Using Design of Experiments to Characterize the Design Space of "Like-for- Like" Packaging Equipment

José R. Marrero Figueroa
Master of Engineering in Manufacturing Engineering
Edgar Torres, Ph.D.
Industrial Engineering Department
Polytechnic University of Puerto Rico

Abstract *— Industries with high volume demand tend to have various equipments that can perform the same task. Many of these tasks produce items that in order to sample them would require destroying them. Validation is necessary in this case. Using Design of Experiments, various treatments are applied, during a validation process, to the equipment, to prove that the outcome from each one will be repeatable, reproducible, expected and within the acceptable range. By using equipments from the same manufacturer, same brand, same model, and same capacity but in another location of the plant, the industries use the term "like-for-like" to lower or ease the validation process. However, the design space used by the manufacturer of the equipments allows it to provide the same equipment various times but with difference in their expected outcome; resulting in differences in the items produced by the industries in those equipments.*

Key Terms *— Design of Experiments, Design Space, "Like-for-Like", Validation.*

DEFINE

Many companies exclude "like-for-like" changes from their change control program. The FDA and other international regulators are becoming more skeptical to these changes because a product can be very sensitive to even the smallest changes. For example, when changing a defective pump that produces the input flow rate of product to a dryer; a pump of the same model, same design and same manufacturer was purchased to replace the defective pump, by definition a "like-for like" change. However, the new pump operates more efficiently, causing a difference in the particle size distribution of the final product that fails to meet the expected outcome. [1]

In that regard, using two equipments from the same manufacturer, same model and same design, which at the same time use the same single accessory that allows packaging the specific form of the product would still be exposed to two different variables, the location at the plant and the difference of efficiency of both equipments.

The packages as outputs from the two equipments will not be exactly the same. In order to test such outputs, tests are performed that result in destroying the package and damaging the integrity of the product inside. Therefore, validation, the process of establishing documented scientific evidence that a process or equipment is capable of consistently delivering quality results [2], is required for the packaging equipment.

During the validation process, the Design of Experiments, an information gathering exercise where variation is present that applies different treatments to understand the effects on the product, is used as a way to objectively present the scientific evidence required. In the case of the packaging equipments, treatments are formed by varying the parameters used as inputs that are, generally, temperature, pressure, and exposure time.

Design Space, the combination of input variables and process parameters that have been demonstrated to provide assurance of quality [3], is intended to be characterized as a result by the Design of Experiments; to further explain the differences that the two variables, location and efficiency are having on "like-for like" packaging equipments.

To understand the "like-for-like" presumption a table will be shown using two different equipments and the same set of accessories. The equipments

used will be named: A and C; on the other hand, the two accessories will be: A0 and A1.

MEASURE

Using the packaging equipment settings, a matrix of 16 treatments is created to understand the effects that it has on the samples and characterize if the design space is significantly different. Using an industrial tensile strength tester, a variable response is obtained for each equipment/ accessory arrangement.

Table 1
Parameter Matrix

Parameters			Y (lbf)			
			Equipment A		Equipment C	
T (^{0}F)	P (psi)	time (s)	A0	A1	A0	A1
258	60	4.00	1.768	1.810	1.730	1.758
258	90	4.00	2.051	1.957	1.912	1.900
230	60	4.00	1.403	1.309	1.659	1.717
230	60	9.00	1.744	1.627	1.768	1.567
230	90	4.00	1.475	1.713	1.669	1.579
230	90	9.00	1.874	1.877	1.859	2.061
258	90	9.00	1.870	2.101	1.905	2.174
258	60	9.00	1.969	2.160	1.980	1.958
230	90	4.00	1.576	1.823	1.589	1.446
230	60	9.00	1.626	1.738	1.936	1.773
258	90	4.00	1.795	1.733	2.057	1.789
258	90	9.00	2.114	1.806	1.875	2.150
258	60	9.00	2.104	1.941	2.254	2.077
230	90	9.00	1.865	1.855	1.827	2.029
230	60	4.00	1.441	1.458	1.242	1.582
258	60	4.00	1.729	1.834	1.882	1.959

Using the accessory A0 and then changing to the A1 on one of the equipments, would be considered a "like-for like" change, similarly, using Equipment A and then changing to Equipment C using a specific accessory would also be considered a "like-for-like" change.

