Good data require good field work

A discussion paper on the professional disinterest in transparent, methodologically rigorous quantitative data collection

Roger Henke
Centre for Advanced Study (CAS)
PO Box 2030, # 85, Road 141, Sangkat Veal Vong Khan 7 Makara, Phnom Penh, Cambodia
Tel. +855-23-214494
www.cascambodia.org
rhenke@forum.org.kh

Paper for the sixth socio-cultural research congress on Cambodia, November 18-20, 2003.
Royal University of Phnom Penh

Table of contents

Abstract

This essay addresses data gathering for quantitative analysis in Cambodia. The core argument is that both research practitioners and those commissioning and paying for survey research, from NGOs collecting baseline information for local level programs to international organizations funding nationally representative studies, *undervalue* data collection. The paper compares quantitative data collection in Cambodia and highly educated countries, spelling out similarities and differences. Various reasons for the much higher appreciation of air-con conceptual design and analysis as compared to sweaty field work are offered. Problems caused by this unbalanced perspective on research are described. And institutional strategies to counter this imbalance are suggested.

Introduction

Cambodia is regularly described as a relatively data-rich country. A decade of aid inflow has been accompanied by substantial investments in data collection, environmental data, demographic and health data, socio-economic data and public opinion data.

The relationship between the two is understandable. The bulk of aid money is coming from countries or organizations that have what one may call an evidence-based policy tradition. That is to say policies and practice interventions are normally legitimized by research evidence on their necessity, effectiveness and efficiency. And their implementation is supposed to be accompanied by a monitoring and evaluation framework.

As such, that is certainly laudable. Unfortunately, the emphasis on the importance of a proper and continuously updated evidence base does not go along with a similar emphasis on the *quality* of that evidence base.

It is important to stress that this problem is not specific to Cambodia. But it does have particular local aspects.

In this essay quantitative[1] data collection in Cambodia and highly educated countries is compared, spelling out similarities and differences. Various reasons for the much higher appreciation of air-con conceptual design and analysis as compared to sweaty field work are offered. Problems caused by this unbalanced perspective on research are described. And institutional strategies to counter this imbalance are suggested.

Undervaluing data collection, a universal phenomenon

Although the research cycle story does not in any way imply a hierarchy of importance between the various phases – in the opposite, it implies that the result can only be as good as the weakest link in the chain, and within that, data quality is singled out or recognized as the most likely weakest link, as epitomized by the *garbage in, garbage* out maxim – in actual fact, research professionals perceive a hierarchy. Desk analysis is valued higher than field work. Theorizing, conceptual design, data analysis and telling the story, are seen as the hart of research, actually getting (good quality) data is seen as a necessary but rather mundane part. It is repetitive, often outsourced[2] to less qualified[3] others, and it generates unwieldy kinds of material that are useable only once turned into numbers.

Some indications for the disregard of data collection

Obviously it is impossible to prove a blunt assertion like this. I cannot do more than provide some facts that make it plausible.

---

[1] This paper actually addresses the 'easy' part of the research enterprise. Assessing quality of *qualitative data* is much more difficult.

[2] The expression is used both because this part of a project *is* often outsourced, i.e. subcontracted *and* because, when it is not subcontracted in the technical sense of the term, it is normally executed by research 'partners'; but the term partner suggests much too much real participation in the whole project. Partners are often treated as if hired by the *real* researchers for nothing else but work in the field.

[3] I.e. less well, or at least less specifically educated. However, for field work as such these others are often *better qualified* because of substantial practical experience.

The core instrument of quality control for research, *peer review*, never extends to the actual data collection. It will cover design and methodology as they are described on paper, but not the actual data collection process. As with mountain climbers, the word of researchers on data collection is to be believed. Sure, there is a very practical reason for this exception[4], but all the rest being open to scrutiny, it does mean that there are no external incentives to ensure field work quality.

Also, in as far as *public criticism* – as opposed to the internal criticism of peer review – is concerned, the research profession, hardly ever exposes bad work produced by one of its own. True, no profession lives up to the we-keep-our-own-stables-clean rhetoric that professions use to fend off outside interference into what they consider their own affairs. But it is telling that *if* someone is criticized, it is bound to be for conclusions, *not* for data collection.

