20th International Congress on Modelling and Simulation, Adelaide, Australia, 1–6 December 2013
www.mssanz.org.au/modsim2013

A Framework for Determining the Validation of Analytical Campaigns in Defence Experimentation

Fred DJ Bowden and Peter B Williams

Defence Science and Technology Organisation
Email: fred.bowden@dsto.defence.gov.au

Abstract: The problem of looking at modernisation for the Australian Army is that it is wicked in its nature. These problems are generally insoluble and the purpose of analysing them is to better understand them and allow more informed decisions to be made. The recognition of this requires an Analytical Campaign approach to these problems. In these campaigns a balance is sought between three aspects of analytical validity, within resource constraints.

The first two aspects of analytical validity are the conflicting aspects of Internal and External Validity. These are addressed through an iterative campaign that moves across the spectrum applying methods at different levels of each to try and understand the impact. Within Army modernisation methods with high External Validity are used to identify potential areas that will impact Army modernisation. These are then transitioned to methods with high Internal Validity to try and understand cause and effect within these areas before returning to high External Validity methods to validate and contextualise the results.

The third paradigm is that of Philosophical Validity. Generally Army modernisation problems have many ways they can be viewed and have multiple competing stakeholders it is important to try and capture all relevant perspectives to the problem. Within the Army modernisation program Philosphical validity is viewed in terms of the five types of triangulation; Data, Investigator, Theory, Method and Environment.

The aim of this paper is to provide a framework for viewing the overall analytical validity of Analytical Campaigns. The paper then uses two examples to show the balancing of Internal and External Validity and provide a newly introduced framework for understanding Philosophical Validity.

Keywords: *Defence experimentation, analytic validity, multi-disciplinary, trans-disciplinary*

1. INTRODUCTION

The process of evolving the Australian Army spans problems such as: major systems procurement; changing structures; and developing future concepts as well as the how they are sensibly brought together. This is further complicated by any number of potential adversaries who might endeavour to adapt in order to exploit weaknesses in each these three aspects. The types of problems involved in Army modernisation are thus inherently complex and generally poorly defined. There can be no definitive formulation of the problems themselves. Solutions cannot be determined to be correct or incorrect, simply 'good' or 'bad'. Each problem is essentially unique. And each problem can be viewed as a symptom of another, larger, problem.

That is to say, evolving the Army is achieved via a series of *wicked* problems (Rittel and Webber (1973)). As such, they cannot effectively be solved by the application of a single method.

For each problem – be it capability acquisition or concept exploration – determining a cause-and-effect is the analysts' panacea. However, such a definite answer is worthless if it is not relatable or applicable to the real world. But the more realistic the problem is made, the harder it becomes to measure the impact of variables (such as the capabilities in question).

How does the analyst ensure a rigorous answer and maintain relevance to the real world? The solution is to balance the conflicting paradigms of Internal Validity and External Validity, whilst maximising Philosophical Validity, via a carefully designed Analytical Campaign. This paper describes these paradigms, and provides exemplars of how in the past they have been balanced to ensure valid results can be achieved.

2. ANALYTICAL VALIDITY

In 2006, The Technical Cooperation Program published the Guide in Experimentation (GUIDEx) (TTCP (2006)) as a Code of Best Practice to conduct military experimentation. This document has become the standard by which experimentation is judged within the defence domain. Shortly following this, The Logic of Warfighting Experiments (Kass (2006)) was published by one of the authors of the GUIDEx. Together with the GUIDEx these two documents outline a process to try and address many of the issue facing Military experimentation. These two documents focus around the conflicting ideas of Internal Validity and External Validity. Although this paper considers the broader question of Analytical Campaigns these ideas from the experimentation literature form a good starting point for discussion of this problem.

The complex nature of the problems faced by Army Modernisation mean that it is not sufficient to just consider internal and external validity. Thus, the idea of Philosophical Validity has been introduced. This has been done to highlight the need to consider not just the empirical aspect of the problem but also the human. It provides a measure of the degree of completeness of the analysis behind the individual components and the Analytical Campaign overall.

Construct Validity (Shadish *et al.*, 2002) is key in bringing together an analytical campaign. However, it rests outside the scope of this paper as it is assumed here that the analyst team will ensure Construct Validity as part of good practice.

