

A Potential Framework for Integration of Architecture and Methodology to Improve Statistical Production Systems

John L. Eltinge¹, Paul P. Biemer², and Anders Holmberg³

This article outlines a framework for formal description, justification and evaluation in development of architectures for large-scale statistical production systems. Following an introduction of the main components of the framework, we consider four related issues: (1) Use of some simple schematic models for survey quality, cost, risk, and stakeholder utility to outline several groups of questions that may inform decisions on system design and architecture. (2) Integration of system architecture with models for total survey quality (TSQ) and adaptive total design (ATD). (3) Possible use of concepts from the Generic Statistical Business Process Model (GSBPM) and the Generic Statistical Information Model (GSIM). (4) The role of governance processes in the practical implementation of these ideas.

Key words: Adaptive Total Design (ATD); Evolutionary Operation (EVOP); Generic Statistical Business Process Model (GSBPM); paradata; survey quality, cost and risk; stakeholder utility; total survey error.

1. Introduction: Systems Development, Implementation and Maintenance in the Context of Total Survey Design

1.1. Definition and Evaluation of Systems for Statistics Production and Their Architectures

Over the past two decades, large-scale production of official statistics has become increasingly complex. Efforts to manage the resulting business processes and statistical production systems have led to development of a literature on “architecture” for these systems. Ideally, an architecture provides a reference point (used between and within statistical organisations) for efficient coordination of common business processes, information flows, production components, and development decisions. Architecture is considered especially important for reduction of development and maintenance costs and risks, for improvement of data quality, and for management of transitions from old to new production systems. See, for example, [Dunnet \(2007\)](#), [Penneck \(2009\)](#), [Braaksma \(2009\)](#), [Pink et al. \(2010\)](#), [Sundgren \(2010\)](#), [Finselbach \(2011\)](#), and [UNECE \(2011\)](#). In addition, a system perspective has helped some statistical organizations to improve standardization in

¹ Bureau of Labor Statistics, Postal Square Building, 2 Massachusetts Avenue, NE Washington DC, U.S.A. Email: Eltinge.John@bls.gov

² RTI, Research Triangle Park, NC 27709-2194, U.S.A. Email: ppb@rti.org

³ Statistics Sweden, Box 24300, SE-701 89 Örebro, Sweden. Email: anders.holmberg@scb.se

Acknowledgments: This article benefited from many discussions with Mike Hidioglou and several other participants in the 2011 International Total Survey Error Workshop and the Statistics Canada Symposium 2011. In addition, Boris Lorenc provided invaluable contributions to the content and exposition of this article.

order to reduce system complexity, and to identify redundant production systems. See, for example, [Sundgren \(2007\)](#), [Field \(2009\)](#), [Renssen and Van Delden \(2008\)](#), [Renssen et al. \(2009\)](#), [Penneck \(2009\)](#), [Jug \(2009\)](#), [Gloersen and Saeboe \(2009\)](#) and [Studman \(2010\)](#). Work with system architecture has also strengthened international collaborations through the use of common frameworks. See [Vaccari \(2009\)](#), [Todorov \(2011\)](#), [ABS \(2012\)](#), [Hamilton and Tam \(2012a, 2012b\)](#), and [UNECE \(2012a\)](#). [European Commission \(2009\)](#), [Statistics Canada \(2012\)](#), [UNECE \(2011\)](#), [Sundgren \(2011\)](#) and [Munoz-Lopez \(2012\)](#) provide additional background on statistical architecture; and [Sowa and Zachman \(1992\)](#), [CIOC \(1999, 2001\)](#), [Rozanski and Woods \(2005\)](#), [OMB \(2009\)](#) and [The Open Group \(2009\)](#) provide background on architectural frameworks and architecture of systems in general. In addition, the articles in this special issue of the *Journal of Official Statistics* discuss some important aspects of architecture for statistical production systems.

The abovementioned references use somewhat varying definitions of the terms “statistical production system” and “architecture.” For the current article, “statistical production system” will include systems used for sample surveys (for which the statistical organization exercises a high degree of control) and for statistical work with administrative record systems and registers (for which the statistical organization may have a lesser degree of control). In addition, the term “architecture of a statistical production system” (or simply “architecture” for short) will mean an overall description of the system structure (input, output and logic), its components (subsystems), their interrelationships, and the performance characteristics of the components, the integrated system, and system outputs. Inclusion of the performance characteristics will be crucial if the architecture is to be useful to assist in improving a production system. For system output these would typically be an overall account of system capability and ability to deliver what is required by the business owner (ideally the same as the clients/users want). Section 1.2 below considers performance assessments that contribute to the architecture. These assessments simultaneously depend on methodology factors as well as factors of system design and development.

Careful descriptions of systems architecture generally require balanced discussion of features related to management, methodology and information technology. It is important that they contain views from all of these fields and that the descriptions are complementary and compatible. For some purposes, descriptions at a very high conceptual level can be useful, but practical evaluation and implementation generally require a substantial amount of technical detail. These details often involve interactions among heterogeneous subsystems, as well as evaluation of quality, cost, and risk factors over the full life cycle of the system ([Camstra and Renssen 2011](#); [Munoz-Lopez 2012](#)). Important system features include efficiency, reliability, durability, maintainability, vulnerability, flexibility, scalability, adaptability, and extensibility to future systems. These features are not unique to statistical systems; see [MIT \(2004\)](#) and [Rozanski and Woods \(2005\)](#) for related comments in a broader context.

1.2. Improving the Balance of Total Survey Quality, Cost and Risk in the Context of Stakeholder Utility

Because statistical system architecture has developed to improve the balance of quality, cost, and risk in production systems, it is useful to consider this architecture in the context

of total survey design and stakeholder utility. For general background on total survey design, see, for example [Biemer \(2010\)](#), [Groves and Lyberg \(2010\)](#) and references cited therein. For the current development, we note that a “total survey design” approach can provide a useful conceptual framework in which to develop, implement, and evaluate a wide range of methodological features, including work with frames, sampling, instruments, fieldwork, estimation, and inference. In addition, schematic extensions of the total survey design approach may help to provide a stronger and more systematic basis for the integration of methodology with production systems and survey management. This would involve three long-term goals:

- (1) Provide a framework in which to explore the impact that specific features of the system design may have on the balance among multiple performance criteria, including data quality, cost, risk, and stakeholder utility. Each of these performance criteria are multidimensional, and require evaluation over the full lifecycle of a production system, including initial design and development, implementation, continuing operations, maintenance, and periodic upgrades. This exploration may involve direct effects of the system design as such, as well as interactions between the systems design and relevant features of the methodological and administrative design.
- (2) Use the framework in (1) to guide practical decisions related to resource allocation and governance processes for methodological, systems, and administrative design.
- (3) Use the results from (1) and (2) to identify specific areas of research that are most likely to have a substantial impact on work at the interface of methodological, systems, and administrative design.

