

Augmented Measurement System Assessment

NATHANIEL T. STEVENS

*Business and Industrial Statistics Research Group, Department of Statistics and Actuarial Sciences,
University of Waterloo, Waterloo, N2L 3G1 Canada*

RYAN BROWNE

Department of Mathematics and Statistics, University of Guelph, Guelph, N1G 2W1 Canada

STEFAN H. STEINER and R. JOCK MACKAY

*Business and Industrial Statistics Research Group, Department of Statistics and Actuarial Sciences,
University of Waterloo, Waterloo, N2L 3G1 Canada*

The standard plan for the assessment of the variation due to a measurement system involves a number of operators repeatedly measuring a number of parts in a balanced design. In this article, we consider the performance of two types of (unbalanced) assessment plans. In each type, we use a standard plan augmented with a second component. In type A augmentation, each operator measures a different set of parts once each. In type B augmentation, each operator measures the same set of parts once each. The goal of the paper is to identify good augmented plans for estimating the gauge repeatability and reproducibility (GR&R), a ratio that compares the contribution of the measurement system to the overall process variation. We show that, if there are three or more operators or if we include the possibility of part-by-operator interaction, then use of an appropriate augmented plan can produce substantial gains in efficiency for estimating GR&R compared with the best standard plan with the same total number of measurements.

Key Words: Gauge Repeatability and Reproducibility; Measurement Variation; Part-by-Operator Interaction; Repeatability; Reproducibility; Study Design.

Introduction

IN MANY manufacturing processes, parts will be measured to ensure that certain specifications are

Mr. Stevens is a Masters student in the Department of Statistics and Actuarial Science. His email address is nstevens@uwaterloo.ca.

Dr. Browne is a Post Doctoral Fellow in the Department of Mathematics and Statistics. His email address is rbrowne@uoguelph.ca.

Dr. Steiner is a Professor in the Department of Statistics and Actuarial Science and Director of the Business and Industrial Statistics Research Group. He is a senior member of ASQ. His email address is shsteiner@uwaterloo.ca.

Dr. MacKay is an Associate Professor in the Department of Statistics and Actuarial Science. He is a member of ASQ. His email address is rjmackay@uwaterloo.ca.

met. However, these measurements may be misleading if the measurement system (the devices, people, and protocol used to measure a part) is itself not adequate. Accordingly, many quality systems require the periodic assessment of critical measurement systems used to qualify parts or to make decisions about process control.

Here we deal with the assessment of a nondestructive measurement system that determines a single continuous characteristic or dimension. In assessing such a system with r operators, it is standard practice to randomly sample k parts from the current process and have each of the r operators measure each part n times for a total of $N = krn$ measurements. We call this a standard plan (SP), denoted by $SP(k, n)$. Common choices suggested by the forms in AIAG (2003, pp. 216–217) are $k \leq 10$; $r = 2, 3$;

$n = 2, 3$ so that $40 \leq N \leq 90$. In our experience, most measurement assessment studies in industry follow these guidelines closely. In this article, we look carefully at the standard plan and suggest a different allocation of the total number of measurements. We also consider augmented (unbalanced) plans in which some parts are only measured once by an operator.

There has recently been considerable activity in the development of different study plans and the corresponding analyses for measurement system assessments in a wide variety of situations. De Mast et al. (2010) look at ordinal measurement systems. Van Wieringen and De Mast (2008) and Danila et al. (2010) consider the assessment of binary measurement systems. Van Der Muelen et al. (2009) examine the assessment of a destructive measurement system. Browne et al. (2009b, 2010) look at the use of two-stage plans for assessing gauge repeatability & reproducibility (GR&R) where parts for the second stage (a standard plan) are selected based on an initial measurement.

For a standard plan, we use the following model to specify the attributes of interest in the measurement system and to analyze the data collected for the assessment:

$$Y_{ijl} = P_i + \mu_j + PO_{ij} + M_{ijl} \tag{1}$$

where $i = 1, 2, \dots, k$ (parts), $j = 1, 2, \dots, r$ (operators), $l = 1, 2, \dots, n$ (repeats), Y_{ijl} is a random variable representing the observed response for the repeated measurement l by operator j on part i , P_i is a random variable representing the variation of the true dimension of part i , M_{ijl} is a random variable that represents the measurement error when the same operator repeatedly measures the same part (i.e., measurement repeatability), and μ_j represents the effect of operator j . We also include the random variable PO_{ij} to allow the operator effect to change from part to part, i.e., to allow for interaction between the operator and parts. We make the additional distributional assumptions that $P_i \sim N(0, \sigma_p^2)$, $PO_{ij} \sim N(0, \sigma_{po}^2)$, $M_{ijl} \sim N(0, \sigma_m^2)$ and that all of these random variables are independent.

Many authors treat the operator effects as random. That is, they suppose the μ_j 's are a sample from a normal distribution. This assumption makes sense when there is a large number of operators and only a sample is available for inclusion in the assessment. In this instance, efficient plans require a large number of operators who may not be available in most manufacturing settings. Here we consider only

the fixed-effects case when a small number of operators are part of the measurement system and are all included in the assessment study. Because the true part value is a random effect, we adopt the traditional approach of describing the possible part-by-operator interaction using a random effect. The interaction is quantified parsimoniously by the single parameter σ_{po} . Burdick et al. (2005) discuss the issue of fixed versus random operator effects and provide an analysis for both cases. They also provide a large number of references.

We define

$$\sigma_o^2 = \frac{\sum_{j=1}^r (\mu_j - \bar{\mu})^2}{r}$$

as in Burdick et al. (2005, p. 83). Note that σ_o^2 quantifies the measurement variation due to the relative biases of each operator (i.e., measurement reproducibility) but is not a variance in the usual sense. Similarly, we define and interpret the total variation $\sigma_t^2 = \sigma_p^2 + \sigma_o^2 + \sigma_{po}^2 + \sigma_m^2$. If the system is automated with no operator effects or has a single operator ($r = 1$), we have $\sigma_o^2 = 0$ and we cannot estimate σ_{po}^2 separately from σ_m^2 , so we also set $\sigma_{po}^2 = 0$ in this case. We also separately consider the case when there is no part-by-operator interaction by setting $\sigma_{po}^2 = 0$.