If one considers the actual *instrument* used in the field as part of the field work, and that makes sense because turning generic or draft instruments into their final versions is a normal part of outsourced 'field work', it is telling that peer review hardly ever addresses the instrument. There are *no practical constraints* involved here. Yes, it would require more time of the reviewer, but not much more, provided s/he is qualified[5] to assess survey instruments.

However, most reviewers are not. I offer some speculations *why* this is the case.

A major reason is that *methodology* – including the nitty-gritty of how to phrase questions and how to optimize the flow of the whole instrument – has been 'outsourced' to a discipline of its own. Most social scientists only receive limited general training on this crucial subject, certainly a lot less than the time they are required to spend on digesting theory. From a division of labor perspective this might make good sense *if* the methodology discipline were a popular one. If anything however, the opposite is the case: methodology is a *fringe specialization*, more akin to philosophy of science than to economics, sociology or anthropology. And it shows in the number of practitioners: only a fraction of those working in 'mainstream' social science.

It has always struck me as very odd that what *distinguishes social science* from other kinds of information[6], *transparent methodologically rigorous data gathering*, seems of no major concern to the research community itself. Usually professions jealously guard control of their distinctive mark. E.g. the medical profession has not relegated the production of its evidence-base, clinical trials, to a fringe discipline. On the contrary, the amazingly swift acceptance of Evidence Based Practice (EBP) shows an overvaluing of the scientific foundation.

A second reason for reviewers not being qualified is that *assessing instruments requires a lot of knowledge of the context* that the instrument is used within, (sub-)cultural knowledge, socio-cultural knowledge, etc. And that part of the knowledge spectrum has been delegated to

[4] One cannot do it post hoc. If field work monitoring would be a normal part of research quality control and/or a research funding condition it would involve considerable cost, at least compared to the minimal costs involved in current post hoc review procedures.
[5] Here, 'qualified' refers to both formal education and practical experience.
[6] E.g. 'expert opinion', 'investigative journalism', 'trend-watching', etc.

anthropologists. So unless one happens to be a national of the study country or an area specialist one should be careful in judging the appropriateness of phrasings etc.

The only real fence-sitters between the methodology discipline and the mainstream are *hard-core survey researchers* for whom survey research is the only thing they do. Most countries have at least one institute generating the kind of national socio-economic and social indicators that accountable policy-makers need. But more survey researchers are probably employed by contract research institutes and market research agencies. It s telling that in the academic pecking order these 'data-crunchers' are looked down upon as unimaginative.

It is equally telling that the hard-core itself, although often very sophisticated in instrument design, will normally outsource actual data collection to interviewers paid by the interview[7].

Researchers, funders and users alike

To summarize, I argue that the *social science research community* shows a relative disregard for data collection.

This disregard is made all the more easy because it is *shared by funders and users of research*. The latter are interested in analytic conclusions, not the raw data underlying them. Data collection is a black box to them that is preferably available on short notice at minimal cost. So those commissioning and paying for survey research, especially if they are also end-users of the results, often apply time schedules and grant budgets that are a clear invitation if not an order to cut corners.

Sure, it is tempting to accept the invitation. The spoils are direct, be it personal or institutional income, career advance, or the potential to influence real world decisions. The risks are minimal because those paying are not in the game of questioning data quality per se[8]. They are interested in summary conclusions, recommendations, generalizations, best practice lessons, and see themselves as paying for intelligent analysis. The raw material underlying it is seen as process not content.

But in the end, it is up to the research professionals to indicate what it takes to do their work properly. If anything, that is the essence of a profession: being in charge of a particular expertise[9]. So the asserted undervaluing of data collection is ultimately to *blame on the research community itself*. Therefore, it is up to the research community to start ringing the alarm bell. Other stakeholders, users and funders of research, can make a significant difference (as I will show below). But, similar to the EBP movement in the medical world, an effective process of change is most likely to occur if it is initiated from within the research community itself.