ANALYZE

Using various commands from data sheet software, we obtain the following table:

Table 2
Basic Statistics

Equipment	A		C	
Acc	A0	A1	A0	A1
Max	2.114	2.160	2.254	2.174
Min	1.403	1.309	1.242	1.446
St Dev	0.22645	0.21279	0.2243	0.22702
Mean	1.775	1.796	1.822	1.845
Variance Coefficient	0.12756	0.11845	0.12314	0.12305

According to the coefficient of variance, equipment A with accessory A1 is the most consistent combination. Overall, changing from equipment A to C would not be exactly the same. However, if the "like-for-like" change is done with the accessories instead, changing accessory A0 for A1, on equipment C is almost equal, the same cannot be said for Equipment A, showing the highest different when changing from accessory A0 to A1.

Using Statistical Software Minitab, a graph is shown for each equipment/ accessory combination that shows how the three factors affect the response:

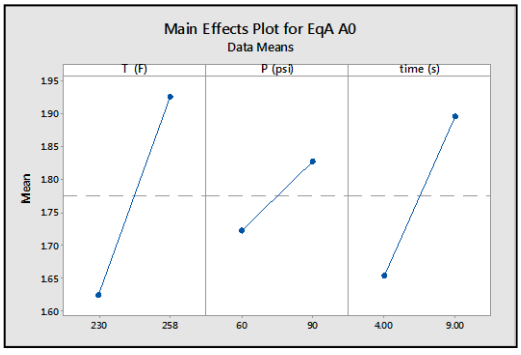

Figure 1
Main Effects Plot for Equipment A with Accessory A0

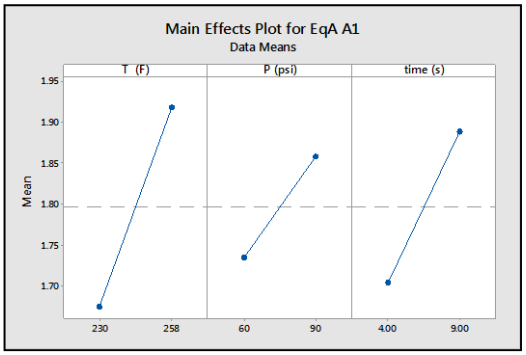

Figure 2
Main Effects Plot for Equipment A with Accessory A1

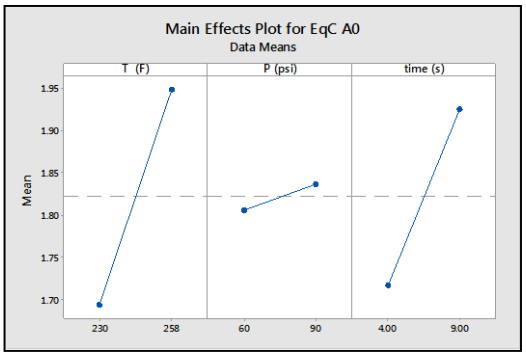

Figure 3
Main Effects Plot for Equipment C with Accessory A0

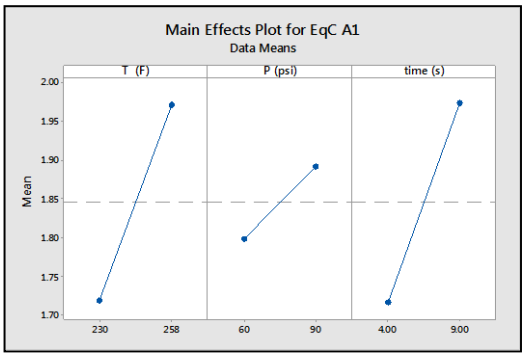

Figure 4
Main Effects Plot for Equipment C with Accessory A1

The temperature appears to affect the response in the same way for each combination. The exposure time has the same slope for the temperature only in the last combination, equipment C with accessory A1. In the other three combinations, the exposure time has the same inclination which is different to all other lines. The pressure is different in all combinations and does not have the same slope as any other line. In a main effects plot, a horizontal line means that the different levels of the factors do not affect the response. Therefore, the pressure affects the response the least of the three factors followed by the exposure time and then, by the temperature. The temperature is the factor that most affects the response.

In some of following analysis, there is the requirement that the data follows a normal distribution; four graphs are shown to prove the normality of the data.