A common metric for assessing a measurement system compares the variation due to the measurement system (repeatability and/or reproducibility) to the overall variation due to the process (i.e., due to differences in the true dimensions and the measurement system). We define the gauge repeatability and reproducibility (GR&R) (AIAG, 2003) as

$$\gamma = \sqrt{\frac{\sigma_o^2 + \sigma_{po}^2 + \sigma_m^2}{\sigma_t^2}}. \tag{2}$$

According to AIAG (2003, p. 77), a measurement system is deemed to be acceptable if γ is less than 0.1, unacceptable if γ is greater than 0.3, and is in need of improvement if $0.1 \leq \gamma \leq 0.3$. If the estimate of γ is large, we can examine the estimates of σ_o , σ_{po} , and σ_m separately to identify the source of the large measurement-system variation.

To estimate γ , we need a plan that provides an estimate of the underlying process variation σ_p . Alternate metrics, such as the precision to tolerance (PT) ratio, depend only on the measurement-system variation $\sqrt{\sigma_o^2 + \sigma_{po}^2 + \sigma_m^2}$ and do not require an es-

timate of σ_p . The optimal design of an assessment study to estimate the PT ratio and other such metrics will be different from what we propose. In other situations, the goal may be to estimate the individual variance components. This change of goal will lead to different assessment plans. In this paper, we focus on finding good plans for estimating γ while preserving some information about the separate variance components.

Other authors have compared balanced plans in the context of measurement system-assessment. If the measurement system is automated or has a single operator, Shainin and others (Shainin (1992), Traver (1995)) recommend an Isoplot[®] study, where $k = 30$ parts are selected and each is measured twice, i.e., $n = 2$. The Shainin plan provides better balance between the number of degrees of freedom available for estimating the measurement and process variation, and, as we shall see, is the optimal SP for estimating γ in this case. Walter et al. (1998) look at the optimal standard plan for estimating the interclass correlation using power calculations. They consider only plans with random operator effects that are subsumed into the repeatability component σ_m . Varde-man and Van Valkenberg (1999) look at the design of standard plans for N fixed with a flexible criterion that depends on which parameters are of most importance. They point out, somewhat facetiously, that if the goal is to estimate only the numerator of γ^2 (i.e., the total variance attributable to the measurement system), as is required for assessment criteria like the PT ratio, then it is best to use a single part. They also note that, if operator effects are random, then the usual standard plan with $r = 2$ or 3 cannot estimate σ_o^2 with any usable precision. We can add the same comment about estimating σ_p^2 with relatively few parts (i.e., 10), as is typical in practice. The reference manual of the AIAG (2003) on page 99 notes that their suggested number of appraisers, trials, and parts “represents the optimal conditions for conducting the study”. We shall see that this statement is far from the truth if we use a criterion based on the precision of the estimate for γ .

As noted by a referee, there is often available information from process records that can be incorporated into the analysis. For example, the GR&R analysis in Minitab[™] (2007) allows the substitution of an available estimate of the total variation. Browne et al. (2009a) and Danila et al. (2010), in the binary situation, demonstrate the considerable value of this extra information. In the current context, with fixed

operator effects, we may have additional process information by operator, i.e., estimates of μ_j for each operator (and hence σ_o) and $\sqrt{\sigma_p^2 + \sigma_{po}^2 + \sigma_m^2}$. The presence of this information will change the recommended plans described below. We do not pursue this issue further here.

In this paper, we compare standard plans with two types of augmented plans in which not all parts are measured the same number of times. In all cases, the number of operators r is fixed. The augmented plans have two components. One component is a standard plan using k parts with n repeated measurements by each operator. There are two possibilities for the other component

Type A: Randomly sample k_A parts (different from those selected in the SP component) where k_A is a multiple of r . Each operator measures k_A/r different parts once. We call this an A plan, denoted by $A(k, n, k_A)$.

Type B: Randomly sample k_B parts (different from those selected in the SP component). Each operator measures each of these parts once. We call this a B plan, denoted by $B(k, n, k_B)$.

Plan $A(k, n, k_A)$ has a total of $N = krn + k_A$ measurements using $k + k_A$ parts. Plan $B(k, n, k_B)$ has a total of $N = krn + rk_B = r(kn + k_B)$ measurements using $k + k_B$ parts. If we set k_A or k_B to zero, the corresponding augmented plan is an SP. Note that the second component of plan B corresponds to an SP with $n = 1$. The two components of an augmented plan can be conducted simultaneously or in any order. Within each component, every part is measured the same number of times.

The goal of this work is to identify augmented and standard plans that efficiently estimate γ when N , the total number of measurements available, is fixed. We measure efficiency of any augmented plan at a particular set of parameter values by comparing the asymptotic standard deviations (not variance) of the maximum likelihood estimates of γ from the augmented plan relative to the best SP. Here “best” means the standard plan with the smallest asymptotic standard deviation at the given parameter values. We search for augmented plans that have efficiency greater than 1 over a whole range of values for the unknown parameters.

If we set $n = 1$ in an SP or in the SP component of either plan A or B, then no part is measured more

than once by any operator. Looking at the model (1), we see that, in this case, σ_{po} and σ_m are not separately identifiable or estimable but $\sigma_{po}^2 + \sigma_m^2$ and, hence, γ can be estimated. Also, any plan B is now an SP with each of the r operators measuring $k + k_B$ parts. If we suspect that there is part-by-operator interaction that we want to identify separately, then we include only plans with $n > 1$ in the comparisons.

Here is the outline of the paper. In the next section, we derive the likelihood function and the Fisher information of the standard and augmented plans. Then we use the marginal asymptotic standard deviation of γ to rank various possible plans. We consider the special cases when there is no part by operator interaction (i.e., $\sigma_{po}^2 = 0$) and when there are no operator effects ($r = 1$) separately. In each case, we recommend specific plans and, when appropriate, calculate the efficiency of the recommended plans relative to the best standard plan. Because our choice is based on an asymptotic criterion, we also check the performance of the recommended plans using simulation. We provide access to software that a reader can use to evaluate other plans with N and r fixed. We also provide MATLAB[®] (2008) code for the analysis of data from an augmented plan based on maximum likelihood estimation. A brief discussion and summary follows.

Likelihood, Fisher Information, and Asymptotic Standard Deviations

We rank standard and augmented plans using the asymptotic precision as given by the inverse of the Fisher information matrix. We focus on estimation of γ with the parameters σ_o , σ_{po} , and σ_m considered secondary. With other goals, different plans would be recommended. To derive the information matrix, we need the likelihood.