The current article will focus on schematic extensions related to goal (1), while goals (2) and (3) will be considered in future work. Consider a general survey design characterized by a vector X , where the term “design” includes both formal design decisions and implementation of those decisions. Components of X generally would include all features of standard survey procedures, for example: development and refinement of frames and related administrative record systems; sample designs; collection methods; review, edit, imputation, and estimation methods; and procedures for dissemination of the resulting estimates. Additional components of X include features of the computational systems used to implement the abovementioned survey procedures, and features of administrative processes that may have an effect on survey quality, cost, and risk.

In the development of this article, it will be useful to distinguish among factors associated primarily with methodological (M), systems (S), and administrative (A) features of the design, and we will partition the design vector $X = (X_M, X_S, X_A)$ and related model parameter vectors accordingly. In addition, performance of our survey design may be affected by uncontrolled factors Z . These factors may include general features of the statistical environment (e.g., quality of frame information, accessibility of sample units through prospective collection modes, and the willingness of specific subpopulations to cooperate with survey requests), availability of resources, operational constraints, and changes in stakeholder needs for data. The factors Z also include process data and other forms of metadata and paradata (e.g., data on contact histories, costs, and effort expended at the unit level of effort, and incomplete-data patterns) as discussed in, for example, [Couper \(1998\)](#) and [Kreuter et al. \(2010\)](#).

We emphasize that each of X_M , X_S and X_A may depend heavily on the statistical production system components of the architecture. Also, the sensitivity of these components to changes in Z may depend on the architecture.

We can characterize the performance of the resulting survey design in terms of its quality, cost, risk, and stakeholder utility. In most practical cases, each of these performance measures will be multidimensional. First, let Q represent standard measures of data quality. This might include the components of a standard total survey error model, as well as more qualitative properties like timeliness and relevance, as reviewed by [Brackstone \(1999\)](#) and others. The quality measures may also include criteria developed originally in the computer science and information systems literature, for example, use of generally accepted systems design features, transparency of system architecture and code, or performance of the system with specified test cases.

Second, let C represent the costs associated with the development, implementation, maintenance and operation of the procedures; this includes costs for field operations, systems work, dissemination and methodology. One may consider schematic models for quality and cost, for example,

$$Q = g_Q(X, Z, \beta_Q) \tag{1.2.1}$$

and

$$C = g_C(X, Z, \beta_C) \tag{1.2.2}$$

where $g_Q(\cdot)$ and $g_C(\cdot)$ have (nominally) known functional forms, and β_Q and β_C are parameter vectors. In many cases, some or all components of β_Q and β_C are unknown.

Third, note that standard approaches to survey quality have tended to focus on issues that arise from cumulative effects of a large number of random or quasi-random events, such as separate household-level decisions on whether to respond to a survey request. However, the performance of a statistical production system can be seriously degraded by a small number of discrete events. For example, in many large-scale surveys, managers devote considerable effort to prevention of catastrophic degradation in one or more dimensions of quality. These efforts may include reduction of the probability of a given failure, as well as development of system features that will ameliorate the effects of failures that do take place. As a second example, [Linacre \(2011\)](#) and others have noted that in the development of survey systems, there are substantial risks that the development project might be aborted for a variety of reasons. For example, in many cases, termination took place when it became clear that initial assessments of cost or quality effects were unrealistic.

In principle, one could expand the definition of “survey quality” to include the abovementioned problems. However, since previous literature on survey quality has generally not included these issues, we will instead use a separate definition of *survey risk* R as the combined impact of a substantial and discrete failure in one or more components of a survey process. This would include degradation in survey quality, as well as failures in database management; in system development, testing and maintenance; in confidentiality protection for disseminated data; or in standard electronic data security. Characterization of these components of risk generally will include the severity of the failure, the duration and trajectory of recovery, effects of resource re-allocation arising from the failure, and

the related impact on institutional credibility and on perceived stakeholder utility. Thus, for this discussion, the term “risk” incorporates the probability of a specific failure, its direct and indirect costs, and its broader impact.

As in the case of quality and cost, the survey risk R generally will be multidimensional, and we can consider a model

$$R = g_R(X, Z, \beta_R) \tag{1.2.3}$$

where $g_R(\cdot)$ is a function of general form and β_R is a vector of unknown parameters.

Fourth, let U represent the perceived stakeholder utility of the resulting survey design as implemented, and consider a model

$$U = g_U(Q, C, R, X, Z) \tag{1.2.4}$$

intended to link perceived utility with measures of quality, cost, and risk. In some relatively simple settings, R is perceived to be negligible, $g_U(\cdot)$ is the identity function and $U = Q$. However, in many cases, perceived utility is influenced by components of the design or environmental factors that may not be captured directly through standard quality measures Q . Furthermore, perceptions of utility may be influenced by components of risk and cost. For example, the risk that data releases will not occur on schedule may reduce the perceived value of these releases. Similarly, if cost structures are unpredictable from one year to the next, then data availability in a future year may be problematic, which may in turn reduce the perceived utility of the proposed production system.

In addition, note that Model (1.2.4) includes Q , C , and R (which are functions of X and Z), and also incorporates X and Z directly. For example, let X_1 be the design factor that determines the level of geographical granularity in the publication; this may have a direct effect on the perceived utility of the data release. On the other hand, let X_2 be the design factor that determines sample allocation across strata; this factor may affect U only through the components of Q related to estimator variance.

Fifth, note that each of the performance measures Q , C , R and U will have distinct subvectors associated, respectively, with development and implementation of the design, as well as with ongoing production and maintenance. To simplify the exposition, we will use “performance” as an umbrella term for all of our quality, cost, risk, and utility measures at any point in the cycle of design, implementation, operations, and maintenance, and will use

$$P = g_P(X, Z, \beta_P) \tag{1.2.5}$$

to define the resulting “performance surface,” which is somewhat analogous to a multivariate “response surface” considered in industrial quality control. In addition, we emphasize that there are often complex multivariate relationships among Q , C , R and U , and one could expand Model (1.2.5) to reflect these relationships.

Finally, in keeping with the term “schematic” we emphasize that Models (1.2.1)–(1.2.5) are intended primarily to provide some structure for discussion of the complex trade-offs among cost, quality, risk, and perceived utility encountered in large-scale statistical work, and should not be interpreted in an excessively literal or reductionist form. For example, some components of P may be measured with a reasonable level of precision, while other important components may allow only very general qualitative

assessments. Similarly, for some components of P , relationships with factors X and Z may be reasonably approximated by, for example, a linear or nonlinear regression model, while in other cases the relationships are highly fluid and not reasonably characterized or estimated through standard statistical models with available data.

1.3. *The Combined Effects of Methodological, Systems, and Administrative Components of the Design Vector X*

To address goals (1) through to (3) presented in Section 1.2, consider a relatively simple linear form of the performance model:

$$P = g_P(X, Z, \beta_P) = \beta_{P0} + \beta_{PM}X_M + \beta_{PS}X_S + \beta_{PA}X_A + (\text{interactions}) + e_P \quad (1.3.1)$$

Under Model (1.3.1), β_P contains the intercept vector β_{P0} , the main-effect coefficient vectors β_{PM} , β_{PS} , β_{PA} , and β_{PZ} and the coefficient vectors for the two-factor and higher-order interactions as needed. In addition, the error term e_P is assumed to have a mean of zero and a variance-covariance matrix V_{eP} that may depend on the design variables X and the environmental and process variables Z .