2.1. Internal Validity

In this paper, Internal Validity is defined as the ability to identify cause-and-effect within a problem solution. The GUIDEx refers to three aspects to Internal Validity: the ability to use the tested capability; the ability to detect change; and the ability to isolate the reason for change. Here it is only the final aspect that is considered to be Internal Validity. The first two relate to good analytical design which clearly impacts the ability to isolate cause and effect, but are considered by the authors not to relate directly to Internal Validity.

Mathematical models are the pinnacle of Internal Validity, providing an explicit relationship between cause and effect. A long-standing example for combat models is the Lanchester equations (Taylor (1983)). However, these models have little External Validity as they have been abstracted from reality making their relationship to it very remote. This does not mean they have no use, just that there is a need for careful consideration when applying the results of these models into the real world as shown in Pincombe and Pincombe (2007) and Pincombe *et al.* (2009).

2.2. External Validity

External Validity is defined as the ability to relate the results of an analytical campaign to the real world. This varies slightly from that provided in the GUIDEx, where results are only applied to operational scenarios, but holds its overall intent. It covers the analysis of real events both current as well as historical operations. As

these events encompass the full range of variables within a complete context, results from these studies can give the analyst great confidence that they will be applicable when fully realised. They provide the optimal External Validity to a solution.

However, in general it is extremely difficult to establish Internal Validity from these studies due to the complexity and non-linearities involved. It is the role of historians to attempt to establish Internal Validity by studying aspects of these events and trying to establish cause and effect, such as the work done by Rowland *et al.* (1996), Speight *et al.* (1997) and Speight and Rowland (1999).

2.3. The Internal-External Validity Spectrum

There exists a natural conflict between Internal and External Validity. One paradigm seeks to abstract problems into well-defined constructs and isolate dependent variables. The other considers all possible factors, up to and including the human element and random, uncontrollable events. . In essence, Internal Validity can be thought of as adopting a positivist approach, whereas External Validity prefers a rationalistic mentality.

However, the paradigms are not mutually exclusive sets, but rather form a spectrum where problems – and solution methods – can lie at any point within it. Figure 1 represents varying levels along this spectrum with example methods that can be used to solve a wicked problem as part of Army modernisation.

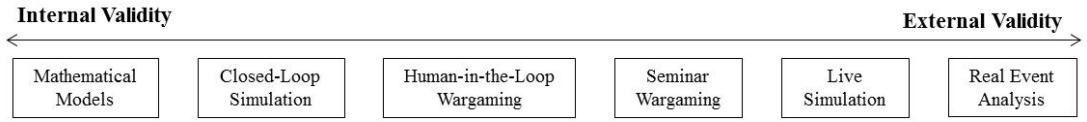

Figure 1: Example analytical techniques on the Internal-External Validity spectrum[1]

2.4. Philosophical Validity

In conjunction with Internal and External Validity, Philosophical Validity provides a way of viewing how completely a problem has been explored/considered. Complete Philosophical Validity is considered to be achieved if all possible perspectives are accounted for within the solution. By definition of a wicked problem this cannot be achieved, however, to gain an understanding of the level of Philosophical Validity of an analytical campaign the five types of triangulation are used; Data, Investigator, Theory, Methodological and Environmental (Guion *et al.* (2011) and Ismail (2013)). It should be noted that in the context of Philosophical Validity, triangulation has been used to try and ensure completeness of information as outlined in Patton (2002) and Adami and Kiger (2005) rather than as confirmatory technique as originally described by Webb (1966) and other such as Kelle (2004).

Data Triangulation is defined as the use of multiple sources of data. These sources may include different stakeholders or subject matter experts. In relation to Army Modernisation this relates largely to ensuring the participants represent all aspects and domains of the problem being considered.

Investigator Triangulation comes from having multiple investigators (analysts), from the same field, use the same approach to study the data. This provides multiple ways of viewing the same data to ensure the outcomes any one analyst considers is justifiable. If it is not possible to have all the data analysis by different analysts, the aim is to provide some degree of overlap in the analysis as was done in the thematic analysis as described in Pincombe *et al.* (2012).

Theory Triangulation uses multiple perspectives to interpret the same data. In this, analysts from different fields with different constructs are used to view the same problem. This is in essence a multi-disciplinary team approach to the problem.