Following the derivation in the Appendix and denoting the observed data y_{ijl} $i = 1, 2, \dots, k$, $j = 1, 2, \dots, r$ and $l = 1, 2, \dots, n$, the log-likelihood contribution from a standard plan or the SP component of an augmented plan is

$$l_1(\vec{\mu}, \sigma_p^2, \sigma_{po}^2, \sigma_m^2) = -\frac{1}{2} \left\{ b_1 \sum_{i=1}^k \sum_{j=1}^r \sum_{l=1}^n (y_{ijl} - \mu_j)^2 + b_2 \sum_{i=1}^k \sum_{j=1}^r \left[\sum_{l=1}^n (y_{ijl} - \mu_j) \right]^2 \right\}$$

$$+ b_3 \sum_{i=1}^k \left\{ \sum_{j=1}^r \sum_{l=1}^n (y_{ijl} - \mu_j) \right\}^2 - \frac{k}{2} \ln \left[(\sigma_m^2 + n\sigma_{po}^2 + rn\sigma_p^2)(\sigma_m^2 + n\sigma_{po}^2)^{r-1} \times (\sigma_m^2)^{r(n-1)} \right]. \tag{3}$$

We can find the log-likelihood contribution for the data from the augmented component of plan B by setting $n = 1$ in Equation (3). If we denote the observed data in this component by z_{ij} , $i = 1, \dots, k_B$; $j = 1, \dots, r$ we have

$$l_B(\vec{\mu}, \sigma_p^2, \sigma_{po}^2, \sigma_m^2) = -\frac{1}{2} \left\{ (b_1 + b_2) \sum_{i=1}^{k_B} \sum_{j=1}^r (z_{ij} - \mu_j)^2 + b_3 \sum_{i=1}^{k_B} \left[\sum_{j=1}^r (z_{ij} - \mu_j) \right]^2 \right\} - \frac{k_B}{2} \ln \left[(\sigma_m^2 + n\sigma_{po}^2 + r\sigma_p^2)(\sigma_m^2 + \sigma_{po}^2)^{r-1} \right]. \tag{4}$$

For an augmented plan A, each operator measures different parts once, and measurements on all parts are independent. From model (1), we have, for any measurement made by operator j , $Z_j \sim N(\mu_j, \sigma_p^2 + \sigma_{po}^2 + \sigma_m^2)$ and so, denoting the observed measurements by z_{ij} , $i = 1, \dots, k_A/r$; $j = 1, \dots, r$ the log-likelihood contribution from the augmented component of plan A is

$$l_A(\vec{\mu}, \sigma_p^2, \sigma_{po}^2, \sigma_m^2) = -\frac{1}{2} \frac{\sum_{j=1}^r \sum_{i=1}^{k_A/r} (z_{jl} - \mu_j)^2}{\sigma_p^2 + \sigma_{po}^2 + \sigma_m^2} - \frac{k_A}{2} \ln(\sigma_p^2 + \sigma_{po}^2 + \sigma_m^2). \tag{5}$$

The overall log likelihood is then the sum of the contributions from Equations (3) and (5) for plan A and Equations (3) and (4) for plan B.

To calculate the asymptotic standard deviations for any assumed values for the unknown parameters, we find the Fisher information matrix symbolically using MapleTM to calculate the appropriate second derivatives of the overall likelihood and then substitute the expected values and change variables to give results in terms of γ . The derivation is sketched in the Appendix.

We also consider the special case where we assume there is no part-by-operator interaction. We set $\sigma_{po}^2 = 0$ in the above calculations up to the stage of finding the partial derivatives of the log-likelihood function. We then proceed as before, except we purge the appropriate row and column corresponding to σ_{po}^2 from the information matrix and the change of variables matrix D (given in the Appendix). We also use this calculation for the case $n = 1$ in the SP component for both type A and B plans, where σ_m^2 is now the sum of the repeatability and part-by-operator components of the variation. If we consider the case with a single operator (or no operator effects), we set $r = 1$ and $\sigma_{po}^2 = 0$ and $\sigma_o^2 = 0$ and alter the information matrix and the matrix D accordingly. Note the simplification of the parameter γ given in Equation (2) in these instances.

Without displaying the information matrix explicitly, we note two of its properties that have important consequences. First, because of the block diagonal form of both the information in the original parameterization and the matrix D , we can show that the final information matrix does not depend on the specific values of μ_1, \dots, μ_r but only on σ_o^2 . Second, because we can calculate the overall information by summing the components for each part, a scale change in the number of parts in a standard plan or in the SP component of a plan A or plan B produces the same scale change in the information that then acts inversely on the asymptotic variance. So, for example, if we have $r = 3$ operators, the relative efficiency of $A(5, 2, 30)$ to $SP(10, 2)$ (two plans each with 60 measurements) is the same as the relative efficiency of $A(7, 2, 42)$ compared with $SP(14, 2)$ (two plans each with 84 measurement).

Comparison of Plans

To compare plans, we suppose that the total number of measurements N and the number of operators r are fixed. Here we consider values of N between 60 and 100 with $1 \leq r \leq 4$. With each combination, we examine each possible SP, plan A (integer values k, n, k_A with $krn + k_A = N$), and plan B (integer values for k, n, k_B with $r(kn + k_B) = N$). We substitute a range of possible values for the unknown parameters and then rank all possible plans according to the asymptotic standard deviation of γ . Because γ is defined as the square root of a ratio of variances, with no loss of generality we can set $\sigma_t^2 = \sigma_p^2 + \sigma_o^2 + \sigma_{po}^2 + \sigma_m^2 = 1$. Note throughout the comparisons of augmented plans, we conservatively

use the asymptotic standard deviation of the estimate of γ (not its variance) from the best standard plan at the particular parameter values as the basis to calculate relative efficiency.

Plans with One or No Operator

Many measurement systems are automated with no operator effects. This also corresponds to a system with a single operator. In our formulation of the problem, we then have $r = 1$, $\sigma_o = \sigma_{po} = 0$, and γ , the parameter of interest, simplifies to $\gamma = \sqrt{\sigma_m^2 / (\sigma_p^2 + \sigma_m^2)}$. Also with $r = 1$, augmented plans A and B are equivalent. Both augmented plan types start with a standard plan with k parts each measured n times. Then, in the augmented component, we measure an additional k_A (or k_B) parts once.

Suppose $N = 60$ and the true value of γ equals 0.3. In Table 1, we list the best four plans in increasing order of the asymptotic standard deviation of γ . For purposes of comparison, we also include the standard plan with $k = 10$ parts, each measured $n = 6$ times.

The best plan is $SP(30, 2)$, the Shainin proposal for an IsoplotTM study. Not surprisingly, we see similar results (not presented here) for other values of N and γ . In general, when there are no operator effects, it is best to use a standard plan that balances the degrees of freedom for estimating σ_p and σ_m by minimizing the number of repeated measurements, i.e., choosing $n = 2$. There is a substantial improvement in precision for estimating γ for the Shainin plan over the default AIAG (2003) standard plan with 10 parts. Because of the scaling property, for example, we also have a 23% improvement using $SP(45, 2)$ over $SP(15, 6)$. Augmentation provides no benefit here. We can use standard ANOVA methods to analyze the data from the recommended standard plan.