Eltinge (2012) and Eltinge and Phipps (2009) explored variants of Model (1.3.1) for general survey design work. The current article focuses primary attention on the effect that the system design vector X_S has on the mean and variance functions of the performance measures P . Specifically, in the use of Model (1.3.1), three topics of special interest are:

- (a) System features X_S that allow the measurement or (partial) control of specified components of P . Examples include features that facilitate the collection and analysis of data from a contact history instrument (CHI); computer-assisted recording (CARI); cost accounting records; audit trails produced by microdata edit and review processes; or disaggregated reports of interviewer activity linked with survey responses.
- (b) System features X_S that impose constraints on some components of the methodological and administrative design vectors (X_M , X_A), or on the use of information regarding environmental factors Z . For example, use of certain standardized system features may preclude some types of local optimization of a methodological design; or may limit the amount of process data that one may collect on specific measures of cost, quality, or risk. Moreover, system constraints that do not allow for poststratification or response-propensity adjustments would restrict the use of such environmental information in Z for the improvement of survey weights.
- (c) Direct effects of X_S on P , through the main-effect coefficient β_{PS} and related interaction coefficients. For example, some system features may allow for improved editing of collected data; timelier follow-up for nonrespondents; more flexible on-demand production of estimates for special subpopulations; or real-time monitoring of estimates based on data collected to date, as considered in some forms of responsive design. In some cases, one may also consider the effect of X_S on V_{eP} , in keeping with the study of dispersion effects in the response surface methodology literature (e.g., Brennerman and Nair 2001, Bisgaard and Pinho 2003 and references cited therein). However, dispersion effects appear to have received relatively little attention in the survey and systems design literature to date.

The remainder of this article explores issues (a), (b) and (c) through the following steps. Section 2 considers the use of the performance model (1.3.1) to inform design decisions. Section 3 explores the relationship between the general model (1.3.1) and more specific models developed for Total Statistical Quality and Adaptive Total Design. Section 4 links the model, and related design decisions, with the Generic Statistical Business Process Model (GSBPM) and the Generic Statistical Information Model (GSIM). Section 5 outlines related issues with governance processes that link the managerial design features X_A with the system design features X_S .

2. Use of Performance Models to Inform Design Decisions

2.1. General Questions for Exploration of Design Decisions

Ideally, one would have extensive information on all relevant aspects of the performance model (1.3.1), including identification of the applicable predictors X and Z (incorporating both main effects and interactions); sufficiently precise estimates of the coefficients β_p ; and the values of R^2 and other goodness-of-fit measures. Under such conditions, selection of a preferred design X (including the systems design features X_S) is arguably a straightforward exercise in optimization or satisficing, based on the specified multivariate performance criteria.

In most practical cases, however, currently available empirical information falls far short of the abovementioned ideal. Consequently, it is important to apply standard concepts from social measurement and response surface methodology to describe the more limited types of design information that may be feasible to obtain, and that may be useful in practical work with the systems design features X_S . The following questions are of particular interest:

(A) What practical design decisions are under consideration? Examples include:

(A.1) Design of a survey that is mostly or entirely new. For this case, one would seek to understand the extent to which previous experience with somewhat similar surveys may offer partial insights into the performance surface for the proposed new survey. Beyond that limited information, evaluation of the performance surface will depend primarily on new empirical work, such as small-scale tests of system components and of the interfaces of those system components with specific methodological or managerial design components.

(A.2) Comprehensive revision of one or more survey design features, while leaving other design features largely unchanged. Here, it will be important to identify the extent to which design components that are subject to change will have substantial interactions with other factors in the performance surface (1.3.1). If the changed components influence the performance surface only through main effects, then it may be relatively simple to assess the likely performance of the survey under the proposed new design. On the other hand, if the changed components enter Model (1.3.1) through complex interactions, then evaluation of performance under the new design may require extensive amounts of additional empirical work.

(A.3) Incremental revision of one or more design features. In this case, detailed empirical information on the performance surface (1.3.1) under the previous design may

provide important insights into likely performance under the new design. This final case is somewhat analogous to “evolutionary operation” approaches considered in industrial response surface methodology. See, for example, [Box \(1957\)](#) and [Box and Draper \(1969\)](#).

(B) Out of the many potential performance characteristics reviewed in Section 1, which features are of primary interest to the survey organization? Are there practical and objective methods available to measure these performance characteristics? If not, what methodological, systems or management work should one carry out to develop and implement such measures?

(C) Within the context defined by parts (A) and (B):

(C.1) What information is currently available regarding Model (1.3.1), including its functional form, dominant main effects and interactions, and parameters β_p and V_{ep} ?

(C.2) Does the information in (C.1) provide sufficient information to guide the practical decisions identified in (A)? The definition of “sufficient information” will depend heavily on the decisions of primary interest in (A). For example, for the extensive redesign work in cases (A.1) or (A.2), it may suffice to have very rough order-of-magnitude indications of the improvements in quality or cost likely to follow from the proposed changes in X . Conversely, one may consider incremental revisions considered in Case (A.3) as part of a modest effort to obtain moderate improvements in quality or efficiency; and the corresponding definition of “sufficient information” may depend primarily on the variances of available estimators of the coefficient vector β_p . Similar comments apply to cost, quality and risk measures related to the system development process itself.

(C.3) If the answer to (C.2) is no, what additional information is needed, and what methods should be used to obtain that information (e.g., formal experiments, pilot tests, or comparisons with industry benchmarks)?

(C.4) During development, implementation, operation and maintenance of a designed production system, what are the ways in which available information will be used to monitor performance and identify needed mid-course adjustments? This information may include model parameter data from (C.1)–(C.3), the performance measures from (B), and related paradata. In keeping with the cautionary note by [Linacre \(2011\)](#) cited above, it is especially important to use relevant measures of cost, quality and risk incurred during the process of system development and implementation, as well as performance measures that apply to operation and maintenance of the production system after implementation.

(D) Are there external constraints that have an important effect on the design project? These constraints may include standard methodological and managerial constraints, as well as imposition of limits on certain system design features. For example, senior management may mandate the use of specified security software, increased use of standardized system features, or reduced funding available for specified types of design work.

2.2. *Questions Arising from Dynamic Features of Quality, Cost, Risk, and Utility*

One may use Model (1.3.1) to characterize a survey process that is static, in the sense that (after completion of design implementation) the coefficients β_p are essentially fixed, the design vector X is generally not subject to change, and the distribution of the paradata Z is relatively stable. For this reason, some persons who specialize in one area of methodology,

management, or systems may tend to perceive the other areas as readily characterized by a fixed set of “requirements.”