---

[7] Obviously, in many countries most surveys are not face-to-face interviews but mail or telephone surveys.

[8] Data are questioned if they do not fit pre-conceptions and/or are politically inconvenient. But then questioning is near certain to target data quality – that is the nature of the game – and the party being questioned is near certain to loose in one way or the other. Research, like politics, is in the end a trade of the possible, and the possible is hardly ever the perfect. There is always methodological and practical fault to be found in a data collection process. And even apart from that, questions sow doubt, the if-there-is-smoke-there-must-be-fire phenomenon.

[9] I am cutting corners here myself. There is a lot more to professionalism/zation.

Garbage in, garbage out, Cambodia is no exception

If *undervaluing field work* did not *affect its quality* this essay were superfluous. Unfortunately, that is not the case. As in most areas and walks of life, something less valued is bound to be sub-optimally prepared, receive less supervisory attention than necessary, gets away with more sloppyness regarding the application of protocols, and tends to be considered not worth the trouble of independent controls.

In Cambodia, this is as true as anywhere else. In a way, *Cambodia's situation is advantaged* in the sense that data quality or more accurately, the lack of it, is quite openly discussed in various sectors (socio-economic data, e.g. on poverty incidence, environmental data, e.g. on forest cover, etc.) and regarding data sources at different levels (both nationally representative data and local level, usually NGO produced data).

Because the earlier mentioned maxim *garbage in, garbage out* is all too true, this makes the underlying cause, the disregard for field work a worrisome phenomenon. All the more worrisome, because also in a country like Cambodia where data quality is at least an issue of stakeholder concern, concern does not seem to translate into more attention to and regard for the data collection process.

Similarities and differences between the Cambodian situation and other countries

The described undervaluing of data collection is as evident in Cambodia as anywhere else. Similarly, quality of Cambodian data are as dependent upon the quality of the instrument used as anywhere[10].

However, surveys being face-to-face interviews, Cambodian studies stay clear from the problem of non-participation/response associated with mail and telephone surveys. Refusal rates can be limited to numbers that do not seriously affect the representativenss of one's sample (which is not to claim that all studies actually succeed in doing so!). In that sense the survey instrument potentially still lives up to its major promise, generating genuinely representative data, a promise that is long lost in many well-educated countries.

But Cambodia is not so advantaged in other respects.

The majority of Cambodian respondents have just a couple of years of formal schooling at the most, are often not familiar with the kinds of questions that survey instruments contain, or feel not confident or embarrassed to disclose their background, knowledge or opinion to outsiders. These are different but at least as difficult problems as the survey fatigue often encountered in well-educated settings. In other words, *doing field work properly is anything but easy in Cambodia*. However, much speaks against these difficulties being dealt with seriously in the Cambodian context.

[10] Slight variations in the phrasing of questions can generate substantially different answer patterns, asking the same question in a positive or a negative way also produces different results, question order and context are important determinants of answers, all of this applies in Cambodia as anywhere else.

Most survey research is *managed by a non-Cambodian principal investigator*, who does not know Khmer, and often even does not know much about Cambodia. This leaves much of the actual data collection, including the finalization of the instrument, effectively beyond supervision.

Much research that is done by *NGOs* is designed and executed by *non-qualified researchers*. The organizations don't have the funds to outsource their research, or, erroneously, look upon research as something that anyone with some higher education can do.

Even when instrument design, etc. is properly done, the actual data collection is often delegated to *teams of inexperienced enumerators, controlled by uninterested supervisors,* sent to the field after just a few days of training. The enumerators lack basic understanding of the objectives of the study, of the way each individual question relates and/or contributes to these objectives, they lack knowledge of how to deal with problems not foreseen in the training protocol, they display inappropriate and counterproductive behavior in the field, and, in general have no idea of the importance proper field work for the validity of the project as a whole. Supervisors equally lack understanding of the importance of what they are asked to control, if they actually accompany their team to the field that is. Both enumerators and supervisors are often recruited in-house, from the NGO or government department or agency involved in the project. If per diems are interesting the job might be a coveted one, if not failure is build in from the start. But even with financial interest guaranteed real commitment, proper behavior, rigorously following protocols, etc. is far from standard.