Methodological Triangulation comes from applying multiple methods to the same problem in an attempt to better understand it. This is often discussed in terms of the use of both qualitative and quantitative methods. This aspect of triangulation has strong ties with issues related to Internal and External Validity.

Environmental Triangulation ensures that different settings are considered when addressing a problem. This may include factors such as terrain, adversary force, population and political climate. The aim is to ensure that all key possible options are explored. Generally, Army Modernisation will take a Morphological analysis approach to this aspect of Philosophical Validity as can be seen in Williams and Bowden (2013). As a basis

[1] It should be notes that this representation is not necessarily linear.

from determining the key Environmental Triangulation characteristics the dimensions of future warfare as defined by Pincombe *et al.* (2012) are used. Initially all the dimensions and components as defined in this study are considered. Then those not relevant to the study at hard are removed.

3. ANALYTICAL CAMPAIGNS

It has been noted that the wicked problems inherent in evolving Army's capabilities and concepts require a delicate balance between Internal and External Validity, and to try to maximise the Philosophical Validity. The only way to adequately incorporate all of these requirements is through the use of Analytical Campaigns. Within a single Analytical Campaign, a variety of different methods from different places on the Internal-External Validity spectrum are brought together in an iterative fashion. This is an extension of the battle lab process as defined in Bowley and Lovaszy (1999). This process has been modified to include the need to not just refine the model but also to adjust the Internal/External Validity of each approach In addition to this, the methods used, engagement of participants and stakeholders, and careful construction of the analytical team need to be considered in order to maximise Philosophical Validity.

The process in general, as adopted by the authors, is to begin an Analytical Campaign near the External Validity end of the spectrum. This ensures that as the campaign applies methods of greater Internal Validity the right characteristics of the problem are considered. Then an iterative series of activities traverses to the left; by systematically removing the least important variables and abstracting the problem to focus on those identified as being key. It is preferable to begin with the analysis of real events, however, in looking at possible futures it is not always possible to do this so often human-in-the-loop discussions, seminars and wargaming are used. As the understand of the problem is better understood and the abstracted version of the problem is 'solved', these solutions are then injected into models with greater External Validity to ensure the solutions are applicable in the greater context of the problem. The ultimate check for this is to be able to relate them to the analysis of real events. As a minimum they are critically analysed by Subject Matter Experts to ensure face validity (TTCP, 2006). Depending on the outcomes of this there may be a need to then return to methods with greater internal validity.

At each step in this process, the Philosophical Validity is examined to ensure that the overall campaign minimises any chance of missing any perspectives of the problem. The type of Philosophical Validity that is relevant will depend on the progress of the Analytical Campaign; for example, whilst running a seminar wargame, it is essential to maximise Data and Environmental Triangulation whereas in interpreting results from a mathematical model it is more important to focus on Theory Triangulation.

4. CASE STUDIES

In this section two case studies are used to show how the principles outlined in Sections 2 and 3 can be applied to analyse wicked problems from the Defence domain. The first of these related to a large campaign stretching over a year with considerable resources used to ensure validity in all three forms. The second is a smaller campaign that shows even within a smaller campaign it is possible to reach a high level of validity.

4.1. Offensive Support Requirements Study

In 2002, the Deputy Chief of Army requested a study be made into the requirements for Offensive Support systems. The aim was to minimise the number of calibres, whilst maximising the effects generated by them. Ten options were created from permutations of different mortars and artillery systems.

The 2003 Army Modernisation year was dedicated to this problem, with a series of activities culminating in data analysis. Figure 2 shows the major activities and their position on the Internal/External Validity spectrum.

The Philosophical Validity of this Analytical Campaign can be considered by looking at the five types of triangulation:

- **Data triangulation** was achieved by ensuring participants of the preliminary activities and HEADLINE were representative of Army as a whole. This involved including representatives from the full range of corps in the wargaming and discussions. A similar process was used for the Closed-Loop Simulation step.
- **Investigator triangulation** and **theoretical triangulation** were ensured via a multi-disciplinary analyst team. From the core group of five analysts, two were from an Operations Research background, two were Systems Engineers and one Historian. The analytical process and outputs of the results at each level was analysed by multiple people before conclusions were drawn.

