Is that really so? Some guidelines when evaluating research

Professor dr.polit

Ragnar Audunson
Oslo University College
Faculty of Journalism
Library and Information Science
Oslo, Norway

1. Introduction

LIS is a research-based profession. Professional practice and professional reflections, therefore, should be based upon research and research-based evidence. That makes it necessary for practitioners to evaluate research in order to sort out good and useful research from flawed and low-quality research. What criteria should we use when evaluating a piece of research? What questions should we ask to the report we are reading in order to sort out reliable research from that which is not so reliable? That is the topic of this paper.

Research can be a powerful and efficient tool. In public debates on the role and impact of libraries or when trying to convince those in charge of public or organisational budgets to allocate resources to libraries, few arguments are more efficient than those which can refer to research-based documentation. It is an efficient tool internally: It can help us make optimal decisions when planning and designing library and information services and it can produce adequate and relevant solutions to technical problems. And, maybe most important of all: High-quality research might help us meet the future in adequate ways and transcend present practices, thereby developing and refining librarianship instead of iterating established practices, norms and standards which will be outdated and irrelevant in tomorrow’s society. Being able to develop the profession by expanding and transcending the repertoire, standards and norms we have inherited (while at the same time building upon those norms and standards) is a precondition for the profession’s survival in today’s environment characterised by profound and rapid changes in the technological as well as in the social and political sphere. Research is probably the most powerful tool we dispose of in generating the
knowledge-base on which new practices can be developed and thoughtful reflection on the role of librarianship can take place.

If however the research undertaken does not match reasonable and generally agreed-upon standards of quality, the effects of research might be exactly adverse of those referred to above. Probably nothing is more dangerous than basing your argument on mediocre or low quality research. It will, in the end be revealed, with negative consequences both for the concrete cause for which you are arguing and – and that it even more serious – for the profession and the legitimacy of its knowledge base in general. A Norwegian example from a different professional field than ours might serve to illustrate my point. Approximately one year ago, some months before we had local elections in Norway in September 2003, the Norwegian association for graduated nurses undertook and published an investigation on the living conditions for old-aged people living in institutions for elderly who are not any longer capable of taking care of themselves.

(http://www.sykepleierforbundet.no/nettside/nsfmain.nsf/webShow/01B8B4DC661684B5C1256D8D004CAFAF?OpenDocument&cat=NSF%20mener%7CTema%20i%20tiden%7CGi%20kvaliteten%20stemme!). The methodology and data-collection was based upon gathering reports from nurses on instances where they had to deviate from established norms of quality. The report revealed that such deviations seemed to be extensive and, in many cases, ugly and shocking. At first, the findings of the research were given a broad coverage in the press. As one might expect, it was used as an efficient argument in the election campaign for the necessity of granting more resources to institutions for old-aged people. But then someone started to look more thoroughly into the methodology used in the research. It turned out that the research design and methodology was so full of serious flaws that it could not be said to document anything. It turned out, then, to be a boomerang which turned back and hit the nurses severely. It might very well be that it described the situation in the institutions in question correctly, but that did not help. It had been presented as a kind of research-based documentation, a thing which it turned out not to be. Not only did the cause for which the nurses where fighting, a cause for which the majority of the population has sympathy, suffer. Trust in nursing as a research based profession and research produced by institutions affiliated with the profession might suffer, and in the long rund that is far more serious. Can we find similar cases of would-be research within our own professional field? I believe we can, and we shall return to that later.

So, then, how can we evaluate a piece of research? Is a given piece of research of a quality that can help us develop and refine our practices and reflections, or is it inflicted with flaws that will make it turn into a boomerang if we base our arguments and thinking upon it. That is the topic of this paper.

I will now reflect for a moment upon the relationship between the field of practice and the field of research in LIS.

2. The relationship between practice and research: a fruitful tension between independence and dependencies on the one side – a dangerous trap on the other

In the example referred to above, an organization – in our case the association of graduated nurses – had ordered and financed a piece of research to be used in a concrete decision-making situation (local elections) in which it had interests. The distance between the organisation ordering the research and those undertaking it was small to the extent of being non-existent, and the organisation ordering the research had strong vested interests in the
results. Probably they would have no interest in financing a piece of research documenting that the allocation of resources to the institutions in question seemed to be optimal and that the causes of problems found could not be traced back to the resource-situation. The organisation was interested in some results and not in others. That is a very usual situation. In a market-driven society, researchers and research-institutions are dependent upon contracts from organisations and enterprises whose agenda it is to promote certain interests. The research they order is, so to speak, a part of their lobbying activity. The researchers sense this, and in a society where research is becoming more and more market driven, there is a danger that they, consciously or unconsciously, adapt to it in order to keep their « customers ». The boarders between research, consultancy work and lobbying become blurred. That is a very dangerous development. It might produce some reports that are useful, even very useful, in the short run, but in the long run it will be destructive for librarianship as a research based profession. I am not saying that research financed and organised in such a way cannot produce high-quality research according to the criteria we will discuss below, but we have every reason to be particularly rigorous when evaluating it. The following question is crucial when evaluating research:

Are we dealing with a consultancy report disguised as research to (mis)use the prestige of research in order to promote certain interests, or are we dealing with a sound piece of research?