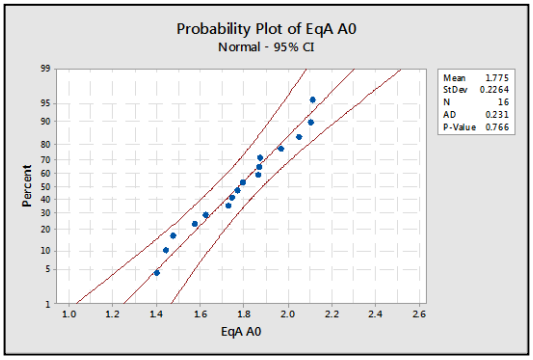

Figure 5
Normal Probability Plot for Equipment A with Accessory A0

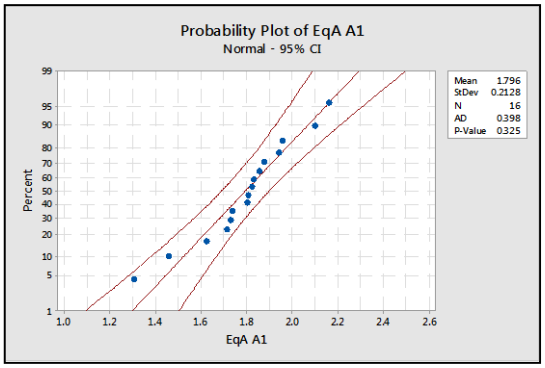

Figure 6
Normal Probability Plot for Equipment A with Accessory A1

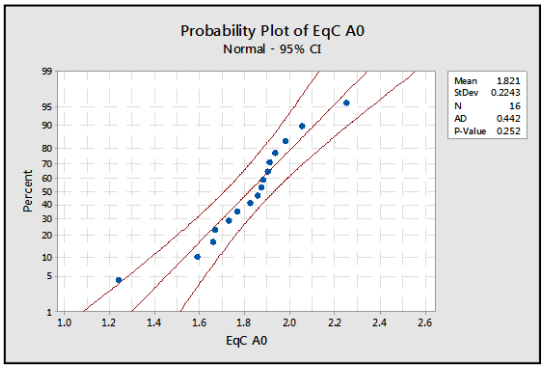

Figure 7
Normal Probability Plot for Equipment C with Accessory A0

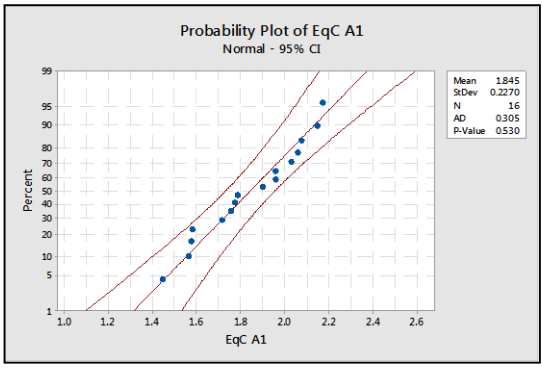

Figure 8
Normal Probability Plot for Equipment C with Accessory A1

Graphically, if most of the points are within the two edge lines then the data follows a normal distribution. Statistically, if the P-Value is higher than 0.05, then the data follows a normal distribution. Therefore, all of the combinations follow a normal distribution.

The following analysis is a Hypothesis Test, used to prove a null hypothesis where the mean from population 1 is different from population 2. The following lines were extracted from Minitab 17. Note: A confidence level of 95%, Error type I of 0.05 is used.

Two-Sample T-Test: EqA A0 vs EqA A1

	N	Mean	StDev	SE Mean
EqA A0	16	1.775	0.226	0.057
EqA A1	16	1.796	0.213	0.053

Difference = μ (EqA A0) - μ (EqA A1)
Estimate for difference: -0.0211
95% CI for difference: (-0.1800, 0.1378)
T-Test of difference = 0 (vs ≠): T-Value = -0.27
P-Value = 0.788 DF = 29

Two-Sample T-Test: EqC A0 vs EqC A1

	N	Mean	StDev	SE Mean
EqC A0	16	1.821	0.224	0.056
EqC A1	16	1.845	0.227	0.057

Difference = μ (EqC A0) - μ (EqC A1)
Estimate for difference: -0.0234
95% CI for difference: (-0.1866, 0.1397)
T-Test of difference = 0 (vs ≠): T-Value = -0.29
P-Value = 0.771 DF = 29

