Discussion

*David Dolson*¹

Amongst other things National Statistical Offices (NSO) and their coordinating bodies strive for information production strategies that are fit for their intended uses, cost efficient and can control respondent burden to an acceptable level. A common theme is survey integration ([High-Level Group for the Modernisation of Official Statistics 2014](#); [Priest, G. 2010](#); and [United Nations Statistics Division and the Statistical Office of the European Union 2015](#)), which in addition has the desirable effect of leading towards improved coherence of statistical information across survey programs and between countries. These objectives are often characterized in terms of standardized production processes, commonly accepted concepts, and harmonized survey questionnaire content.

[Reis \(2013\)](#) describes a strategy being taken for the improved integration of European social surveys. The essence of the approach is an integrated system of surveys based upon a set of standardized modules of survey questions which can be combined as needed into a set of different survey instruments to meet different needs in terms of frequency and precision of estimates.

The current article by Ioannidis et al. addresses a general approach to survey integration for social surveys conducted by a NSO. The situation described by Reis would be a specific instance of this challenge. The article describes how, compared to the status quo of more stove piped approaches, such a strategy has the potential to reduce both respondent burden and costs while continuing to produce estimates to meet precision requirements. The main focus of the article is development of a solution to the difficult dual optimization problem of 1) determining the survey modules to be included in each instrument and 2) the related sample sizes given various constraints on the modules and instruments (such as required precision of estimates, required joint observations of modules and periodicities of modules and instruments).

Their solution is an appealing and powerful if perhaps somewhat complex (for me!) method using simulated annealing and the simplex algorithm. It provides advances relative to split questionnaire design approaches that make it a superior solution for this survey integration challenge. Its power and flexibility provide a valuable method for use by any organization contemplating the kind of large scale survey integration considered here and in [Reis \(2013\)](#).

However, as the authors quite appropriately point out, an initiative towards this kind of integration across an entire program of social surveys would comprise several other major elements.

¹ Statistics Canada, 100 Tunney's Pasture Driveway, Ottawa, Ontario, Canada K1A 0T6. Email: david.dolson@canada.ca

Although not currently contemplating this kind of large scale survey integration, Statistics Canada has made significant headway on several of the fronts that would be precursors. The unifying strategic direction is use of solutions that are globally optimal rather than locally optimal for each individual survey program. All of these approaches are directed towards reduced cost and risk associated with development, maintenance and use of multiple survey methods and systems as well as improved coherence of the information produced by the social survey program ([Statistics Canada 2005](#)). Through the development and use of generalized methods and systems for data collection, data processing, sampling, estimation and dissemination, risks and costs associated with using and maintaining multiple systems are reduced and better controlled ([Brisebois and Dufour 2015](#)).

Through use of our Address Register, we have implemented a common household survey frame service ([MacNabb et al. 2011](#)). It provides efficiencies in survey development and operation and allows for both negative and positive, as appropriate, sample coordination strategies to manage respondent burden while also facilitating sample overlap between surveys as needed.

In an initiative started several years ago, harmonized content modules with standardized concepts and question wording have been developed for use in paper questionnaires and for electronic questionnaires ([Nadwodny and Best 2011](#)). This work continues and has been particularly long and challenging as various surveys, each according to its particular redesign cycle, have adapted to possibly slightly different concepts, changed questionnaire wording and possible mode effect changes associated with updated data collection modalities. Creating common questionnaire modules takes time and careful testing to ensure acceptable usability across the surveys that will use them.

An Integrated Collection Operations System ([Statistics Canada 2012](#)) has been implemented that provides standardized questionnaire development techniques using the standardized modules, a consistent look and feel for all electronic questionnaires and common data collection operation and management tools.

In a related development, social and household surveys now use a common social survey processing environment that implements consistent methods and systems for steps such as data capture, editing, coding and imputation ([Nadwody and Best 2011](#)). Over many years a suite of generalized systems ([Deguire et al. 2011](#)), applicable for both household (social) and business surveys, has been developed, implementing commonly used statistical methods for sample design, edit and imputation, estimation, disclosure avoidance and tabulation.