- **Methodological triangulation** was achieved not only in using methods across the full Internal-External Validity spectrum, but also within each level multiple methods were used. For example, a large number of seminar wargames were used in the scoping phases, and the mathematical modelling itself contained a range of disparate methods to cover the dimensions of effects generation, responsiveness, sustainment, maintenance, survivability and deployability.
- **Environmental triangulation** was included in the analytical campaign to a limited extent. The analysis focused on a single Brigade plan in a single terrain set. However, within this context, four very different fire missions were selected to which the entire analysis was applied.

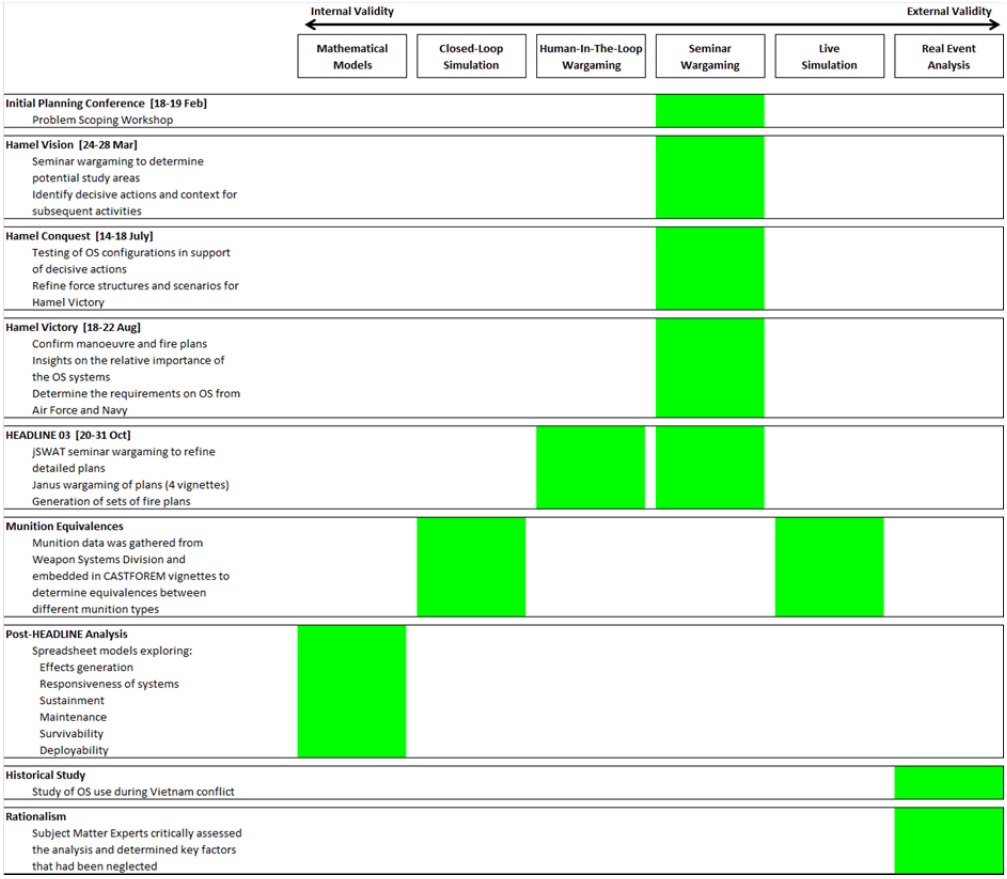

Figure 2: Analytical techniques used for Offensive Fires study

The Offensive Support study Analytical Campaign spanned the complete Internal-External Validity spectrum and had good levels of Philosophical Validity. The greatest weakness was in the use of only limited environmental and theory triangulation. The results produced by this Analytical Campaign were used to alter several projects in the Defence Capability Plan.

4.2. Reconnaissance by Fires Concept

In this study there was a need to analyse the ability of different force structures to conduct a future concept by which reconnaissance would be conducted by fires. For the purposes of this case study, this concept can be simplified to be expressed as the following three steps being repeated on the battlespace:

1. Manoeuvre to force the enemy above the discrimination threshold[2],
2. Hold the enemy above the discrimination threshold,
3. Use joint fires and organic firepower to neutralise the enemy.

An operation or mission using the Reconnaissance by Fires concept ends when the enemy either drops below the discrimination threshold or is neutralised.

This study involved three analytical techniques, discrete Markov chains (Norris (1977)), CASTFOREM simulation (TRADOC (2005)) and jSWAT wargaming (Millikan et al. (2005) and Menadue et al. (2009)).