In many practical settings, however, the schematic models presented in Section 1.2 are fundamentally dynamic in nature. For example, some environmental factors Z may vary over time, and thus may have an effect on quality, cost, risk, and stakeholder utility. To take one methodological case, changes in cell telephone and landline telephone use over the past two decades have had a substantial effect on the quality characteristics of standard random digit dialing surveys. In addition, during the recent recession, many stakeholders had increased interest in unemployment rates and other labor-force characteristics, which in turn led to changes in the linkage between perceived utility and standard quality measures like accuracy and timeliness.

Similarly, both survey methodology and systems work are dynamic in nature, so the schematic models for quality, cost, risk, and utility may change over time. Some relatively simple examples of changes include the maturing of specific methodological features (e.g., more sophisticated use of incentives or nonrespondent-conversion techniques); improved testing and robustness of complex skip patterns in CATI or CAPI instruments; or amortization of costs related to methodology, training or systems development and implementation.

Dynamics in the development and maturation of methodology imply that system architecture work often must go beyond specification and satisfaction of static “system requirements.” Conversely, dynamics in systems architecture, and in the needs of data users, imply that methodologists must be open to considering a wide range of performance criteria, some of which may differ substantially from traditional statistical performance criteria like bias and mean squared error.

Consequently, it may be important to supplement questions in areas (A) through (D) with additional questions that are linked directly with the dynamics of the survey process. Some examples include:

(E.1) In keeping with a suggestion by Hidioglou (2011, personal communication), what are appropriate practical criteria to use to determine whether a given methodological or managerial component has reached a level of “maturity” that warrants incorporation into standard production systems?

(E.2) To what extent, and in what ways, should one try to build flexibility into standardized systems to increase the likelihood that they will accommodate further methodological refinements or managerial changes at a relatively low cost?

2.3. *Approximations Arising from Standardization Processes*

As noted in the preceding sections, work with system architecture often leads to discussion of standardization of some or all of the components of a survey process across surveys. Potential benefits of standardization include improvement in some components of data quality (especially those related to failure to execute design features as specified), and reduction of certain components of cost and risk. On the other hand, in some cases standardization may prevent a given survey from making optimal use of survey-specific information on, for example, characteristics of the underlying target population. In that sense, an excessively rigid application of standardization may lead to degradation of some

components of survey performance. Similar issues have been noted for standardization in some areas far removed from survey work. For example, [Hall and Johnson \(2009\)](#) discuss distinctions between environments in which a process may be improved through the application of “art” or “science” respectively. Extending their general ideas to the statistical performance model (1.3.1), one could suggest that extensive use of standardization may be warranted when one can account for most of the variability of cost, quality, risk and utility through some simple predictors X that correspond to highly standardized design features. Conversely, standardization may be problematic when much of the variability of cost, quality, risk and utility are attributable to the uncontrolled environmental factors Z , unobserved remainder terms e_p , or design components X that correspond to highly customized design decisions.

3. Relationship With Models for Total Statistical Quality and Adaptive Total Design

Sections 1 and 2 were developed in the context of general performance measures P . To date, the survey literature has devoted a large amount of attention to some quality measures Q , some attention to cost structure (e.g., [Groves 1989](#)), and relatively little attention to survey risk R and stakeholder utility U . Consequently, this section reviews some previously developed strategies for continually improving the quality of system outputs in real time as the outputs are being generated. In previous literature, this approach has been given several different labels, including active management ([Laflamme et al. 2008](#)), responsive design ([Groves and Heeringa 2006](#)) and adaptive design ([Schouten et al. 2011](#)). It may be of interest to develop similar approaches for C , R and U in future work. As with many cases involving high-dimensional objective functions, a focus on specific subcomponents can lead to suboptimization, but nonetheless may be necessary for tractability.

For the current discussion, we will focus on a strategy known as Adaptive Total Design (ATD), as presented in [Biemer \(2010\)](#). ATD aims to reduce the risks of poor data quality in survey processes by using concepts embodied in the total survey error paradigm. In the notation defined in Section 1, ATD seeks to maximize Q under a constraint on C by selecting design features X so that $Q(X,Z)$ is maximized subject to the constraint $C(X,Z) < C_T$ where C_T is the data collection budget.

Alternatively, the ATD may seek to minimize the total risk, $R(X,Z)$, subject to the same budget constraint. At the present time, ATD has only been applied for the data collection process. See, for example, [Carly-Baxter et al. \(2011\)](#) and [Cunningham et al. \(2012\)](#). In addition, some important features of ATD have been implemented in other data collection work. For example, [Calinescu et al. \(2013\)](#) apply some elements of ATD to control nonresponse and measurement error in the initial design, but their approach does not incorporate quality monitoring and improvement during data collection. [Wagner \(2008\)](#) uses adaptive design to reduce the effect of nonresponse on the estimates during data collection; however, his work does not attempt to control other types of nonsampling errors, such as measurement error or frame error. Further generalizations to other components of the GSBPM framework are conceivable and will be discussed subsequently.

To illustrate the principles of ATD for real-time, quality improvement we begin with a preliminary assessment of Z , say Z_0 , and available budget C_T . We choose an initial design

(or, as in the current case, a data collection approach) X_0 which we believe will produce the optimal quality level Q^* . However, typically, the quality level produced by X_0 will not be optimal because either g_Q and/or g_C are usually misspecified in the initial design or the initial assessment of Z was not current or otherwise did not describe the true, current process environment. Suppose that, during the implementation of the process, new information on Z , say Z_1 , is obtained. Based upon Z_1 , the initial design X_0 may no longer be optimal; that is, the quality level that can be achieved by X_1 is $Q_1 < Q^*$. In order to achieve a quality level that better approximates Q^* , the preliminary design must be altered to some extent, say, to X_1 while still ensuring that the budget, C_T , is not exceeded. As the process continues to completion, there may be new assessments of Z , say, Z_2, Z_3, \dots with consequential design alterations, X_2, X_3, \dots .

ATD relies on one’s ability to obtain updated information on Z during production and perhaps better specifications of g_Q and/or g_C so that $Q(X,Z)$ can be re-estimated, leading to a re-optimization of the design. For example, in the data collection setting, information on interviewer field performance may be continually updated as the interviewers work on their assignments. This information may relate to the cost of interviewing by area, difficulties in contacting and interviewing certain types of sample units, sample unit cooperation as a function of the unit characteristics, and so on. This information and other paradata streaming in from the field constitute Z_k for $k = 1, 2, \dots$.

As the design is altered during implementation to achieve quality levels that approach Q^* , the budget must be continually assessed requiring precise estimates of the cost to completion at each stage. For example, at the start of data collection, the full budget C_T was available. However, by the time that Z_1 is observed, only C_1 of the original budget remains; that is, $C_T - C_1$ has been spent by the time Z_1 is observed. Therefore, in determining what interventions $X_1 - X_0$ are needed to achieve Q^* under Z_1 , the currently available budget, C_1 , must not be exceeded. Thus, the shift from X_0 to X_1 is made that will produce a final level of quality Q_1 where $Q_0 < Q_1 \leq Q^*$ where the costs of X_1 are constrained by C_1 . Note that the difference $X_1 - X_0$, referred to above as design “interventions,” are changes in the original design necessary to achieve the desired level of quality. This process of iteratively re-evaluating Z (i.e., Z_1, Z_2, \dots) and applying interventions to achieve designs X_1, X_2, X_3, \dots such that $Q_1 < Q_2 < Q_3 < \dots \leq Q^*$ while holding total costs fixed at C_T is referred to as Adaptive Total Design (Biemer 2010).