Two-Sample T-Test: EqA A0 vs EqC A0

	N	Mean	StDev	SE Mean
EqA A0	16	1.775	0.226	0.057
EqC A0	16	1.821	0.224	0.056

Difference = μ (EqA A0) - μ (EqC A0)
Estimate for difference: -0.0463
95% CI for difference: (-0.2092, 0.1167)
T-Test of difference = 0 (vs ≠): T-Value = -0.58
P-Value = 0.566 DF = 29

Two-Sample T-Test: EqA A1 vs EqC A1

	N	Mean	StDev	SE Mean
EqA A1	16	1.796	0.213	0.053
EqC A1	16	1.845	0.227	0.057

Difference = μ (EqA A1) - μ (EqC A1)
Estimate for difference: -0.0486
95% CI for difference: (-0.2077, 0.1105)
T-Test of difference = 0 (vs ≠): T-Value = -0.62
P-Value = 0.537 DF = 29

There are four comparisons shown, the first two are for the "like-for-like" change of one accessory to the other in the same equipment, the other two, are for the "like-for-like" change of one of the equipments to the other using the same accessory. Using the value of 29 degrees of freedom, given by Minitab in the last line of each comparison, a critical value of ±2.045 is obtained.

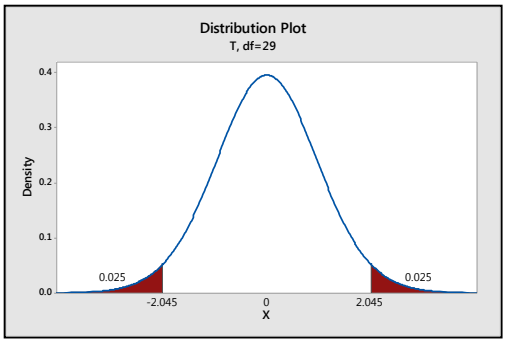

Figure 9
Critical T Value

Any negative T-value lower than -2.045 or any positive T-value higher than 2.045, would fall in the rejection area giving enough evidence to reject the null hypothesis. Likewise, if the P-value is lower than 0.025 then there is enough evidence to reject the null hypothesis. In all of the comparisons, these two

are not met; therefore, there is not enough evidence to reject the null hypothesis. In other words, the difference between the two population means for the "like-for-like" change of the two cases, changing an accessory for the other in any equipment or changing to one of the two equipments, that uses any of the two accessories, is not significant.

The following analysis is to show that all responses are within control and capable of performing under specifications.