TABLE 1. Five Plans for Estimating γ when $r = 1$, $N = 60$, and $\gamma = 0.3$

Plan	SE($\hat{\gamma}$)	SE($\hat{\sigma}_m$)	Relative efficiency
SP(30, 2)	0.0523	0.0387	1.00
A(29, 2, 2)	0.0525	0.0394	1.00
A(28, 2, 4)	0.0527	0.0401	0.99
A(16, 3, 12)	0.0529	0.0375	0.99
SP(10, 6)	0.0680	0.0300	0.77

Plans with More than One Operator and No Part-by-Operator Interaction

Now suppose we have more than one operator and we assume, as is commonly done in GR&R studies, that there is no part-by-operator interaction.

We consider three cases where the number of operators is $r = 2, 3$, or 4 , each with two values of N close to 60 and 96 . We use $N = 64$ for $r = 4$ so that there is a large number of possible augmented plans. We specify three values of $\gamma = 0.5, 0.3$, and 0.1 , corresponding to a poor, acceptable, and good measurement system. Because

$$\gamma = \sqrt{\frac{\sigma_o^2 + \sigma_m^2}{\sigma_p^2 + \sigma_o^2 + \sigma_m^2}}$$

in this case, we set $\sigma_p^2 + \sigma_o^2 + \sigma_m^2 = 1$ (without loss of generality because γ is not changed by a scale change) so that $\sigma_o^2 + \sigma_m^2 = \gamma^2$. We then specify $\sigma_m^2 = \delta\gamma^2$, where $\delta = 0.1, 0.5, 0.9$, to look at situations when the repeatability (σ_m) makes up a small, medium, or large proportion of the overall measurement system variation as captured by γ .

For each value of r, N , and the nine pairs of values for γ and δ , we rank all possible SP A, and B plans using the asymptotic standard deviation of the MLE of γ as described in the previous section. For example, with $r = 2$ and $N = 60$, there are 103 plans of type A and B and 8 standard plans. Table 2 presents a comparison of the best plans of each type when $r = 2, N = 60, \gamma = 0.3$, and $\delta = 0.1$ (σ_m is relatively small compared with σ_o). We also include the widely used AIAG (2003) standard plan SP(10,3).

Here there is a small gain over the best SP (about 7% reduction in standard error) in estimating γ with an augmented plan A with 5 parts measured twice by both operators and then two sets of 20 parts measured once by each operator, i.e., the $A(5, 2, 40)$ plan. All of the best plans are substantially better than the

SP with $k = 10$. The plan $A(5, 2, 40)$ has the same 7% gain over the best SP when $\delta = 0.1$ for other values of γ . However, for $\delta = 0.5$ or 0.9 , the plan $A(5, 2, 40)$ is 12% to 25% less efficient than the best SP. In this case, we recommend the SP(30, 1) plan.

We see a similar pattern for $r = 3$. For example, when $N = 60$, in all cases except when $\gamma = .5$ and $\delta = 0.9$, there is a plan A that is superior to the best SP. Unfortunately, the best plan A varies as the parameters are changed. When $\delta = 0.1$, the plan $A(3, 2, 42)$ has a relative efficiency of about 1.26. However for larger values of δ , this plan is 2% to 12% less efficient than the best SP. Accordingly, with $r = 3$, we recommend the SP(20, 1) plan.

In Table 3, we present a second example with $r = 4, N = 64, \gamma = 0.3, \delta = 0.5$. Here there is an 18% gain in estimating γ by using the best plan A rather than the best SP. With 4 operators, as shown by the right-most column of Table 4, the plan $A(4, 2, 32)$ does well over the entire parameter space compared with the best standard plan SP(16,1). There is significant improvement relative to the best SP in all cases except when $\gamma = 0.5$ (the measurement system is highly variable) and $\delta = 0.9$ (most of the measurement variability is due to repeatability, i.e., $\sigma_m \gg \sigma_o$). In this case, there is no loss in efficiency. Also (not shown here), there is no material difference between $A(4, 2, 32)$ and the best plan A for any value of γ and δ in our array.

Because these comparisons are based on an asymptotic criterion, we also checked the relative efficiency of $A(4, 2, 32)$ versus SP(16, 1) using simulation. We generated 10,000 samples for each plan using all pairs of values for $\gamma = 0.5, 0.3, 0.1$ and $\delta = 0.1, 0.5, 0.9$. We then calculated the maximum likelihood estimates of the parameters for each sample. We provide a summary of these estimates for both plans in Table 4. The plan $A(4, 2, 32)$ provides a less biased estimate of γ with smaller standard deviation over

TABLE 2. Comparison of Plans with $r = 2, N = 60, \gamma = 0.3, \delta = 0.1$

Plan type	SE($\hat{\gamma}$)	SE($\hat{\sigma}_m$)	SE($\hat{\sigma}_o$)	Relative efficiency
A(5, 2, 40)	0.0347	0.0173	0.0210	1.07
B(2, 2, 26)	0.0383	0.0119	0.0122	0.97
SP(30, 1)	0.0371	0.0122	0.0122	1.00
SP(10, 3)	0.0621	0.0095	0.0122	0.60

TABLE 3. Comparison of Plans with $r = 4, N = 64, \gamma = 0.3, \delta = 0.5$

Plan type	SE($\hat{\gamma}$)	SE($\hat{\sigma}_m$)	SE($\hat{\sigma}_o$)	Relative efficiency
A(4, 2, 32)	0.0456	0.0283	0.0366	1.18
B(2, 2, 12)	0.0567	0.0212	0.0265	0.95
SP(16, 1)	0.0537	0.0217	0.0265	1.00
SP(8, 2)	0.0720	0.0200	0.0265	0.75

TABLE 4. Simulated and Theoretical Comparison of $A(4, 2, 32)$ to $SP(16, 1)$ for $r = 4, N = 64$

γ	δ	$A(4, 2, 32)$		$SP(16, 1)$		Simulated efficiency	Theoretical efficiency
		Average	Standard deviation	Average	Standard deviation		
0.5	0.1	0.511	0.052	0.525	0.072	1.38	1.37
0.5	0.5	0.507	0.070	0.522	0.080	1.14	1.12
0.5	0.9	0.509	0.086	0.522	0.086	1.00	0.99
0.3	0.1	0.309	0.038	0.321	0.057	1.50	1.39
0.3	0.5	0.307	0.048	0.320	0.060	1.25	1.17
0.3	0.9	0.305	0.055	0.312	0.065	1.18	1.09
0.1	0.1	0.103	0.014	0.108	0.022	1.57	1.39
0.1	0.5	0.102	0.017	0.108	0.023	1.35	1.20
0.1	0.9	0.101	0.018	0.107	0.024	1.33	1.14

all of the parameter values except when $\gamma = 0.5, \delta = 0.9$, as predicted by the theoretical information calculations. In all cases, the simulated efficiency is higher than predicted by the asymptotic calculations.