To summarize, ATD begins with an initial design X_0 which is optimal given Z_0 , and then seeks to find a sequence of designs X_1, X_2, \dots, X_k based upon the updated information Z_1, Z_2, \dots and updated functions g_Q and g_C such that $Q(X_k, Z_k) = Q^*_k$, the maximum quality that can be achieved for a total cost of C_T . ATD can be thought of as a form of Evolutionary Operation (EVOP; e.g., Box 1957; Box and Draper 1969 and Hahn 1982) for data collections where the goal is to find the maximum of the performance surface under cost constraints.

Some of the implications of ATD can be summarized as follows:

- a. Paradata feeds the EVOP process by providing updated information on Z and the specifications of g_Q and g_C .
- b. Therefore, the goal of collecting paradata should be to:

- i. best inform those aspects of Z , g_Q , and g_C that are most predictive of Q and C and
 - ii. best capture the dynamic nature of the data collection process.
- c. It is important to monitor these paradata in order to maximize data quality because the difference $Q_k - Q_1$ (i.e., the quality improvement after k ATD cycles) may be quite large.
- d. When the budget is reduced, the reduced budget, say $C_{T1} < C_T$, is entered into the ATD optimization and the design X is reoptimized. This usually results in a lower level of quality that can be achieved.

4. Possible Integration With the Generic Statistical Business Process Model (GSBPM) and the Generic Statistical Information Model (GSIM)

Sections 2 and 3 explored the ideas of Section 1 in the context of general performance models and Adaptive Total Design, respectively. Detailed conceptual and empirical expansion of Section 1 also would require a detailed characterization of survey processes. Over the past decade, several organizations have developed forms of this characterization. See [Dunnet \(2007\)](#), [Renssen and Van Delden \(2008\)](#), [Gloersen and Saeboe \(2009\)](#), and [Pink et al. \(2010\)](#). An effort to standardize these approaches and related terminology through a Common Metadata Framework has led to a Generic Statistical Business Process Model (GSBPM; [Vale 2009](#)) based largely on Statistics New Zealand's business process model.

The GSBPM has quite rapidly become a standard instrument to communicate *between* organizations. This outcome arose from the resemblance of GSBPM to already existing descriptive models, its successful union of terminology for higher level statistical production processes, and good communication and anchoring. The GSBPM was identified by the High-Level Group for Strategic Developments in Business Architecture in Statistics (HLG-BAS) as a key standard for modernizing official statistics ([UNECE 2011](#) and [UNECE 2012a](#)).

It would be premature to forecast specific impacts of the GSBPM, but its shared process terminology has the potential to accelerate and improve development of common areas. For example, it has already been used to pinpoint collaboration areas and to a certain extent it helped in the harmonization of architectures (see [Vaccari 2009](#)). In addition, some authors have suggested that the GSBPM also could be a framework for process quality assessment and improvement ([Todorov 2011](#)) and that it will facilitate the sharing of statistical software and components ([Hamilton and Tam 2012a, 2012b](#)).

An important complement to the GSBPM is the Generic Statistical Information Model (GSIM). GSIM is described as "a reference framework of internationally agreed definitions, attributes and relationships that describe the pieces of information that are used in the production of official statistics (information objects). This framework enables generic descriptions of the definition, management and use of data and metadata throughout the statistical production process" ([UNECE 2012b](#)).

In particular, GSIM aims to define and describe in a harmonized way the *information objects* that flow through and control the statistical production processes. The resulting common framework may help development of shared production systems by improving

communication between distinct disciplines, for example, subject matter statisticians, IT specialists and methodologists. GSIM is also aligned to standards such as DDI (Data Documentation Initiative; <http://www.ddialliance.org/>) and SDMX (Statistical Data and Metadata eXchange; <http://sdmx.org/>), but it does not depend on them.

In the future, the GSBPM and GSIM (or similar frameworks) could be valuable for exploration of the schematic approach outlined in Section 1, and related topics considered in this special issue of the Journal of Official Statistics. Two cross-cutting areas may warrant special attention. First, GSBPM (or similar national models) could lead to standardized and perhaps highly modular production systems and architectures, and GSIM could support standardized handling of information objects in a complex production system. Stated in terms of Models (1.2.1)–(1.2.5), the resulting modularity could have an effect on system performance through both the X_S and Z factors. Some effects may be related to X_S because modularity (more subsystems/components) will require system design with higher attention directed toward the system's life cycle factors. If a system fails when one of its components fails, then increased standardization and reuse of components will increase the importance of managing the full spectrum of risks. This includes risks associated with individual components, integration of these components, and performance of the overall production chain. In addition, if standards and harmonization of information objects lead to more system components that share data, then the resulting production systems may have performance that is less subject to uncontrolled environmental factors Z . An arguably desirable development would be that some factors that today are observable but not controlled (i.e., Z terms) can be replaced by controllable design factors X_M or X_S . One such example could be automation of certain parts of a production system. Hamilton and Tam (2012a, 2012b) and Borowik et al. (2012) provide thoughts about possible future modular production systems where attention to these matters is needed.

Second, one could consider expansions of the GSBPM that would incorporate in-depth collection and use of information on costs and data quality throughout the process chain.

Although quality management is present in the model as an overarching process, it is not reflected specifically in the model phases or subprocesses. For example, as noted above, an important part of the design work is the design of ATD monitoring and improvement components that run simultaneously with, and gather data from, the processes. These components are critical for updating Z and respecifying X to optimize Q in real time; and are implemented later, during data collection. Likewise, our ability to conduct post-survey data quality evaluations requires that components are in place to capture the relevant data on survey errors that aid these evaluations. In addition, systems for quality control are key components of work with initial development, as well as data collection and processing. Quality control components need to be developed and implemented to assure both product and process quality. Finally, documentation on data quality is a critical component of data dissemination and archiving. These are just a few areas of the GSBPM where quality management could be emphasized by providing specific references to subprocesses that address ATD-related improvements to data quality. In this regard, the current version of the GSBPM seems to emphasize statistical standards (i.e., standards related to data production). However, it provides relatively little coverage of quality standards (i.e., standards related to quality improvement) and of quality functions like the measures Q

considered in Section 1 above. The European Code of Practice (available from http://epp.eurostat.ec.europa.eu/portal/page/portal/product_details/publication?p_product_code=KS-32-11-955), which focuses primarily on developing, producing and disseminating European statistics, provides some examples of statistical standards. An example of a quality standard is Eurostat's Quality Assurance Framework (available from http://epp.eurostat.ec.europa.eu/portal/page/portal/quality/documents/QAF_version_1.0_EN.pdf), which focuses primarily on quality assurance and quality management.