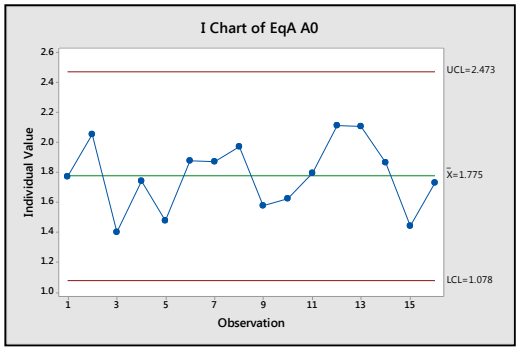

Figure 10
Equipment A with Accessory A0 Control Chart

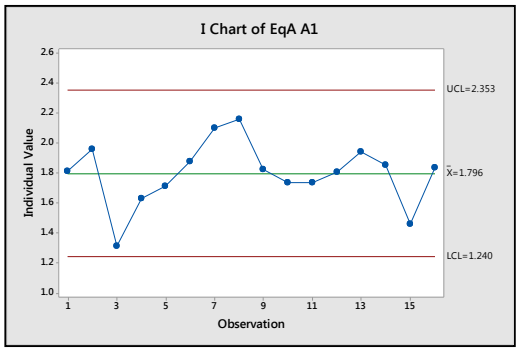

Figure 11
Equipment A with Accessory A1 Control Chart

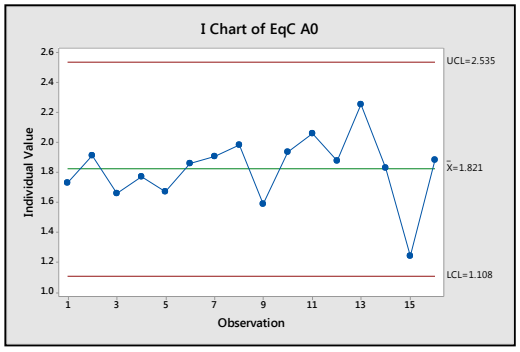

Figure 12
Equipment C with Accessory A0 Control Chart

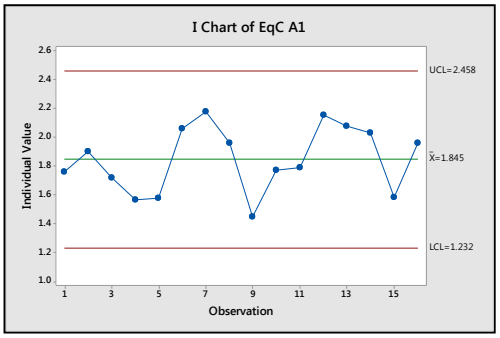

Figure 13
Equipment C with Accessory A1 Control Chart

The last 4 figures show that the data for all the combinations are within statistical control.

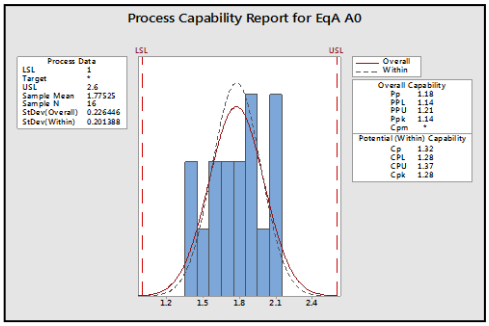

Figure 14
Equipment A with Accessory A0 Capability Analysis

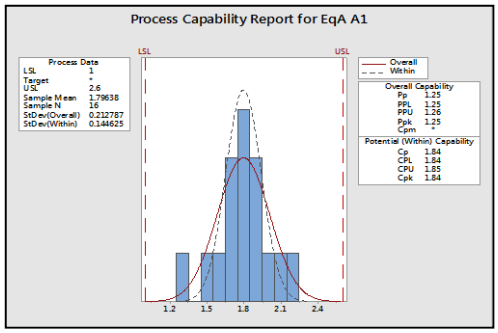

Figure 15
Equipment A with Accessory A1 Capability Analysis

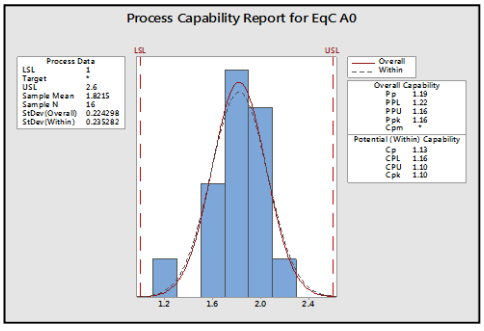

Figure 16
Equipment C with Accessory A0 Capability Analysis

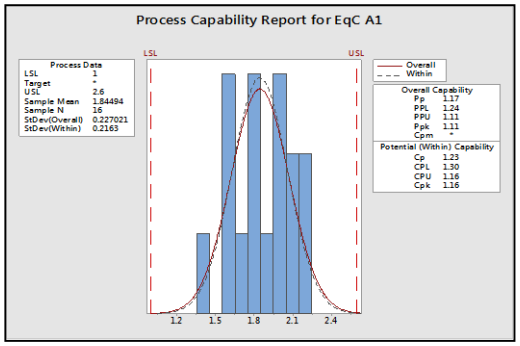

Figure 17
Equipment C with Accessory A1 Capability Analysis

In order for a process to be capable of performing under specifications, it must have a Cpk or process capability index of 1.33 or higher. The magnitude of Cpk relative to Cp is the direct measure of how off-center the process is operating. Cp may be equal or higher than Cpk. Cpk takes process centering into account. In other words, Cpk deals with the case of a process with a mean that is not centered between the specification limits. On the other hand, Cp only measures the spread of the specifications relative to the six sigma spread in the process [4]. This analysis is dependant of the specification limits used. This process only specification is that the response must be higher than 1.0 lbf. As an upper specification limit, a value of 2.6 lbf is used, however, it is not required; if a higher value is chosen the values of Cpk change. The most capable combination is equipment A with accessory A1 with a Cpk of 1.84. In this case, the value of Cp or Process Capability is equal to Cpk, this only happens when the mean of the process is centralized in respect to the specifications established. The other three combinations may have difficulty meeting customer requirements with Cpk values lower than 1.33; the Cpk values are 1.28, 1.16 and 1.10 for combinations equipment A with accessory A0, equipment C with accessory A1 and equipment C with accessory A0, respectively.