We investigated and compared the best plans of each type for $r = 2, 3, 4$ and $N \approx 60, 90$ with γ and δ as described above. To save space, we do not present all of the results here. We draw the following conclusions when there is no part-by-operator interaction, based on this empirical investigation:

- The best standard plans have $n = 1$ and $k = N/r$. That is, we maximize the number of parts in the study. We can justify this conclusion by noting that maximizing the number of parts maximizes the degrees of freedom for estimating the part variation and, because we are using the part-by-operator sum of squares to estimate σ_m^2 , increasing the number of parts also increases the degrees of freedom for estimating σ_o^2 . Because each part is measured by each operator, parts act as blocks, so we also get good estimates of the operator means μ_1, \dots, μ_r and hence σ_o^2 by increasing the number of parts. Note that this conclusion is contrary to the AIAG (2003) recommended plans (see sample forms pp. 216–217) that suggest setting $n = 2$ or 3.
- In all cases, the best standard plan is superior to the best plan B.
- The best augmented plans have $n = 2$ and use a small number of parts in the SP component.
- For two or three operators, augmentation provides little gain unless σ_m is relatively small

compared with σ_o . If $\delta = 0.1$ and $r = 2, 3$, then the best plan A is about 6% ($r = 2$) and 20% ($r = 3$) more efficient in estimating γ than the best SP. These results are independent of N in the range $60 < N < 100$.

- With four operators and $N = 64$, the plan $A(4, 2, 32)$ is (almost) uniformly better than the best SP and the gains in efficiency are relatively large when $\delta < 0.5$. For any value of N , we can scale this plan and see the same gains in efficiency. For example, if $N = 96$, the plan $A(6, 2, 48)$ has the same good properties.
- The simulated results show that the asymptotic calculations are conservative. The actual efficiency of the recommended augmented plans is better than predicted by these calculations.

Plans with More than One Operator and Possible Part-by-Operator Interaction

Now we consider a measurement system with two or more operators in which we allow for the possibility of part-by-operator interaction. We proceed as in the case with no interaction with $r = 2, 3, 4$ and $N \approx 60, 90$. There are two added complications. First, we have an extra parameter, σ_{po} , and γ is given by Equation (2). We set $\sigma_p^2 + \sigma_o^2 + \sigma_{po}^2 + \sigma_m^2 = 1$ without loss of generality so that $\gamma^2 = \sigma_o^2 + \sigma_{po}^2 + \sigma_m^2$. The first two terms are due to operator-to-operator differences, so, for a given value of γ , we look at three cases of $\sigma_m^2 = \delta\gamma^2$ with $\delta = 0.1, 0.5, 0.9$, so the repeatability contribution to γ is relatively small to large. Then, for given values of γ and δ , we consider three cases, $\beta = 0.1, 0.5, 0.9$, where the contribution of σ_o^2 is a relatively small to large propor-

TABLE 5. Four Plans with $r = 2$, $N = 60$, $\gamma = 0.3$, $\delta = 0.5$, and $\beta = 0.5$

Plan type	SE($\hat{\gamma}$)	SE($\hat{\sigma}_m$)	SE($\hat{\sigma}_{po}$)	SE($\hat{\sigma}_o$)	Relative efficiency
A(11, 2, 16)	0.0552	0.0320	0.0678	0.0445	1.10
B(2, 2, 26)	0.0494	0.0713	0.1097	0.0341	1.23
SP(15, 2)	0.0607	0.0274	0.0581	0.0387	1.00
SP(10, 3)	0.0713	0.0237	0.0570	0.0433	0.85

tion of the total operator contribution ($\sigma_o^2 + \sigma_{po}^2$) to the measurement system variation. Algebraically, we have $\sigma_o^2 = \beta(1 - \delta)\gamma^2$ and $\sigma_{po}^2 = (1 - \beta)(1 - \delta)\gamma^2$.

The second complication is that we must decide if we are going to entertain plans with $n = 1$. In this case, we cannot separately estimate σ_{po}^2 and σ_m^2 , but we can estimate $\sigma_{po}^2 + \sigma_m^2$ and hence γ . Including $n = 1$ plans is equivalent to assuming that there is no interaction (or more accurately, that any interaction is subsumed by the repeatability σ_m^2), so that we should compare $n = 1$ plans with those alternatives considered in the previous section. Here we do not allow $n = 1$ plans, so that we can get separate estimates of σ_{po}^2 and σ_m^2 .

For given values of N and r , we look at the best plans of each type as we sweep across the 27 combinations of γ , δ , and β . Consider the case with $r = 2$, $N = 60$, $\gamma = 0.3$, $\delta = 0.5$ (the repeatability σ_m^2 is the same as the reproducibility $\sigma_o^2 + \sigma_{po}^2$), and $\beta = 0.5$ (operator and part-by-operator effects are equal). Table 5 gives the best plans of each type. For the sake of comparison, we also include the AIAG (2003) recommended SP(10, 3).

The asymptotic standard error for estimating γ is about 23% smaller for the best plan B compared with the best standard plan and about 18% smaller than the best plan A. In the best plan B, we use only two parts with two repeated measurements by each operator in the SP component. In the augmented component, we have a large number of parts (26) measured once by each operator. For this plan, the estimates of σ_{po}^2 and σ_m^2 are highly correlated because most of the information is about their sum. The plan SP(10, 3) is much less efficient than the best plan B. We found that a plan B with 2 parts measured twice by each of the two operators in the SP component was uniformly the best plan. This result is not surprising because this plan is very close to the corresponding standard plan with $n = 1$ that is more efficient for estimating γ but cannot separately estimate σ_{po}^2 and σ_m^2 .

For example, when $N = 60$ and $r = 2$, we see in Table 6 a comparison of $B(2, 2, 26)$ with the best standard plan SP(15, 2). We also include the results of a simulation with 10,000 samples to demonstrate how well the asymptotic calculations rank the plans.