[2] The discrimination threshold can be thought of as the event horizon at which an enemy element can be identified by EXFOR.

Within the Internal External validity spectrum this related to the use of mathematical models, closed-loop simulation and seminar wargaming as shown in Figure 3. This particular Analytical Campaign was aimed at addressing a set of very constrained questions. Of particular interest was, how many missions each force structure was able to complete. Thus, as the campaign moved from External Validity towards Internal Validity as it tried to identify cause and effect of what was observed within the seminar wargame. Once the initial results were determined these were presented back to a forum of SMEs. This forum was used to confirm External Validity and allowed the next iteration of the process.

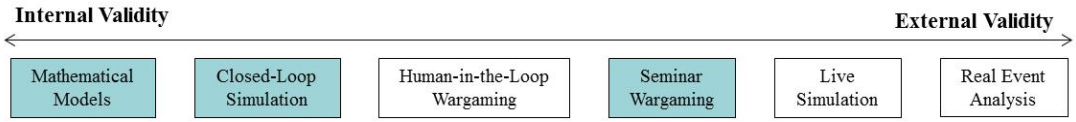

Figure 3: Analytical techniques used for Reconnaissance by Fires concept study

The Philosophical Validity of this Analytical Campaign can be considered by looking at the five types of triangulation:

- **Data triangulation** was achieved by ensuring participants of the Seminar Wargame were representative of Army as a whole. The initial activity included two independent syndicates. Additional SMEs viewed the results of the seminar wargame providing additional detail where required. Similar processes were used for the Closed-Loop Simulation step.
- Multiple analysts were used throughout the analysis process to generate **investigator triangulation**. A team of analysts were used to analyse the data from the seminar wargame with overlap occurring in terms of scenarios, force options wargamed and phases of operation. The team that built the closed-loop simulation model built with each checking aspects of each other work. Outputs of the results at each level were analysed by multiple people before conclusions were drawn.
- The analytical team was mainly made up of natural scientist and one historian. To help counter this weakness the results were presented to Army SMEs at various stages of the process SMEs. In addition to this the construct of the analysis was presented to various SMEs prior to and post the initial step in the Analytical Campaign. So although the analytical team composition limited the theoretical validity this external exposure provided confidence that the campaign did provide **theoretical triangulation**.
- In addition to the methods provided as part of the Internal/External Validity treatment **methodological triangulation** was achieved via the application of a wide variety of methods. These included the use concept papers, structured planning, interviews and AARs during the seminar wargaming process. In developing closed-loop simulations further SME input was sought to confirm the finding of the earlier data capture.
- The concept was tested within three terrain types and the transition between these terrain types as part of the seminar wargame. This provided a degree of initial **environmental triangulation** in this dimension. However, due to resource constraints only the most dangerous of these was taken into the later part of the Analytical Campaign. So the final results have limited validity in this sense.

Overall this the reconnaissance by fires concept Analytical Campaign had strong Internal Validity, good External Validity and reasonable Philosophical Validity. The greatest weakness was in the use of only limited environmental triangulation. This placed limited validity in terms of both External and Philosophical Validity. Further limitations in Philosophical Validity were due to the reduced theoretical validity.

5. CONCULSIONS

This paper has presented the paradigm used in Army Modernisation to determine the validity of its analytical campaigns. This comes down to balancing Internal and External Validity within an iterative process. In addition to this, the idea of Philosophical Validity has been described. This aspect of the framework provides a key addition ensuring the breadth of the wicked problems analysed are taken into account. To judge the level of Philosophical Validity within an analytical campaign the paper puts forward the five types of triangulation: Data, Investigator, Theory, Methodological and Environmental. These serve as guides to ensure the analysts generates the highest level of confidence in addressing a wicked (*i.e.* insoluble) problem.

It should be the aim of the analyst to maximise Philosophical Validity for any study, however it becomes difficult and expensive to attempt to increase all types simultaneously. What is left as the focus of future work is to determine the optimal utility of increasing Philosophical Validity in each of the above areas. That

is, to determine both the minimum triangulation required to add strength to the Analytical Campaign, and at which point adding resources yields decreasing gains.

6. ACKNOWLEDGEMENTS

We acknowledge the inputs into the ideas expressed in this paper from the following people Dean Bowley, Mike Brennan and Paul Gaertner.