Finally, a Design of Experiment is performed using the matrix of 16 treatments. The following results are extracted from Minitab.

Factorial Regression: EqA A0 versus T (F), P (psi), time (s)

Analysis of Variance

Source	DF	Adj SS	Adj MS	F-Value	P-Value
Model	7	0.683933	0.097705	9.17	0.003
Linear	3	0.634806	0.211602	19.86	0.000
T (F)	1	0.358801	0.358801	33.68	0.000
P (psi)	1	0.043681	0.043681	4.10	0.077
time (s)	1	0.232324	0.232324	21.81	0.002
2-Way Interactions	3	0.026627	0.008876	0.83	0.512
T (F)*P (psi)	1	0.006241	0.006241	0.59	0.466
T (F)*time (s)	1	0.015625	0.015625	1.47	0.260
P (psi)*time (s)	1	0.004761	0.004761	0.45	0.523
3-Way Interactions	1	0.022500	0.022500	2.11	0.184
T (F)*P (psi)*time (s)	1	0.022500	0.022500	2.11	0.184
Error	8	0.085234	0.010654		
Total	15	0.769167			

Coded Coefficients

Term	Effect	Coef	SE Coef	T-Value	P-Value	VIF
Constant		1.7752	0.0258	68.80	0.000	
T (F)	0.2995	0.1497	0.0258	5.80	0.000	1.00
P (psi)	0.1045	0.0523	0.0258	2.02	0.077	1.00
time (s)	0.2410	0.1205	0.0258	4.67	0.002	1.00
T (F)*P (psi)	-0.0395	-0.0198	0.0258	-0.77	0.466	1.00
T (F)*time (s)	-0.0625	-0.0312	0.0258	-1.21	0.260	1.00
P (psi)*time (s)	-0.0345	-0.0173	0.0258	-0.67	0.523	1.00
T (F)*P (psi)*time (s)	-0.0750	-0.0375	0.0258	-1.45	0.184	1.00

Figure 18
Equipment A with Accessory A0 DOE Analysis

The DOE Analysis has two sections, the first is ANOVA or Analysis of Variance and the second is the Coded Coefficients. Both these sections are based on a Hypothesis Test. In the analysis of variance section if the F-value is higher than the critical F-value, or if the P-value is lower than 0.05, then there is enough evidence to reject the null hypothesis. The null hypothesis would be that the factor has no influence over the response. Using 1 degree of freedom for the numerator and 8 for the denominator, the critical F-value is 5.318. On the other hand, in the coded coefficients section, if the T-value is higher than the positive critical value or lower, than the negative critical value, or if the P-value is lower than 0.025, then there is enough evidence to reject the null hypothesis. The null hypothesis would be that the factor has no impact over the response. To find the critical T-value, it is necessary to calculate the degrees of freedom; using

error type I = 0.025, k = factors = 3 and n = repetitions = 2.

$$degrees\ of\ freedoms = 2^k(n-1) = 2^3(2-1) = 8 \quad (1)$$

The critical T-value would be ±2.306. For the first combination, equipment A with accessory A0, the factors time and temperature influence the response. On the other hand, the factor pressure, the interaction between temperature and pressure, the interaction between temperature and time, the interaction between pressure and time, as well as the interaction between the three factors do not influence the response. The same conclusion can be reached using the critical T-value in the coded coefficients section.

Factorial Regression: EqA A1 versus T (F), P (psi), time (s)

Analysis of Variance

Source	DF	Adj SS	Adj MS	F-Value	P-Value
Model	7	0.562750	0.080393	5.52	0.014
Linear	3	0.431408	0.143803	9.88	0.005
T (F)	1	0.235710	0.235710	16.20	0.004
P (psi)	1	0.061009	0.061009	4.19	0.075
time (s)	1	0.134689	0.134689	9.26	0.016
2-Way Interactions	3	0.129701	0.043234	2.97	0.097
T (F)*P (psi)	1	0.103041	0.103041	7.08	0.029
T (F)*time (s)	1	0.000900	0.000900	0.06	0.810
P (psi)*time (s)	1	0.025760	0.025760	1.77	0.220
3-Way Interactions	1	0.001640	0.001640	0.11	0.746
T (F)*P (psi)*time (s)	1	0.001640	0.001640	0.11	0.746
Error	8	0.116422	0.014553		
Total	15	0.679172			