On the other hand, I believe it is an asset both for LIS as an academic field and research field and for librarianship as a field of practice that there are links between research and practice. In 1999 I attended an international research conference on cultural politics in Bergen, Norway. I remember many of the cultural sociologists participating at that conference expressing their disinterest in culture as a field of practice. They were researchers. Armed with their theoretical and methodological competence they exploited the field of culture as an empirical field for testing out and refining their theoretical approaches. If their research turned out to be useful, then that was a good thing. But that was by no means their main point but more of an accidental side-effect. No medical researcher, of course, would express his disinterest in the field of medical practice. It is the refinement of medical practice which is the raison d’être of medicine as a research field. The same goes for LIS. The development and refinement of library and information practice is the raison d’être of library and information science as a research field. For LIS that is an asset. Through its relationship with the field of practice LIS is capable of generating knowledge that the pure disciplines are not. The links between research and practice is, therefore, necessary and fruitful. Practitioners, however, tend to stress and evaluate the immediate and instrumental usefulness of research. Therefore practitioners tend to argue for a too close and intimate relationship between research and practice. That is dangerous. Good research shall help us transcend present practices. That implies a critical attitude to present practices, norms and standards, which, in turn, presupposes a certain distance and independence between research and practice. Researchers, on the other hand, often want to loosen ties with the field of practice and develop LIS as a pure and general discipline. That is also a dangerous development, first and foremost because we will then lose the possibilities of generating that kind of knowledge that only practice oriented research is capable of. But it also contains a danger that LIS will disintegrate as a research field and instead become sub-disciplines of and appendices to computer science, literary sociology, the management sciences, pedagogics, political science etc, depending upon the perspectives of the individual researchers and research institutions. There is, thus, within LIS a tension between the field of research and academia on the one hand and practice on the other. If one manages to strike a balance, this tension is highly fruitful. It contributes to
developing the field of practice and it contributes to generating knowledge that disciplines not affiliated to practice is not capable of.

That the problem-formulation of a piece of research is related to and relevant for the field of practice, is a quality criterion against which research can be evaluated.

3. What kind of research are we dealing with? Which criteria are relevant when evaluating it?

Research is a complex concept. That is also the case with LIS. Do researchers developing and testing e.g. mathematically based models for information retrieval or academics preoccupied with professional ethics or the role and potential of public libraries as a meeting place in a multi-cultural and digital context have anything in common? In LIS as a practice-related field they should have. In my ideal LIS-project, researchers with competency in computer science and mathematics work together with researchers with competence in social science and the humanistic sciences in order to develop an integrated arena of social practice. We have not reached that far yet. But even though there are characteristics that are common for every undertaking calling itself scientific and research-based. Among these are:

- Research is based on theory. Theoretical approaches might differ and might even be a source of conflict; disciplines might differ with respect to how developed their theories are. LIS belongs probably to the disciplines with a relatively weakly developed theoretical platform. But all research is guided and informed by theory. It is on the basis of a theory that research questions and hypotheses are formulated. It is on the basis of our theoretical concepts our empirical findings are interpreted. And in the meeting between our theoretical approach, the questions and hypotheses formulated on basis of that and our empirical results, our theories will be refined and developed. So, then, good research has to account for its theoretical point of departure and how the theory in question lead to concrete research questions to be analysed in the given project.

- The researchers first and foremost loyalty lies with with the norms of science and research and that which promotes truth and understanding. In many other spheres of life, other values are more important. The field of public librarianship, for example, is constituted by and organised around certain values – the principle of free borrowing being one of them – to which one expects the members of the field to be loyal. A researcher undertaking a project on the effects of free borrowing must, however, be capable of decoupling himself from such loyalties and see and report what his material really indicates, even if it goes contrary to values he cherishes. So, then, one question to ask when evaluating research is: Does the researcher and his research seem biased by interests, values and ideologies, or has the researcher managed to decouple himself from values irrelevant for research?1

- The researcher is expected to show all his card. In most instances in life, we will try to hide that which is not to our advantage. Whether we are engaged in a political discussion or in wooing that is regarded as legitimate. Not so in research. If your response-rate is critically low, you have to inform about it; if one of the vital questions in your

---

1 In Norway smoking has been totally forbidden on all cafes and restaurants from the 1st of June 2004. The law is based upon a summary of research on the effects of passive smoking produced by a centre for evidence-based medicine. The researchers who have written the report summing up relevant research, have been accused of being ideologically biased and of being more loyal to the political value of reducing smoking than to what research on passive smoking really indicates. I am not in a position to judge whether or not these accusations are correct, but refer to them as an example.
questionnaire turned out to be a bad and invalid operationalisation of the variable you want to measure, you must inform about it. The researcher’s duty is, as far as possible, to give the reader all the information he needs in order to follow and evaluate every step of the research: Is the chosen methodology relevant to investigate the problem in question? Have the variables been operationalised in a valid way? What about the sampling procedures and response rate? Is a given finding really significant? A vital question to ask, then, when evaluating research is: Has the researcher given me all this information? Is it possible to follow him step by step and look into his cards?