5. Governance Processes and Management of Stakeholder Expectations

The development, implementation, testing, operation and maintenance of survey production systems require cooperative work by several distinct groups within a survey organization. These groups tend to have different skill sets, institutional cultures and incentive structures. Consequently, governance processes for architecture and systems work can involve a complex set of factors. In a formal sense, these factors can be viewed as some of the components of the “management design” vector X_M introduced in Section 1. This section presents some of these management factors. Subsection 5.1 outlines some general features of a governance process. Section 5.2 reviews some related features of institutional and professional environments. Throughout this section, we use the term “management” as the omnibus term for all relevant management activities.

5.1. Features of a Governance Process

Roles and responsibilities of management. These include the following:

- Resource allocation, including (a) initial investments; (b) internal pricing for use of production systems (e.g., a free good; partially subsidized use; or full assignment of costs to users; and procedures for amortization of costs, based on anticipated longevity of a given system or component); (c) mid-course adjustments and project termination.
- Acquisition and evaluation of relevant and actionable information regarding the quality, cost and risk factors that are most directly relevant to decisions on all aspects of the survey process. This generally will require management personnel to have extensive training and experience in the relevant technical areas.
- Resolution of disagreements among stakeholders, especially in relation to trade-offs between distinct components of cost, quality and risk. This will require senior personnel to have broad skills for leadership and consensus-building in environments for which ambiguity and incomplete information may be quite prominent.
- Communication and implementation of decisions. Depending on the specific decisions, communication may include a wide range of internal and external stakeholders. Implementation work includes all of the steps required to convert an abstract decision into concrete results, and may be especially challenging when management has limited information on the dominant features of the performance models (1.2.5).
- Personnel management. This includes determination of skills and experience required for specific tasks; full-time or part-time assignment of specific personnel to

these tasks; integration of these assignments with general career paths in the survey organization; and evaluation of individual and group performance for these tasks.

Mechanisms for management. This includes both control of X and the methods used to control specific components of X . Examples of management mechanisms include standard direction through a supervisory chain and management through a committee structure. For committee-based management, clearly defined decision processes may be especially important, and may include decisions by formal committee vote, decisions by formation of a consensus, and decisions by a line manager based on advice from the committee.

Feedback loops for management of specific design components. This includes feedback loops to provide information on (1) adequacy or limitations of models for C , Q , R and U ; (2) operational risk (e.g., identification of cases in which there are substantial deviations from the nominal design during implementation); (3) environmental factors Z with large coefficients for the main effects on Z , or for the interactions between X and Z .

Incentives for individuals or groups. Successful revisions of system architecture generally require that the relevant individuals and groups have an incentive structure that is aligned clearly with the survey organization's need to improve its overall balance of cost, quality and risk, and to overcome institutional inertia that may stand in the way of that improvement. Examples of implicit or explicit incentives include: direct mandates for standardization and systems use; subsidized or free access to systems developed under the revised architecture; allowing the participating groups to re-invest or otherwise control some of the savings obtained through the revised architecture; or linkage of participation with individual promotions or other employment-related benefits. Mechanisms for implementation of these incentives include standard personnel management processes; internal competitions for resources used in systems development, testing and implementation; and allowed flexibility for customization.

5.2. Relevant Features of the Institutional and Professional Environment

In addition to the general features outlined above, the success of a governance process may depend on several qualitative features of the institutional and professional environment, and the extent to which those features are coherent with the performance characteristics of the proposed system.

Expectations regarding objective and measurable criteria for performance. Under a (generally unattainable) ideal, this would include all relevant dimensions of the C , Q , R and U vectors. More realistically, this would include dimensions of these vectors that are considered especially important for key stakeholders.

Professional norms and standards for transparency, credibility, intellectual property and management confidentiality. These norms and standards appear to vary substantially across different survey organizations, and across different professional groups within a survey organization. For example, among professionals in mathematical statistics and methodology, professional credibility generally depends on a very high degree of transparency regarding methods used and the strengths and limitations of those methods. Intellectual property rights generally center on publication copyrights and acknowledgement of priority, rather than formal patent or property rights as such, and

confidentiality restrictions (with respect to management) generally are limited to certain business processes (e.g., bid pricing).

External benchmarking of system features and performance measures. Some organizations expect that most features of a survey process (including the general architecture, methodology, fieldwork, production systems and management processes) will be rigorously compared with relevant external benchmark standards. Other organizations carry out little or no external benchmarking, in some cases due to perceptions that their processes are too specialized to allow for meaningful external comparisons.

Use of objective external peer review processes. In parallel with the discussion on external benchmarking, some survey organizations make extensive use of external peer review processes, while others do not. For survey groups that do use external peer review processes, important issues include the selection and management of the peer review group (e.g., strong standards for expertise and credibility, and procedures to prevent conflicts of interest); the degree to which review recommendations are binding on the survey organization; and specific resource allocation decisions and institutional incentives that are linked with review recommendations. In addition, the effectiveness of an external peer review process will depend heavily on the extent to which the members of the review group have a carefully nuanced understanding of the relevant issues, based on their skills and experience, as well as sufficient access to survey-specific information. Architecture based on extensive sharing of system components may provide a degree of internal peer review for some system components, but is not a substitute for systematic external reviews. See [Lyberg \(2012\)](#) for in-depth discussion of external peer review processes for survey organizations.

6. Closing Remarks

Work with architecture for a statistical production system involves a complex balance among a wide range of performance measures that include cost, quality, risk, and stakeholder utility. Each of these measures may be influenced by design decisions regarding methodological, systems, and management factors, as well as other factors that are not subject to control. Models (1.2.1) to (1.2.5) provide a schematic framework through which one may discuss performance measures and design decisions for statistical production work.

The preceding five articles in this special issue of Journal of Official Statistics provide some snapshots of the current state of work at the interface of methodology with the architecture of statistical production systems. These articles report on recent progress at the abovementioned interface, and also identify areas in which additional work will be warranted. We believe that four general topics will be especially important: more systematic empirical assessment of factors that affect cost, quality, risk, and stakeholder utility; use of these empirical results to expand previously developed approaches to total survey error and total statistical design; integration of the schematic models (1.2.1) to (1.2.5) with the general process descriptions provided by the GSBPM and the GSIM; and careful development and implementation of governance processes that are tuned appropriately for the architecture of statistical production systems. We hope that this

special issue will contribute to a more systematic study of these topics, and look forward with enthusiasm to further developments in this field.