Standardized methods, systems and strategies are to be used and strong justification is required for any exceptions.

In undertaking this at Statistics Canada, surveys have been able to continue to produce data of similar accuracy as previously with the benefit of improved coherence of statistical products. Use of generalized strategies and systems has been successful in reducing risks associated with multiple survey specific systems. Financially though, important investments must be made up front in harmonization of concepts and questionnaire modules as well as development of generalized systems. Even without the kind of large scale integration contemplated in this article, implementing these generalized methods and tools in a program of surveys is a substantial undertaking over many years with the effect that expected savings are only gradually realized.

Summarizing, strategies and tools such I have just outlined are useful and valuable in their own right to meet the strategic goals I noted above, providing important benefits for a statistical office or set of such offices. They are also an important element in moving towards improved and expanded survey integration, including helping to provide a flexible framework to adapt to changing needs in the social survey program. The development of the methods described in the current article provides a useful addition to the set of methods available in moving forward beyond what I have outlined towards effective large scale survey integration. Although powerful, the technique is also a complex one and the additional challenges in the kind of large scale integration being considered are substantial. It will be interesting to observe statistical offices in their assessment of the benefits of large scale integration and using this technique and their choices in whether and how best to proceed.

References

- Brisebois, F. and J. Dufour. 2015. "Enquêtes santé au Canada : S'adapter aux nouvelles réalités de la société canadienne." Paper presented at the Journées de Méthodologie Statistique de l'INSEE, Paris, France, March 2015. Available at: http://jms.insee.fr/files/documents/2015/S13_5_ACTE_V2_BRISEBOIS_JMS2015.PDF (accessed March 29, 2016).
- Deguire, Y., L. Reedman, and M. Wenzowski. 2011. "Generalized Systems: The Statistics Canada Experience." Proceedings of Statistics Canada Methodology Symposium 2011.
- High-Level Group for the Modernisation of Official Statistics. 2014. "Strategic Vision of the HLG." Available at: <http://www1.unece.org/stat/platform/display/hlgbas/Strategic+vision+of+the+HLG> (accessed March 29, 2016).
- MacNabb, L., M. St-Pierre, and M. Grenier. 2011. "Development of a Common Frame for Household Surveys at Statistics Canada." Proceedings of Statistics Canada Methodology Symposium 2011.
- Priest, G. 2010. "The Struggle for Integration and Harmonization of Social Statistics in a Statistical Agency – A Case Study of Statistics Canada." International Household Survey Network working paper No. 4. Available at: <http://www.ihsn.org/home/integrating> (accessed April, 2016).
- Nadwodny, R. and P. Best. 2011. "Harmonized Content and Common Processing Tools: The New Paradigm in Developing Surveys at Statistics Canada." In Proceedings of Statistics Canada Methodology Symposium 2011.
- Reis, F. 2013. "Links Between Centralisation of Data Collection and Survey Integration in the Context of the Industrialisation of Statistical Production." Working paper presented at the UNECE Seminar on Statistical Data Collection, Geneva, Switzerland. Available at: <http://www.unece.org/fileadmin/DAM/stats/documents/ece/ces/ge.44/2013/mgt1/WP2.pdf> (accessed April, 2016).
- Statistics Canada. 2005. "New Household Survey Strategy: Summary Report." Internal Working Paper.
- Statistics Canada. 2012. "Integrated Collection and Operations System (ICOS) Summary Presentation." Internal presentation from Collection Planning and Research Division.

United Nations Statistics Division and the Statistical Office of the European Union. 2015. "Proceedings of the Global Conference on a Transformative Agenda for Official Statistics: Outcomes and Summaries of Sessions." Background document from the Global Conference on a Transformative Agenda for Official Statistics, New York, USA. Available at: <http://unstats.un.org/unsd/nationalaccount/workshops/2015/NewYork/lod.asp> (accessed March 29, 2016).