TABLE 6. Simulated and Theoretical Comparison of $B(2, 2, 26)$ to SP(15, 2) for $r = 2$, $N = 60$, $\beta = 0.5$

γ	δ	$B(2, 2, 26)$		SP(15, 2)		Simulated efficiency	Theoretical efficiency
		Average	Standard deviation	Average	Standard deviation		
0.5	0.1	0.511	0.069	0.522	0.095	1.37	1.33
0.5	0.5	0.512	0.077	0.524	0.093	1.21	1.20
0.5	0.9	0.513	0.082	0.530	0.093	1.13	1.11
0.3	0.1	0.309	0.049	0.320	0.068	1.39	1.33
0.3	0.5	0.311	0.051	0.322	0.070	1.37	1.23
0.3	0.9	0.309	0.056	0.327	0.068	1.21	1.16
0.1	0.1	0.105	0.018	0.107	0.025	1.39	1.34
0.1	0.5	0.104	0.019	0.106	0.025	1.32	1.24
0.1	0.9	0.103	0.019	0.108	0.026	1.37	1.17

TABLE 7. Relative Efficiencies of Some Good Type A Plans with 3 or 4 Operators

γ	δ	β	$m = 3, N = 60$			$m = 4, N = 64$		
			$A(5, 2, 30)$	$A(6, 2, 24)$	$A(7, 2, 18)$	$A(4, 2, 32)$	$A(5, 2, 24)$	$A(6, 2, 16)$
0.5	0.1	0.1	0.99	1.04	1.07	1.10	1.15	1.16
0.5	0.1	0.5	1.17	1.19	1.19	1.31	1.32	1.29
0.5	0.1	0.9	1.51	1.47	1.40	1.70	1.63	1.49
0.5	0.5	0.1	1.05	1.09	1.11	1.17	1.21	1.20
0.5	0.5	0.5	1.15	1.18	1.18	1.29	1.30	1.27
0.5	0.5	0.9	1.28	1.28	1.26	1.44	1.42	1.35
0.5	0.9	0.1	1.09	1.12	1.13	1.21	1.24	1.22
0.5	0.9	0.5	1.10	1.14	1.14	1.23	1.26	1.23
0.5	0.9	0.9	1.12	1.15	1.15	1.25	1.27	1.24
0.3	0.1	0.1	1.08	1.12	1.15	1.21	1.26	1.26
0.3	0.1	0.5	1.22	1.25	1.24	1.37	1.39	1.35
0.3	0.1	0.9	1.53	1.50	1.43	1.73	1.66	1.52
0.3	0.5	0.1	1.16	1.20	1.20	1.31	1.35	1.32
0.3	0.5	0.5	1.24	1.27	1.26	1.41	1.42	1.36
0.3	0.5	0.9	1.36	1.36	1.33	1.53	1.51	1.43
0.3	0.9	0.1	1.22	1.25	1.24	1.38	1.40	1.35
0.3	0.9	0.5	1.24	1.26	1.25	1.40	1.41	1.36
0.3	0.9	0.9	1.25	1.27	1.26	1.41	1.42	1.37
0.1	0.1	0.1	1.12	1.17	1.19	1.27	1.32	1.31
0.1	0.1	0.5	1.23	1.27	1.27	1.39	1.42	1.38
0.1	0.1	0.9	1.54	1.51	1.44	1.73	1.68	1.54
0.1	0.5	0.1	1.21	1.25	1.25	1.38	1.41	1.38
0.1	0.5	0.5	1.28	1.31	1.30	1.45	1.47	1.41
0.1	0.5	0.9	1.39	1.39	1.36	1.57	1.55	1.47
0.1	0.9	0.1	1.29	1.31	1.30	1.47	1.49	1.42
0.1	0.9	0.5	1.30	1.32	1.31	1.48	1.49	1.42
0.1	0.9	0.9	1.31	1.33	1.32	1.50	1.50	1.43

We display the results only for $\beta = 0.5$ because both the simulated and theoretical calculations do not depend significantly on β . The augmented plan has smaller bias and is more efficient than the best SP. The actual efficiency is again somewhat larger than predicted by the asymptotic calculations.

For $r = 3, 4$, we see a very different behavior. In this case, there are a number of type A plans that are (almost) uniformly better than the best standard plans and always better (over our grid of parameter values) than any type B plan. In Table 7, we show the relative efficiencies of a few type A plans compared with the best SP. There are significant improvements possible over the best standard plan. We checked some of these results using a simulation with 10,000 runs. For example, when $r = 3, N = 60$, we compared $A(6, 2, 24)$ to $SP(10, 2)$ over the complete

grid of values for γ, δ , and β . In all cases, there is less bias with the augmented plan and the estimated efficiencies are substantially higher than those predicted by the asymptotic calculations.

We summarize our findings when we allow for the possibility of part-by-operator interaction as follows:

- The best plans have $n = 2$ in the SP component. That is, there is minimal repeated measurement by the same operator on the same part.
- The best plans use few parts in the SP component and a large number of parts in total.
- With $r = 2$ operators, a good plan for estimating γ is $B(2, 2, 28)$ when $N = 64$ or a scaled version for other values of N . For example, if $N = 96$, the scaled version is $B(3, 2, 42)$.

Note that the best plan in this case is close to SP(30,1), so the estimates of σ_{po}^2 and σ_m^2 are highly correlated.

- With $r = 3$ or 4 operators, there are good type A plans, e.g., $A(6, 2, 24)$ for $r = 3$ and $A(5, 2, 24)$ for $r = 4$ with relatively few parts in the SP component. We can realize substantial benefits in estimating γ using one of these plans.
- Simulations suggest that the asymptotic results are conservative. Efficiencies of the good augmented plans are higher than predicted.

We provide MATLAB[®] code at <http://www.bisrg.uwaterloo.ca/> and also as part of the supplementary material for this article at <http://www.asq.org/pub/jqt/>. The code can be used to select and compare good plans using the asymptotic calculations. In any given situation, a practitioner can investigate a wide variety of potential plans and select one that meets his or her needs. At the same websites, we also provide an example and corresponding MATLAB[®] code that will calculate the estimates and their standard errors for either type of augmented plan.

Discussion and Conclusions

The idea of augmented plans raises several design and analysis issues.