7. References

- Australian Bureau of Statistics (2012). International Collaboration Effort – Statistical Network Progress Report. Paper presented by Brian Pink at the meeting of Directors General of participating countries Australia, New Zealand, Norway, Sweden, Canada, UK in New York, February 2012.
- Biemer, P.P. (2010). Total Survey Error: Design, Implementation and Evaluation. *Public Opinion Quarterly*, 74, 817–848.
- Bisgaard, S. and Pinho, A. (2003). Follow-Up Experiments to Verify Dispersion Effects: Taguchi's Welding Experiment. *Quality Engineering*, 16, 335–343.
- Borowik, J., Henden, M., and Fraser, B. (2012). Riding the Big Data Wave to Streamline Acquiring and Managing Data. Meeting on the Management of Statistical Information Systems (MSIS 2012). Washington, DC, 21–23 May 2012.
- Box, George E.P. (1957). Evolutionary Operation: A Method for Increasing Industrial Productivity, *Journal of the Royal Statistical Society. Series C (Applied Statistics)*. Vol. 6, No. 2 (June 1957), 81–101.
- Box, G.E.P. and Draper, N.R. (1969). *Evolutionary Operation*. New York: Wiley.
- Braaksma, B. (2009). Redesigning a Statistical Institution: the Dutch Case. In *Proceedings of Modernisation of Statistics Production*. International conference organised by Statistics Sweden, Stockholm, 2–4 November 2009. Available at: http://www.scb.se/Grupp/Produkter_Tjanster/Kurser/ModernisationWorkshop/final_papers/H_2_business_models_Braaksma_final.pdf. (accessed February 12, 2013).
- Brackstone, G. (1999). Managing Data Quality in a Statistical Agency. *Survey Methodology*, 25, 139–149.
- Brenneman, W.A. and Nair, V.N. (2001). Methods for Identifying Dispersion Effects in Unreplicated Factorial Experiments: A Critical Analysis and Proposed Strategies. *Technometrics*, 43, 388–405.
- Calinescu, M., Bhulaia, S., and Schouten, B. (2013). Optimal Resource Allocation in Survey Designs. *European Journal of Operational Research*, 21, 115–121.
- Camstra, A. and Renssen, R. (2011). Standard Process Steps Based on Standard Methods as a Part of the Business Architecture. 58th Congress of the International Statistical Institute. Dublin, 21–26 August 2011.
- Carly-Baxter, L., Mitchell, S., and Peytchev, A. (2011). Developing and Implementing Adaptive Total Design. Presented at the 2011 International Field Directors and Technologies Conference (IFD&TC), Scottsdale, AZ. Available at: http://www.ifdetc.org/PC2011/presentation_2011_files/3A-Lisa%20Carley-Baxter.pdf. (accessed February 12, 2013).
- Chief Information Officers Council CIOC (1999). *Federal Enterprise Architecture Framework*. Version 1.1, Chief Information Officers Council, September 1999.
- Chief Information Officers Council CIOC (2001). *A Practical Guide to Federal Enterprise Architecture*. Version 1.0, Chief Information Officers Council, February 2001.

- Couper, M. (1998). Measuring Survey Quality in a CASIC Environment. In: Proceedings of the Section on Survey Research Methods, American Statistical Association.
- Cunningham-Hunter, D., Mitchell, S., Carley-Baxter, L., and Keating, M. (2012). Using Paradata to Monitor Survey Production, Cost, and Quality within an Adaptive Total Design Framework. Paper presented at the 2012 Federal Conference on Survey Methodology, Washington, DC. Available at: http://www.fcs.gov/12papers/Hunter_2012FCSM_V-C.pdf. (accessed February 12, 2013).
- Dunnet, G. (2007). The BmTS: Creating a new business model for a national statistical office of the 21st century. Invited paper at the UNECE/Eurostat/OECD Meeting on the Management of Statistical Information Systems (MSIS). Geneva, 8–10 May 2007. Available at: <http://www.unece.org/fileadmin/DAM/stats/documents/ece/ces/ge.50/2007/mtg1/wp.11.e.pdf>. (accessed February 13, 2013).
- Eltinge, J.L. (2012). Doing More with Less: Balancing Survey Cost, Quality, Risk and Stakeholder Utility. Presentation at the Fourth International Conference on Establishment Surveys, Montreal, Canada.
- Eltinge, J.L. and Phipps, P.A. (2009). Characterization, Evaluation and Management of Prospective Benefits, Costs and Risks in the Development of New Statistical Programs for Energy. Proceedings of the Section on Survey Research Methods, Joint Statistical Meetings.
- European Commission (2009). Communication from the Commission to the European Parliament and the Council on the production method of EU statistics: a vision for the next decade, Document COM(2009)404. Available at: <http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=COM:2009:0404:FIN:EN:PDF>. (accessed February 12, 2013).
- Field, S. (2009). From Red to Green: The role of Enterprise Architecture in the ONS Corporate IT Strategy. UNECE/Eurostat/OECD Meeting on the Management of Statistical Information Systems (MSIS 2009). Oslo, Norway, 18–20 May 2009.
- Finselbach, H. (2011). Implementing Standardised Systems, Processes, Tools and Methods with the ONS Design Charter. 58th Congress of the International Statistical Institute. Dublin, 21–26 August 2011.
- Gloersen, R. and Saeboe, V. (2009). Standardisation for Improvements in Statistics Norway. In Proceedings of Modernisation of Statistics Production International conference organised by Statistics Sweden, Stockholm, 2–4 November 2009. Available at: http://www.scb.se/Grupp/Produkter_Tjanster/Kurser/ModernisationWorkshop/final_papers/H_1_business_models_Gloersen.pdf. (accessed February 12, 2013).
- Groves, R.M. (1989). Survey Errors and Survey Costs. New York: John Wiley & Sons.
- Groves, R.M. and Heeringa, S. (2006). Responsive Design for Household Surveys: Tools for Actively Controlling Survey Errors and Costs. *Journal of the Royal Statistical Society Series A: Statistics in Society*, 169(Part 3), 439–457.
- Groves, R.M. and Lyberg, L. (2010). Total Survey Error: Past, Present and Future. *Public Opinion Quarterly*, 74, 849–879.
- Hahn, G.J. (1982). Statistical Assessment of a Process Change. *Journal of Quality Technology*, 14, 1–9.
- Hall, J.M. and Johnson, M.E. (2009). When Should a Process be Art, not Science. *Harvard Business Review*, March 2009 issue, 58–64.