Benchmarking research

Every trade has its internal critics and whatever I have described above might easily be discounted as 'sounds convincing, but if so few perceive the world as he does, it can't be all that bad'. One might also assume that research is not that different from other sectors and does not stand out as a candidate for necessary change. Or one may shrug off the kind of observations made by arguing that this state of affairs is indeed less than satisfactory but the world is not perfect and arguing for change on this particular issue is just beyond realism.

Therefore I propose to look at the status of field work in social science research through the lens of a very different sector. A benchmarking[11] exercise as it were because I am going to compare the treatment of field work with the treatment of a 'similar'[12] issue in another sector. Above I have described transparent methodologically rigorous data gathering as the hallmark of research and the lack of transparency or accountability of what actually happens in the field as the worrisome result of the relative disregard for field work. Therefore, I propose to define *accountability* as the problem that is to be solved with respect to field work and identify a sector that has to deal with just that same problem.

[11] Normally 'benchmarking' is understood as comparing one's operation with the best practice example or lead organization within one's own sector in order to assess one's performance, determine one's strengths and weaknesses and figure out strategies for improvement.

[12] For benchmarking across sectors to make sense, identifying a sector that has to 'solve' a 'similar' problem as one's own is a crucial first step.

If accountability is at stake, why not look at a discipline for which it is core business: *accounting*. Public and private organizations alike are required to keep accounts. The yearly balance sheet is the result of this requirement. For private enterprise one could even argue that it is the 'result' of the organization tout court. This result is known to be an analytic construction. *Internal accountants* have a toolbox of definitions, and judicial and financial instruments to shape this result in one or the other form. But unlike mountain climbers and researchers, internal accountants are not to be believed upon their word. A large and thriving business sector of *external auditors* has grown around the unquestioned necessity to independently control the work of internal accountants.

How does this external auditing compare with peer review? Well, it doesn't, or rather it is so much more comprehensive that it is no comparison. External auditors spend most of their time on what peer reviewers stay away from: *checking the quality of the raw data* ('you tell me you spent X on consumables in April? Please show me the original bills').

Apart from the plain checking of raw data against reported summaries, the external auditor will also *advise* the client *on adjustments of administrative organization* that improve in-build accountability. Standard advise will either stress creating parallel procedures that generate the same information[13], or having two individuals with independent responsibilities involved in one procedure[14], and preferably both combined.

It is important to stress that, although external audits are accepted as for the common good, they are also *enforced* by a legal and regulatory framework. Or maybe it is better said that *because* they are accepted as for the common good they are enforced. Why this is so is explained by the core tenet of institutional economics: markets need institutional frameworks. Institutions create the guarantees for actors to have the necessary trust in each other and in their environment. External audits are one of those guarantees ensuring that organizations that follow the rules will not loose out to cheaters and free-riders. Obviously, it is only a guarantee if it is enforced. So enforcement is perceived as for the common good.

What this benchmark shows is that in another sector – accounting – the problem of accountability:

- is taken seriously enough ,
- by all actors involved,
- to both implement internal accountability measures,
- and external accountability measures,
- seriously enough that is to make enforcement acceptable

As opposed to social science research where accountability of the data collection process:

- is not taken seriously enough,
- by anyone involved,
- to implement internal accountability measures,
- let alone external accountability measures,

[13] E.g. in a restaurant having the kitchen record all food orders so as to have an independent check on the restaurant bills.

[14] E.g. in a restaurant having one person as cashier, and another to check cash against bills at the end of the day.

- and the idea of enforcing whatever is totally outside the current mindset.

It shows that *the way accountability is dealt with* by social science stakeholders is *sector specific*, that it therefore does not have to be that way, and that suggestions for change are therefore not necessarily pointless.

Changing the incentive structure for research practice

After all these pages on problems, what about solutions? The easy way out is to take a moral stance. So let's get that out of the way.