In many situations, augmented assessment plans provide a means to estimate γ more efficiently than the best standard plan with the same number of operators and total measurements. One drawback of the type A augmented plans is that we cannot use ANOVA in the analysis because some parts are only measured by one of the operators. With type B augmented plans, we can apply the results of Chapter 7 in Burdick et al. (2005) to get approximate confidence intervals. However, in such an unbalanced design, the properties of the ANOVA-based estimates of γ can be examined only by simulation and so are not useful in the planning stage of the study. The easily calculated Fisher information is a convenient basis for comparison of plans and, with maximum-likelihood estimation, provides a method of analysis. To derive approximate confidence intervals for γ , we suggest using the asymptotic standard deviation as given by the Matlab code (perhaps on a transformed scale to avoid problems near 0 or 1) or other likelihood based methods. We have not explored the properties of such approximate confidence intervals.

In a series of papers, Browne et al. (2009, 2010) consider the use of leveraging to increase the efficiency of standard plans. In these plans, the order of the two components is important. In the first stage of a type A leveraged plan, each operator measures a separate set of parts once. Then a standard plan is carried out using extreme parts selected from those measured in stage 1. Browne et al. do not consider the possibility of a part-by-operator interaction. Note that the leveraged plans use fewer parts, so it is not clear how their performance compares with the augmented plans described here. This is another issue for future investigation.

We compared plans for estimating γ under the constraint that the total number of measurements N is fixed. There may be different costs to measuring one part repeatedly rather than a number of parts once and, in this instance, we should compare plans under the constraint that the total cost is fixed. We have not investigated this issue.

Augmented plans are not sequential. We can carry out the components in any order. Another possibility is to carry out the SP component first and then select an augmented component based on a preliminary analysis of the SP data. Such a design may have superior performance over the augmented plans recommended here.

The results of the simulations were surprising to us. Good plans (as ranked by asymptotic standard deviation of the estimator for γ) were typically close to unbiased and the actual standard deviations in the simulations were larger for both the augmented and corresponding standard plans but the efficiencies of the augmented plans were larger than those predicted by the asymptotic calculations.

The idea of augmentation is to use more parts, consistent with the recommendation of Burdick and Larsen (1997). Typical standard plans with 10 parts do not provide sufficient information to adequately estimate σ_p and hence γ . The plans we recommend all increase the number of parts relative to the AIAG (2003) recommended standard plans and reduce the number of repeated measurements on the same part by each operator. That is, we recommend only plans with $n = 1$ or 2. Also note that the recommended augmented plans are almost uniformly more efficient than the best SP over a wide range of the parameter values. We summarize our recommended plans as follows:

1. For a system with a single operator or with no

operator effects, use the standard plan with $N/2$ parts each measured $n = 2$ times.

2. If you are willing to assume no part-by-operator interaction or if you are willing to confound estimation of the part-by-operator interaction and the measurement repeatability,
 - For $r = 2$ or 3 operators, use the standard plan with $n = 1$ to maximize the number of parts in the study.
 - For $r = 4$ operators, use a type A plan with $n = 2$ and a small number of parts in the SP component. For example, use $A(4, 2, 32)$ if $N = 64$, and a scaled version of this plan for other values of N .
3. If you wish to include the possibility of part-by-operator interaction (and wish to separately estimate σ_{po}^2 and σ_m),
 - For $r = 2$ operators, use a type B plan with a small number of parts in the SP component, e.g., $B(2, 2, (N - 8)/2)$.
 - For more operators, use a type A plan with $n = 2$. For $r = 3$, use a scaled version of $A(6, 2, 24)$, and for $r = 4$, use a scaled version of $A(5, 2, 24)$ with scaling depending on the ratio $N/60$.
4. If you have some knowledge of the possible parameter values γ, δ, β , use the provided software to investigate a number of plans over the restricted range of parameter values.

Acknowledgments

We would like to thank the Editor and two referees who made many valuable suggestions to help improve the paper.

Appendix

In this appendix, we show the derivation of the likelihood result given in Equation (3) and discuss the derivation of the corresponding Fisher information. Here we sketch the derivation and avoid the tedious details by using MAPLE™ (2009) and MATLAB® (2008) to carry out the symbolic and numerical calculations. In model (1), we assumed that measurements made on different parts are independent; thus, we can write the log likelihood and Fisher information for each part and then add over all parts. Consider the distribution of all measurements on a randomly selected part i that is measured by r operators n times each, as in the SP (or the SP component of plan A or B). We order the random variables by operator so that $\vec{Y}_i = (\vec{Y}_{i1}^T, \vec{Y}_{i2}^T, \dots, \vec{Y}_{ir}^T)^T$,

where $\vec{Y}_{ij} = (Y_{ij1}, Y_{ij2}, \dots, Y_{ijn})^T$ corresponds to the n measurements by operator j on part i . We let J_a be a column vector of a 1's, $J_{a \times b}$ be an $a \times b$ matrix of 1's, and I_a be the $a \times a$ identity matrix. From model (1), we have $\vec{Y} \sim MVN(\vec{\mu}, \Sigma)$ with

$$\vec{\mu} = (\vec{\mu}_1, \vec{\mu}_2, \dots, \vec{\mu}_r)^T,$$

where

$$\vec{\mu}_j = \mu_j(1, 1, \dots, 1)^T = \mu_j J_n$$

and

$$\begin{aligned} \Sigma &= \sigma_m^2 I_{rn} + \sigma_{po}^2 \begin{pmatrix} J_{n \times n} & 0 & \cdots & 0 \\ 0 & J_{n \times n} & 0 & \vdots \\ \vdots & 0 & \ddots & 0 \\ 0 & \cdots & 0 & J_{n \times n} \end{pmatrix} \\ &+ \sigma_p^2 J_{rn \times rn} \\ &= \sigma_m^2 I_{rn} + \sigma_{po}^2 I_r \otimes J_{n \times n} + \sigma_p^2 J_{rn \times rn}. \end{aligned}$$

The matrix Σ has a special form that allows us to write its inverse explicitly as $\Sigma^{-1} = b_1 I_{rn} + b_2 I_r \otimes J_{n \times n} + b_3 J_{rn \times rn}$, where the Kronecker product \otimes creates the appropriate block diagonal matrix, and we have

$$\begin{aligned} b_1 &= 1/\sigma_m^2, \\ b_2 &= -\frac{\sigma_{po}^2}{\sigma_m^2(\sigma_m^2 + n\sigma_{po}^2)}, \\ b_3 &= -\frac{\sigma_p^2}{(\sigma_m^2 + n\sigma_{po}^2)(\sigma_m^2 + n\sigma_{po}^2 + rn\sigma_p^2)}. \end{aligned}$$

We can also write

$$\begin{aligned} \det(\Sigma) &= (\sigma_m^2 + n\sigma_{po}^2 + rn\sigma_p^2)(\sigma_m^2 + n\sigma_{po}^2)^{r-1} \\ &\times (\sigma_m^2)^{r(n-1)}. \end{aligned}$$