Coded Coefficients

Term	Effect	Coef	SE Coef	T-Value	P-Value	VIF
Constant		1.7964	0.0302	59.56	0.000	
T (F)	0.2428	0.1214	0.0302	4.02	0.004	1.00
P (psi)	0.1235	0.0617	0.0302	2.05	0.075	1.00
time (s)	0.1835	0.0918	0.0302	3.04	0.016	1.00
T (F)*P (psi)	-0.1605	-0.0803	0.0302	-2.66	0.029	1.00
T (F)*time (s)	-0.0150	-0.0075	0.0302	-0.25	0.810	1.00
P (psi)*time (s)	-0.0803	-0.0401	0.0302	-1.33	0.220	1.00
T (F)*P (psi)*time (s)	0.0203	0.0101	0.0302	0.34	0.746	1.00

Figure 19
Equipment A with Accessory A1 DOE Analysis

For the second combination, equipment A with accessory A1, the factors time and temperature, as well as, the interaction between the temperature and pressure, influence the response. On the other hand, the factor pressure, the interaction between temperature and time, the interaction between pressure and time, as well as the interaction between the three factors do not influence the response. The same conclusion can be reached using the critical T-value in the coded coefficients section.

Factorial Regression: EqC A0 versus T (F), P (psi), time (s)

Analysis of Variance

Source	DF	Adj SS	Adj MS	F-Value	P-Value
Model	7	0.589823	0.084260	4.09	0.033
Linear	3	0.438349	0.146116	7.09	0.012
T (F)	1	0.261632	0.261632	12.70	0.007
P (psi)	1	0.003660	0.003660	0.18	0.684
time (s)	1	0.173056	0.173056	8.40	0.020
2-Way Interactions	3	0.139594	0.046531	2.26	0.159
T (F)*P (psi)	1	0.011881	0.011881	0.58	0.469
T (F)*time (s)	1	0.039800	0.039800	1.93	0.202
P (psi)*time (s)	1	0.087912	0.087912	4.27	0.073
3-Way Interactions	1	0.011881	0.011881	0.58	0.469
T (F)*P (psi)*time (s)	1	0.011881	0.011881	0.58	0.469
Error	8	0.164821	0.020603		
Total	15	0.754644			

Coded Coefficients

Term	Effect	Coef	SE Coef	T-Value	P-Value	VIF
Constant		1.8215	0.0359	50.76	0.000	
T (F)	0.2557	0.1279	0.0359	3.56	0.007	1.00
P (psi)	0.0302	0.0151	0.0359	0.42	0.684	1.00
time (s)	0.2080	0.1040	0.0359	2.90	0.020	1.00
T (F)*P (psi)	-0.0545	-0.0273	0.0359	-0.76	0.469	1.00
T (F)*time (s)	-0.0997	-0.0499	0.0359	-1.39	0.202	1.00
P (psi)*time (s)	-0.1483	-0.0741	0.0359	-2.07	0.073	1.00
T (F)*P (psi)*time (s)	-0.0545	-0.0272	0.0359	-0.76	0.469	1.00

Figure 20
Equipment C with Accessory A0 DOE Analysis

For the third combination, equipment C with accessory A0, the factors time and temperature influence the response. On the other hand, the factor pressure, the interaction between temperature and pressure, the interaction between temperature and time, the interaction between pressure and time, as well as the interaction between the three factors do not influence the response. The same conclusion can be reached using the critical T-value in the coded coefficients section.

For the fourth combination, equipment C with accessory A1, the factors time and temperature, as well as, the interaction between pressure and time influence the response. On the other hand, the factor

pressure, the interaction between temperature and pressure, the interaction between temperature and time, as well as the interaction between the three factors do not influence the response. The same conclusion can be reached using the critical T-value in the coded coefficients section.

