with the response. During this step it is easy to identify any outlying results that could have occurred due to a blunder.

The other issue that is inherent to empirical models such as eq 1 is that they do not exactly reflect the complexity of the investigated systems. This unsuitability of the model contributes to the error of the effect's estimation. With the replication of runs, it is possible to partition the error into two elements. The first is the so-called "pure error", which corresponds to the precision of the measurements. The second part is the lack of fit.<sup>20</sup>

In the case of the two-level designs, a central point with all of the factor settings set halfway between the low level and the high level may be added to the plan. It allows for a validation of the linearity of the response by comparing the predicted and the actual response. If they differ significantly, it means that the quadratic terms are important for describing the response accurately and further experiments would be required. Usually it is done by extending the design to three or five levels<sup>10</sup> (central composite). Replication of a central point is also a common way to estimate the random error.

## CONCLUSION

The proposed session has many advantages. Students can see with their own eyes how the different factors influence a visually attractive CL reaction as each student can perform all of the required runs. The proposed reaction is sensitive to concentrations of reagents; therefore, the results are clearly distinguishable and meaningful and can be easily interpreted. The proposed experiments can be performed in almost any setting, and no expensive or hazardous materials are used. The session is not time-consuming and uses only a small amount of low-cost chemicals and simple glassware (beakers and measuring cylinders). Moreover, the required reagents can be easily replaced. For example, instead of hemin any iron salt or complex, such as K<sub>3</sub>Fe(CN)<sub>6</sub>, can be used; UHP can be replaced with any strong oxidant (K<sub>2</sub>S<sub>2</sub>O<sub>8</sub> or H<sub>2</sub>O<sub>2</sub>), whereas

commercial luminol can be obtained from phthalic acid during the organic synthesis classes. The purity of the chemicals is not crucial. The only limitation is working in a partially darkened laboratory in order to see the CL emission but with enough light to take notes and mix the solutions.

The experiment as previously described is suitable for a 2 h laboratory session. If more time is available, it could be easily extended in various ways. The same reaction can be applied to illustrate other types of experimental designs, and more factors can be added. Examples are provided in the Supporting Information.

The proposed session was incorporated into our statistics and numerical methods course for the first time during the winter term of the academic year of 2013/2014. In total 32 students participated. All very much enjoyed the session and all, except one, found it helped them to understand the principles of FE.

## ASSOCIATED CONTENT

### Supporting Information

Detailed instructions for students and teachers and further reading to aid in experimental design, statistical packages list, possible extended versions of the experiment, and PowerPoint slides introducing the design of experiments methodology, and calculation examples using Excel and Statistica. This material is available via the Internet at <http://pubs.acs.org>.

## AUTHOR INFORMATION

### Corresponding Author

\*E-mail: [tomasz.krawczyk@polsl.pl](mailto:tomasz.krawczyk@polsl.pl)

### Notes

The authors declare no competing financial interest.

## REFERENCES

- (1) Owen, M. R.; Luscombe, C.; Lai, Godbert, S.; Crookes, D. L.; Emiabata-Smith, D. Efficiency by Design: Optimisation in Process Research. *Org. Process Res. Dev.* **2001**, *5*, 308–323.
- (2) Tye, H. Application of statistical 'design of experiments' methods in drug discovery. *Drug Discovery Today* **2004**, *9*, 485–491.
- (3) Smallwood, H. M. Design of Experiments in Industrial Research. *Anal. Chem.* **1947**, *19*, 950–952.
- (4) Box, G. E. P. Statistical design in the study of analytical methods. *Analyst* **1952**, *77*, 879–891.
- (5) Öberg, T. Introducing Chemometrics to Graduate Students. *J. Chem. Educ.* **2006**, *83*, 1178–1181.
- (6) Strange, R. S. Introduction to experiment design for chemists. *J. Chem. Educ.* **1990**, *67*, 113–115.
- (7) Lang, P. L.; Miller, B. I.; Nowak, A. T. Introduction to the Design and Optimization of Experiments Using Response Surface Methodology. A Gas Chromatography Experiment for the Instrumentation Laboratory. *J. Chem. Educ.* **2006**, *83*, 280–282.
- (8) Harvey, D. T.; Byerly, S.; Bowman, A.; Tomlin, J. Optimization of HPLC and GC separations using response surfaces: Three experiments for the instrumental analysis laboratory. *J. Chem. Educ.* **1991**, *68*, 162–168.
- (9) Stolzberg, R. J. Screening and Sequential Experimentation: Simulations and Flame Atomic Absorption Spectrometry Experiments. *J. Chem. Educ.* **1997**, *74*, 216–220.
- (10) Palasota, J. A.; Deming, S. N. Central composite experimental designs: Applied to chemical systems. *J. Chem. Educ.* **1992**, *69*, 560–563.
- (11) Gozálviz, J. M.; García-Díaz, J. C. Mixture Design Experiments Applied to the Formulation of Colorant Solutions. *J. Chem. Educ.* **2006**, *83*, 647–650.



- (12) Oles, P. J. Fractional Factorial Experimental Design as a Teaching Tool for Quantitative Analysis. *J. Chem. Educ.* **1998**, *75*, 357–359.
- (13) Smith, R. B.; Billingham, E. J. Factorial design in undergraduate organic experiments. *J. Chem. Educ.* **1968**, *45*, 113–115.
- (14) Van Ryswyk, H.; Van Hecke, G. R. Attaining optimal conditions: An advanced undergraduate experiment that introduces experimental design and optimization. *J. Chem. Educ.* **1991**, *68*, 878–882.
- (15) Bouzidi, N.; Gozzi, C. Experimental Design and Optimization: Application to a Grignard Reaction. *J. Chem. Educ.* **2008**, *85*, 1544–1547.
- (16) Merényi, G.; Lind, J.; Eriksen, T. E. Luminol chemiluminescence: Chemistry, excitation, emitter. *J. Biolumin. Chemilumin.* **1990**, *5*, 53–56.
- (17) Hirsch, R. F. Analysis of Variance in Analytical Chemistry. *Anal. Chem.* **1977**, *49*, 691A–700A.
- (18) Loukas, Y. L. 2(k-p) fractional factorial design via fold over: Application to optimization of novel multicomponent vesicular systems. *Analyst* **1997**, *122*, 1023–1027.
- (19) Nelson, D. Using simple linear regression to assess the success of the montreal protocol in reducing atmospheric chlorofluorocarbons. *J. Stat. Educ.* **2009**, *17* (2), 1–22.
- (20) Coleman, D. E.; Vanatta, L. E. Lack-of-fit testing of ion chromatographic calibration curves with inexact replicates. *J. Chromatogr. A* **1999**, *850*, 43–51.