The log-likelihood contribution from part i with n repeated measurements by r operators is

$$-\frac{1}{2}(\vec{y}_i - \vec{\mu})^T \Sigma^{-1}(\vec{y}_i - \vec{\mu}) - \frac{1}{2} \ln(\det \Sigma)$$

where

$$\begin{aligned} &(\vec{y}_i - \vec{\mu})^T \Sigma^{-1}(\vec{y}_i - \vec{\mu}) \\ &= b_1 \sum_{j=1}^r \sum_{l=1}^n (y_{ijl} - \mu_j)^2 \\ &+ b_2 \sum_{j=1}^r \left[\sum_{l=1}^n (y_{ijl} - \mu_j) \right]^2 \\ &+ b_3 \left[\sum_{j=1}^r \sum_{l=1}^n (y_{ijl} - \mu_j) \right]^2. \end{aligned}$$

Adding over all parts then gives the overall likelihood expression (3).

To calculate the Fisher information, we take second partial derivatives and then expected values of the sum of squares involving the data as given in Equation (6).

$$\begin{aligned}
 E \left[\sum_{i=1}^k \sum_{j=1}^r \sum_{l=1}^n (y_{ijl} - \mu_j)^2 \right] &= k r n (\sigma_p^2 + \sigma_{po}^2 + \sigma_m^2) \\
 E \left[\sum_{i=1}^k \sum_{j=1}^r \left\{ \sum_{l=1}^n (y_{ijl} - \mu_j) \right\}^2 \right] &= k r n (n \sigma_p^2 + n \sigma_{po}^2 + \sigma_m^2) \\
 E \left[\sum_{i=1}^k \left\{ \sum_{j=1}^r \sum_{l=1}^n (y_{ijl} - \mu_j) \right\}^2 \right] &= k r n (n r \sigma_p^2 + n \sigma_{po}^2 + \sigma_m^2), \tag{6}
 \end{aligned}$$

and, for plan A,

$$E \left[\sum_{i=1}^k \sum_{j=1}^r \sum_{l=1}^{k_A/r} (z_{jl} - \mu_j)^2 \right] = k_A (\sigma_p^2 + \sigma_{po}^2 + \sigma_m^2).$$

From the form of the likelihood contributions, because all functions involving the μ_j 's are quadratic,

- The second partial derivative with respect to any of the μ_j 's does not involve the observed measurements and is independent of j .
- The mixed partial derivatives involving μ_j and μ_q are constants not involving the data and are the same for all j and q .
- The expected value of any mixed partial derivative involving only one of the μ_j 's is zero.
- The expected value of any of the second partial derivatives involving $\sigma_p^2, \sigma_{po}^2, \sigma_m^2$ depend only on the expected values of the data-based sums of squares in the likelihoods.

To determine the asymptotic variance (and standard deviation) of the estimator for γ , we substitute the assumed parameter values and invert the matrix numerically with MATLAB® (2008). Finally, because we are interested only in $\gamma, \sigma_p^2, \sigma_{po}^2$, and σ_m^2 , we pre- and postmultiply the information matrix by the matrix of partial derivatives,

$$D = \frac{\partial(\mu_1, \dots, \mu_r, \sigma_p^2, \sigma_{po}^2, \sigma_m^2)}{\partial(\gamma, \sigma_p^2, \sigma_{po}^2, \sigma_m^2)}.$$

Again, we use Maple™ to calculate these partial derivatives and MATLAB® to find their numerical values.

References

AUTOMOTIVE INDUSTRY ACTION GROUP (AIAG). (2003). *Measurement Systems Analysis*, 3rd edition. Southfield, MI: AIAG.

BURDICK, R. K.; BORROR, C. M.; and MONTGOMERY, D. C. (2005). *Design and Analysis of Gauge R&R Studies: Making Decisions with Confidence Intervals in Random and Mixed Effects Models*. ASA-SIAM Series on Statistics and Applied Probability, Philadelphia, PA.

BURDICK, R. K. and LARSEN, G. A. (1997). "Confidence Intervals on Measures of Variability in R&R Studies". *Journal of Quality Technology* 29, pp. 261–273.

BROWNE, R.; MACKAY, R. J.; and STEINER, S. H. (2009a). "Improved Measurement System Assessment for Processes with 100% Inspection". *Journal of Quality Technology* 41, pp. 376–388.

BROWNE, R. P.; MACKAY, R. J.; and STEINER, S. H. (2009b). "Two-Stage Leveraged Measurement System Assessment". *Technometrics* 51, pp. 239–249.

BROWNE, R. P.; STEINER, S. H.; and MACKAY, R. J. (2010). "Leveraged Gauge R&R Studies". *Technometrics*, accepted for publication.

DANILA, O.; STEINER, S. H.; and MACKAY, R. J. (2010). "Assessment of a Binary Measurement System in Current Use". *Journal of Quality Technology* 42, pp. 152–164.

DE MAST, J. and VAN WIERINGEN, W. N. (2010). "Modeling and Evaluating Repeatability and Reproducibility of Ordinal Classifications". *Technometrics* 52, pp. 94–106.

MAPLE 13 (2009). Maplesoft, Waterloo Maple Inc. Waterloo, ON, www.maplesoft.com.

MATLAB 7.7.0 (2008). The MathWorks Inc. Natick, MA, www.mathworks.com.

MINITAB 14.20 (2007). Minitab Inc., State College, PA, www.minitab.com.

SHAININ, P. D. (1992) "Managing SPC, A Critical Quality System Element". *The 46th Annual Quality Congress Proceedings, ASQC*, pp. 251–257.

TRAVER, R. W. (1995). *Manufacturing Solutions for Consistent Quality and Reliability*. New York, NY: American Management Association.

VAN DER MEULEN, F.; DE KONIG, H.; and DE MAST, J. (2009). "Nonrepeatable Gauge R&R Studies Assuming Temporal or Patterned Object Variation". *Journal of Quality Technology* 41, pp. 426–439.

VAN WIERINGEN, W. N. and DE MAST, J. (2008). "Measurement System Analysis for Binary Data". *Technometrics* 50, pp. 468–478.

VARDEMAN, S. B. and Van Valkenberg, E. S. (1999). "Two-Way Random-Effects Analyses and Gauge R&R Studies". *Technometrics* 41, pp. 202–211.

WALTER, S. D.; ELIASZIW, M.; and DONNER, A. (1998). "Sample Size and Optimal Designs for Reliability Studies". *Statistics in Medicine* 17, pp. 101–110.