Yes, all stakeholders *should* take data collection more serious. Yes, this does imply that the major internal quality control mechanism, peer review, *should* include an assessment of the survey instrument used. Yes this again implies that reviewers need to be qualified to do so, and yes that means that methodological expertise *should* be a much more integral part of what it takes to be a social scientist, and that it should certainly not be relegated to the fringes. Yes, the research community *should* be much more willing to publicly disown bad research. And yes, maybe most important of all, data collection is too important to outsource to anyone; enumerators and supervisors *should* be integral members of a project, fully aware of the objectives, down to the level of each individual survey question[15]. And the principal investigator should feel fully responsible for everything that happens in the field, and not stay out of the picture until the cleaned data set is available.

However true all of these appeals and injunctions may be, preaching on correct attitude and behavior is not a very effective instrument of change. Changing attitudes is far from easy, changing behavior is even more difficult.

Thinking in terms of *incentives*, sticks and carrots, is always a good strategy for hitting upon possible change interventions. But, unless one is a hard core behaviorist, this is only really promising as long as the targeted change is not against the grain of current common sense. Else risks of sabotage, evasion and other obstructions are as plausible as the preferred change.

Fortunately, accountable methodologically rigorous data gathering is the distinguishing mark of social science research. This means that *ideologically*, and in the self image of the profession, it is valued. In other words, the basic condition for sticks and carrots to work is definitely fulfilled. The core argument of this essay is that *in practice* it is undervalued. I have tried to show that nothing in the current incentive structure for the research community actively supports valuing proper data collection or actively punishes disregard for it.

The benchmarking exercise showed that one major difference between the world of accounting, in which accountability is thoroughly institutionalized, and the world of social science research, in which it is not, is that external audits are enforced, a stick that all agree to.

[15] In Cambodia this implies a serious reconsideration of the current attitude that data collection can be done by anyone with a two-day training.

I have to admit that I do not have ready-made incentive proposals on offer. But I do have two examples of institutionalized approaches to changing research practice that at least illustrate the feasibility and effectiveness of applying sticks and carrots.

Examples of institutionalized approaches to changing research practice

The first is the earlier mentioned *movement for Evidence Based Practice in medicine*. The core of this movement is a web-accessible data base of reviews of existing information, the so-called Cochrane library[16]. Most of these reviews are meta-analyses of clinical trial studies. Around this data base a whole lot of other 'infrastructure' has developed. Working groups on particular medical topics and on review methodology, training facilities in conducting reviews and using reviews for practice, research councils requiring all applications for funding of a clinical trial to be accompanied by a review along Cochrane lines, showing there is a need for yet another trial, regulations of medical associations starting to include protocolization of treatment based on Cochrane library evidence, etc.

Some will argue that not all of this is positive, e.g. the protocolization certainly has a very real dark side[17]. But this example is not to argue for the best practice of the Cochrane library model. It is to point out that within a very short period of time, we talk a good decade here[18], a largely research community initiated movement, mostly driven forward by donated staff time rather than external funds, can create fundamental institutional changes in the incentive structure of a profession.

This ongoing change is certainly worth some evaluative attention to increase our understanding of why this movement for change is so successful[19]. The outcome of such an evaluation might contain interesting and possibly transferable lessons for the world of social science research. The box below gives some examples of lessons already drawn by others.

[16] See: www.cochrane.org

[17] What is difficult or impossible to standardize in protocol form, a lot of the soft stuff of medicine that is known to account for quite a large share of its effectiveness, runs the risk of being devalued…

[18] See: www.cochrane.org/docs/cchronol.htm

[19] E.g. is there a relationship between the increasing popularity of alternative medicine and this movement which emphasizes the scientific basis of allopathic medicine? If this intuition contains any grain of truth, the increasing competition between social science research and other forms of policy relevant information on the 'information market' can be conceptualized as a 'similar' kind of starting situation.

Existing change interventions for social science research practice based on the Cochrane model
The most direct translation to the world of social science is the *Campbell Collaboration*[20] which promotes the use of controlled social experiments as the equivalent of clinical trials, is creating a library of reviews of such experiments, and has working groups developing methodology for reviewing other than experimental studies. The Campbell initiative is research community, i.c. evaluation researchers, driven.