- Hamilton, A. and Tam, S.-M. (2012a). Toward GSIM V1.0 as a Cornerstone for Common Reference Architecture. Meeting on the Management of Statistical Information Systems (MSIS 2012). Washington, DC, 21–23 May 2012.
- Hamilton, A. and Tam, S.-M. (2012b). Platform for International Collaboration – Preliminary Thoughts From the ABS. Meeting on the Management of Statistical Information Systems (MSIS 2012). Washington, DC, 21–23 May 2012.
- Jug, M. (2009). Uptake of Service Oriented Architecture in Statistics – Are We Really So Different? In Proceedings of Modernisation of Statistics Production International conference organised by Statistics Sweden, Stockholm, 2–4 November 2009. Available at: http://www.scb.se/Grupp/Produkter_Tjanster/Kurser/ModernisationWorkshop/final_papers/C_1_SOA_Jug_final.pdf. (accessed February 12, 2013).
- Kreuter, F., Olson, K., Wagner, J., Yan, T., Ezzati-Rice, T.M., Casas-Cordero, C., Lemay, M., Peytchev, A., Groves, R.M., and Raghunathan, T.E. (2010). Using Proxy Measures and Other Correlates of Survey Outcomes to Adjust for Nonresponse: Examples from Multiple Surveys. *Journal of the Royal Statistical Society, Series A*, 173, 389–407.
- Lafamme, F., Pasture, T., Talon, J., Maydan, M., and Miller, A. (2008). Using Paradata to Actively Manage Data Collection, Proceedings of the ASA Survey Methods Research Section, Denver, Colorado.
- Linacre, S. (2011). Standardising Methods and Tools: A Look Back Over 30-Plus Years of Experience. Paper presented at Statistics Canada Symposium 2011.
- Lyberg, L. (2012). Survey Quality. Waksberg Memorial Lecture delivered at the Fourth International Conference on Establishment Surveys, 14 June 2012.
- MIT (2004). Engineering Systems Division Architecture Committee: The Influence of Architecture in Engineering Systems. Paper presented at the Engineering Systems Symposium, MIT, Cambridge MA, USA, 29–31 March 2004. Available at: <http://esd.mit.edu/symposium/monograph/default.htm>. (accessed February 12, 2013).
- Munoz-Lopez, J. (2012). Value of Systems Architecture for Statistical Organizations. Meeting on the Management of Statistical Information Systems (MSIS 2012). Washington, DC, 21–23 May 2012.
- OMB (2009). Improving Agency Performance Using Information and Information Technology EA Assessment Framework v3.1, Office of Management and Budget. June 2009.
- Penneck, S. (2009). The Office for National Statistics (ONS) Statistical Modernisation Programme: What Went Right? What Went Wrong? In Proceedings of Modernisation of Statistics Production, International conference organised by Statistics Sweden, Stockholm, 2–4 November 2009. Available at: http://www.scb.se/Grupp/Produkter_Tjanster/Kurser/ModernisationWorkshop/final_papers/D_1_management_Penneck_final.pdf. (accessed February 12, 2013).
- Pink, B., Borowik, J., and Lee, G. (2010). The Case for an International Statistical Innovation Program – Transforming National and International Statistics Systems. *Statistical Journal of the International Association for Official Statistics (IAOS)*, 26, Number 3–4. Available at: <http://iospress.metapress.com/content/2h5764574t6318r4/fulltext.html>. (accessed February 12, 2013).

- Renssen, R. and Van Delden, A. (2008). Standardisation of Design and Production of Statistics; A Service Oriented Approach at Statistics Netherlands. Paper presented at the IAOS Congress Reshaping Official Statistics, Shanghai, October 2008.
- Renssen, R., Wings, J., and Paulussen, R. (2009). Processes, Methods and Tools. Statistics Netherlands internal report. Available at: <http://www.cbs.nl/NR/rdonlyres/8EFDD386-0FD9-4567-90BC-6ADAB5C67B3D/0/2009Processesmethodsandtoolsart.pdf>. (accessed February 12, 2013.).
- Rozanski, N. and Woods, E. (2005). Software Systems Architecture. New Jersey, United States of America: Addison Wesley.
- Schouten, B., Calinescu, M., and Luiten, A. (2011). Optimizing Quality of Response Through Adaptive Survey Designs, Statistics Netherlands, Available at: <http://www.cbs.nl/NR/rdonlyres/2D62BF4A-6783-4AC4-8E45-12EF20C6675C/0/2011x1018.pdf>. (accessed February 12, 2013).
- Sowa, J.F. and Zachman, J. (1992). Extending and Formalizing the Framework for Information Systems Architecture, IBM Systems Journal, 31.
- Statistics Canada (2012). Corporate Business Plan – Statistics Canada – 2012–13 to 2014–15. Available at: http://www.statcan.gc.ca/about-apercu/business-plan-affaires_2012-2015-eng.htm. (accessed February 12, 2013).
- Studman, B. (2010). A Collaborative Development Approach to Agile Statistical Processing Architecture – Australian Bureau of Statistics (ABS) Experience and Aspirations. UNECE/Eurostat/OECD Meeting on the Management of Statistical Information Systems (MSIS 2010). Daejeon, Republic of Korea, 26–29 April 2010.
- Sundgren, B. (2007). Process Reengineering at Statistics Sweden. From UNECE/Eurostat /OECD Meeting on the Management of Statistical Information Systems (MSIS). Geneva, 8–10 May 2007. Available at: <http://www.unece.org/fileadmin/DAM/stats/documents/ece/ces/ge.50/2007/mtg1/wp.22.e.pdf>. (accessed February 13, 2013).
- Sundgren, B. (2010). The Systems Approach to Official Statistics. In Official Statistics Methodology and Applications in Honour of Daniel Thorburn. Stockholm University, Stockholm. 225–260. Available at: <http://officialstatistics.files.wordpress.com/2010/05/bok18.pdf>. (accessed February 12, 2013).
- Sundgren, B. (2011). Towards a System of Official Statistics Based on a Coherent Combination of Data Sources, Including Surveys and Administrative Data. Available at: <https://sites.google.com/site/bosundgren/my-life/Towardsasystemofofficialstatisticsbasedonacoherentcombinationofdatasources.doc?attredirects=0>. (accessed February 12, 2013).
- The Open Group (2009). The Open Group Architecture Framework (TOGAF), Version 9. The Open Group: San Francisco, CA. Available from: <http://www.togaf.info/>. (accessed February 12, 2013).
- Todorov, V. (2011). Cooperation Models for Software Development. Meeting on the Management of Statistical Information Systems (MSIS 2011). Luxembourg, 23–25 May 2011.
- United Nations Economic Commission for Europe, UNECE (2011). Strategic Vision of the High-level Group for Strategic Developments in Business Architecture in Statistics. ECE/CES/2011/1. Economic Commission for Europe, Conference for European Statisticians, Geneva, 14–16 June 2011.

- United Nations Economic Commission for Europe, UNECE (2012a). Strategy to implement the vision of the High-level Group for Strategic Developments in Business Architecture in Statistics. ECE/CES/2012/10. Economic Commission for Europe, Conference for European Statisticians, Paris 6–8 June 2012.
- United Nations Economic Commission for Europe, UNECE (2012b). The Generic Statistical Information Model (GSIM): Version 0.8. Available at: <http://www1.unece.org/stat/platform/display/metis/GSIM+v0.8>. (accessed February 12, 2013).
- Vaccari, C. (2009). Common Reference Architecture (CORA) ESSnet: history and next activities, paper presented at the Meeting on the Management of Statistical Information Systems (MSIS 2009). Oslo, Norway, 18–20 May 2009. Available at: <http://www.unece.org/stats/documents/2009.05.msis.htm>. (accessed February 12, 2013).
- Vale, S. (2009). The Generic Statistical Business Process Model, Version 4.0. Available at: <http://www1.unece.org/stat/platform/display/metis/The+Generic+Statistical+Business+Process+Model>. (accessed February 12, 2013).
- Wagner, J. (2008). Adaptive Survey Design to Reduce Nonresponse Bias. Unpublished Ph.D. dissertation, University of Michigan, Ann Arbor, MI.

Received May 2012

Revised November 2012

Accepted November 2012