7. REFERENCES

Adami, M.F. and A. Kiger (2005). The use of triangulation for completeness purposes, Nurse Researcher, 12(4), 19-29.

Bowden, F.D.J., L. Finlay, D. Lohmeyer and C. Stanford (2009). Multi-Method Approach to Future Army Sub-Concept Analysis

Bowley, D. and S. Lovaszy (1999). Use of Combat Simulations and Wargames in Analytical Studies, Proceedings of SIMTect.

Flood, R.L. and M.C. Jackson (1991), Creative Problem Solving: Total Systems Intervention, John Wiley and Sons.

Guion, L.A., D. C. Diehl, and D. McDonald (2011), Triangulation: Establishing the Validity of Qualitative Studies, FCS6014, one of a series of the Department of Family, Youth and Community Sciences, Florida Cooperative Extension Service, Institute of Food and Agricultural Sciences, University of Florida. Original publication date September 2002. Revised August 2011. Original written by Lisa A. Guion, former faculty member; revised by David C. Diehl and Debra McDonald. Visit the EDIS website at http://edis.ifas.ufl.edu.

Ismail, S.A.M.M. (2013), Triangulation, Presentation to postgraduate students, School of Educational Studies, Universiti Sains Malaysia, can be accessed at http://usmpersila.weebly.com/uploads/1/7/6/5/17653075/triangulation.pdf.

Kelle, U. (2004) Computer Assisted Analysis of Qualitative Data, in A Companion to Qualitative

Research, SAGE: London. p. 276-283.

Menadue I., D. Lohmeyer, S. James and L. Holden (2009). jSWAT2 - The Application of Simulation to Support Seminar Wargaming, Submitted to SimTecT 2009

Millikan J., M. Brennan and P. Gaertner (2005). Joint Seminar Wargame Adjudication Tool (jSWAT), Proceedings of the Land Warfare Conference 2005, Commonwealth of Australia, pp. 201-214.

Norris, J.R. (1977). Markov Chains, Cambridge Series in Statistical and Probabilistic Mathematics, Cambridge University Press.

Patton, M. (2002). Qualitative research and evaluation methods, 3rd ed. Thousand Oaks, CA: Sage Publications.

Pincombe, A.H. and B.M. Pincombe (2007). Tractable approximations to multistage decisions in air defence scenarios, ANZIAM Journal, 49 (EMAC2007), pp. C273-C288.

Pincombe, A.H., B.M. Pincombe and C.E.M Pearce (2009). Putting the art before the force, ANZIAM Journal, 51 (EMAC2009), pp. C482-C496.

Pincombe, B., S. Blunden, A. Pincombe and P. Dexter (2012). Ascertaining a Hierarchy of Dimensions from Time-Poor Experts: linking tactical vignettes to strategic scenarios", Technological Forecasting and Social Change, http://dx.doi.org/10.1016/j.techfore.2012.05.001.

Shadish, W.R., T.D. Cook, D.T. Campbell (2002). Experimental and Quasi-experimental Designs for Generalized Causal Inference, Wadsworth, Cengage Learning.

Speight, L.R., D. Rowland and M.C. Keys (1997). Manoeuvre Warfare: Force Balance in Relation To Other Factors and to Operational Success, Military Operations Research, 3 (3), pp. 31-46.

Speight, L.R. and D. Rowland (1999). Modelling the Mobile Land Battle: Combat Degradation and Criteria for Defeat, Military Operations Research, 4 (3), pp. 45-62.

Rowland, D., L.R. Speight, L.R., and M.C. Keys (1996). Manoeuvre Warfare: Force Balance in Relation To Other Factors and to Operational Success, Military Operations Research, 2 (3), pp. 5-16.

Talyor J.G. (1983). Lanchester Models of Warfare, Vol 1, ORSA, Arlington Virginia, 1983.

TRADOC Analysis Center (2005). Combined Arms and Support Task Force Evaluation Model (CASTFOREM) Update: Executive Summary, TRAC-WSMR-TD-05-023, Department of the Army, USA.

The Technical Cooperation Program (TTCP) (2006). Guide for Understanding and Implementing Defence Experimentation (GUIDEx).

Webb, E.J., D.T. Campbell, R.D. Schwartz, and L. Sechrest (1966). Unobtrusive Measures: Nonreactive Research in the Social Sciences, Rand McNally.