The development of review methodology is at the core of an *ESRC*[21]*-funded Evidence Based Policy and Practice network* (EBPP)[22], be it within an explicit context of improving the research-policy interface. Next to data collection, reviewing existing literature and evidence is the other seriously neglected part of the research cycle. Like with data collection, the review process is essentially non-transparent. Like with data collection, the research profession does subscribe ideologically to the importance of systematically reviewing existing information but refrains from putting it into practice, and is also, again very similar tot the data collection issue, not particularly willing to publicly admit that the current treatment of reviewing is problematic. It s telling and possibly instructive, that the EBPP movement in the UK is essentially policy driven[23]. The Blair administration came to power on a EBPP slogan ('what matters is what works') and put money into promoting EBPP initiatives, both within and outside of government.

The second example of the feasibility and effectiveness of applying an incentive strategy to change research practice is the *ISO-certification of contract research institutes*. I am not familiar with the extent to which this is happening in general but in the Netherlands, which has a large contract research and research quango sector, it is certainly happening. What is being certified are research protocols and administration. But getting back to the earlier benchmark example of accounting, the protocol and admin requirements are a lot more elaborate and strict than current practice, and essentially target greater accountability of the actual research process. The organizations going for certification perceive competitive advantage in being certified, because larger research funders start limiting some of their tenders to certified organizations only.

I am not aware of empirical evaluations investigating the institutional changes that this certification movement is bringing about. Again, interesting lessons might be found.

A closing note: applied research as the vanguard?
I want to close this essay with a question. Is it significant that so much of the examples used are initiatives in and by applied research communities? The Cochrane library was initiated by clinical trial researchers, the Campbell initiative emerged from the evaluation research community, the EBPP movement in the UK is rather policy driven but has many evaluation researchers involved. Why is this the case?

One might argue that applied researchers normally work for users and not primarily to contribute to the corpus of academic publications. This means on the one hand that much of

[20] See: www.campbellcollaboration.org
[21] Economic and Social Research Council, UK
[22] See: www.evidencenetwork.org
[23] See: www.evidencenetwork.org/Documents/wp1.pdf

their work is not subjected to the classical internal quality control mechanism of anonymous peer review, and on the other that their environment is to a much larger extent than that of academic researchers ruled by market forces. In other words, while academic research has a mechanism in place, applied research has to invent one. Especially because something credible in place means a competitive advantage in the struggle for projects. And some will add that the fact of an external (commissioned research) rather than an internal motivation (theory-oriented academic research) for investigating a certain subject, with the associated time pressure and financial constraints that come with contract research, also means that applied research is more prone to cutting corners and thus more in need of additional accountability mechanisms.

All of this will certainly be true some of the time. And the much stronger accountability pressure on applied research because of the external funding and user interests involved is bound to make a real difference. But one might as well highlight the positive side of it all and point out *applied research as a prime locus of professional, methodological and sometimes theoretical innovativeness*[24].

If there is any substance to the observation that it is from the world of applied research that moves to revalue data collection are to be expected, settings like the Cambodian one take on a new significance. A fair amount of research, near 100% commissioned, no traditional intellectual hubs of academic research defending their turf: certainly a good environment to experiment with ways to ensure more accountable methodologically rigorous data gathering. Normally a country like Cambodia is looked upon as underdeveloped in terms of its social science research capacity and infrastructure. But for experiments around issues like this its comparative disadvantage might be a comparative advantage.

[24] Within the discipline of sociology of knowledge, some argue that indeed the locus changed and that the traditional, individualistic, university-based mode of scientific knowledge production starts loosing out to a much more collaborative, network-based new mode that is problem-oriented, i.e. applied, in nature. E.g. See Gibbons, M. et. al. (1994) *The new production of knowledge: the dynamics of science and research in contemporary societies*. London, Sage and its 'sequal' Nowotny, H. et. al. (2001) *Re-thinking science: knowledge and the public in an age of uncertainty*. Cambridge UK, Polity.