Factorial Regression: EqC A1 versus T (F), P (psi), time (s)

Analysis of Variance

Source	DF	Adj SS	Adj MS	F-Value	P-Value
Model	7	0.699658	0.099951	10.89	0.002
Linear	3	0.551673	0.183891	20.04	0.000
T (F)	1	0.252758	0.252758	27.54	0.001
P (psi)	1	0.033948	0.033948	3.70	0.091
time (s)	1	0.264968	0.264968	28.87	0.001
2-Way Interactions	3	0.116745	0.038915	4.24	0.045
T (F)*P (psi)	1	0.002889	0.002889	0.31	0.590
T (F)*time (s)	1	0.001463	0.001463	0.16	0.700
P (psi)*time (s)	1	0.112393	0.112393	12.25	0.008
3-Way Interactions	1	0.031241	0.031241	3.40	0.102
T (F)*P (psi)*time (s)	1	0.031241	0.031241	3.40	0.102
Error	8	0.073416	0.009177		
Total	15	0.773075			

Coded Coefficients

Term	Effect	Coef	SE Coef	T-Value	P-Value	VIF
Constant		1.8449	0.0239	77.04	0.000	
T (F)	0.2514	0.1257	0.0239	5.25	0.001	1.00
P (psi)	0.0921	0.0461	0.0239	1.92	0.091	1.00
time (s)	0.2574	0.1287	0.0239	5.37	0.001	1.00
T (F)*P (psi)	-0.0269	-0.0134	0.0239	-0.56	0.590	1.00
T (F)*time (s)	-0.0191	-0.0096	0.0239	-0.40	0.700	1.00
P (psi)*time (s)	0.1676	0.0838	0.0239	3.50	0.008	1.00
T (F)*P (psi)*time (s)	-0.0884	-0.0442	0.0239	-1.85	0.102	1.00

Figure 21
Equipment C with Accessory A1 DOE Analysis

IMPROVE

In the two cases shown, either changing from an accessory to the other in the same equipment, or changing from one equipment to the other while using the same accessory, has been proven that any of the two can be treated as a "like-for-like" change. There is enough scientific evidence to support this fact. However, all this analysis must be done prior to execute such changes. There are some minor differences that have to be taken into consideration that can alter a process outcome if it is very sensitive to minor changes. For instance, taking into consideration the DOE analysis, both cases where the accessory A1 is employed show that it is influenced or impacted by a different interaction of the factors depending on the equipment used. In that regard, changing equipment A for equipment C while using the same accessory A0, appears to be the most adjusted "like-for-like" change in this article. The main effects plot supports this fact for this change and their variance coefficient is different but not that high of a difference. These two combinations were under performing in the capability analysis but it can be adjusted, if the upper specification is higher. All of the combinations showed to be in control, their data followed a normal distribution and the hypothesis test showed that no combination had a significant difference between another. Therefore, it is possible to approve any of the changes analyzed, there has to be documented proof that some adjustments are made and scientific evidence like the analysis made here to support the reason for these changes.

CONTROL

For any process that has been around a long time in any company, there has to be a population data. The population data may be extracted from the various tests done over time or from the validation. If a "like-for-like" change is made it is necessary to test various samples over time from the new equipment/ accessory arrangement to gather information to obtain population data. The reason behind this is to see if the adjustments done in the improve phase were accurate and not only provides the expected outcomes like the previous arrangement but also are within the acceptable range. If this is not true, and the process acquires a new mean, first of all, it is necessary to check if the adjustments made are the cause for this. If they are, it may a possibility reverting the adjustments and check if the process achieves an acceptable mean that satisfies the requirements. If that possibility is discarded, there should be a fishbone diagram analysis to find the root-cause and implement a corrective action/ preventive action. As long as there are no more changes to equipment/ accessory

arrangement there is not necessity for validation. If the root cause is found to be a defective equipment piece that will require some type of validation or the justification analysis shown above to not do the validation for that change.

REFERENCES

[1] Enkap, Inc. (2010, August 10). *Change Control Frequently Asked Questions and Answers Sample Document* [Online]. Retrieved on May 06, 2015 from: http://enkap.org/docs/change_control_faq_and_answers_sample.pdf, last accessed.

[2] FDA (2011, January). *Guidance for Industry Process Validation: General Principles and Practices* [Online]. Retrieved on May 06, 2015 from: http://www.fda.gov/downloads/Drugs/Guidances/UCM070336.pdf, last accessed.

[3] S. Chatterjee (2012, October 14). *Design Space Considerations* [Online]. Retrieved on May 06, 2015 from: http://www.fda.gov/downloads/AboutFDA/CentersOffices/OfficeofMedicalProductsandTobacco/CDER/UCM341156.pdf, last accessed.

[4] Ö. Senvar, *et al.*, (2010, November 02). *Process Capability and Six Sigma Methodology Including Fuzzy and Lean Approaches* [Online]. Retrieved on May 06, 2015 from: http://cdn.intechopen.com/pdfs-wm/12326.pdf, last accessed.