But then research differs: It differs with regard to its classification as basic research, applied research or developmental work (although we probably have very little basic research in LIS); it differs as to which aspect of the library and information field it studies, e.g. the retrieval system, the library as an institution or the information seeker; it differs as to theoretical approach, e.g. cognitivist approaches versus social constructivist approaches; it differs as to which scientific disciplines it draws on, e.g. the social sciences, mathematical and engineering sciences or humanist sciences; it differs as to methodological approaches, e.g. quantitative versus qualitative. That, of course, has nothing to do with the quality of research. A study of information seeking based on a social constructivist approach and using qualitative methods is not per se inferior to or superior to a study of information seeking based quantitative approaches (even though the followers of a given theoretical and methodological perspective might believe their perspective is more fruitful than those of others). Every researcher deserves that those evaluating him or her take a friendly attitude with relation to his/her theoretical and methodological point of departure.

I would, however, like to draw attention to another categorization: That between research aiming at developing and testing new models and solutions or solving problems related to the primary tasks of the information field – first and foremost indexing, organizing, retrieving and promoting knowledge and culture – and research aiming at stimulating a research-based reflection on the role and value-base of librarianship. What is the efficiency of a thesaurus developed on the basis of word-associations compared to thesauri developed on the basis of more conventional methods? (Lykke-Nielsen, 2002). How can we develop automated retrieval systems that are better adapted to the way information seekers formulate questions and approach a stranger, even if that stranger is a computer? (Nordli, 2000)? How can we develop methods of promoting literature making us capable of reaching user groups that are low readers, e.g. young men? Such research-questions aim at using research and theory for finding better answers to practical problems. But then we have research aiming at stimulating a qualified and research-based reflection and continuous learning on questions relating to our profession’s value base. In 2001 we edited a book in Norway called The Civilized Information Society. (Audunson and Windfeld-Lund, 2001). Here we try to present research that might stimulate reflection and discussion on the role of public libraries in a digital world. This is an example of research that is not looking for solutions to be organisationally implemented. Such research is also an important and useful part of research in LIS, not the least in a period of rapid change. When evaluating research we have to be aware of such differences. Is the researcher aiming at presenting solutions to be implemented (and which can be right or wrong as a mean to solve the problem in question) or is (s)he aiming at putting forward some research-based ideas to stimulate discussion and further research (and which can be fruitful or not fruitful). That difference leads to different criteria when evaluating a piece of research.
4. The formulation of the problem

Formulating the problem, i.e. moving from a relatively vague idea to a problem-statement that is theoretically anchored and focused, that links the theoretical perspective and the practical research questions fruitfully and consistently together and that leads to concepts to be operationalised and researched or hypotheses to be tested, is maybe the most vital parts of any research project. It informs and governs later stages of the research process, e.g. selecting methodology.

Some fundamental questions to ask are:

Has the researcher accounted for his theoretical point of departure and has (s)he documented that this theoretical approach is relevant to the theme (s)he is going to research and leads to fruitful research-questions? If I want to study public library use and information seeking in a library context, I can utilise a number of different theoretical approaches: I can make use of economic theories based on rational choice (van House 1985, Aabo, 2004); I can make use of institutional theories where library use is seen as resulting from institutionalised norms and standards (Audunson, 1999; March and Olsen, 1989). I can make use of Bourdieu concepts of habitus, i.e. life-style, and economic versus cultural capital; (Hvenegaard-Rasmussen and Jochumsen, 2000; Savolainen, 1995). I can make use of role-theory, e.g. see if differences between genders or social groups in library use can be fruitfully interpreted on the basis of social roles; I can formulate my research on library use on the basis of marxian inspired theory of social classes. I can relate my research on information and library use to the task and problem of the user, e.g. its complexity and the stage in the problem-solving process (Byström and Järvelin, 1995; Kuhlthau, 1991; Ellis, 1989) The point is that I have to account for that. Too often that is not done. Then we cannot fully do that which we above stated as a standard to judge research against: Look the researcher into his cards. The theoretical perspective leads to the basic concepts that the research employs. Different perspectives lead to different contexts. It provides the framework for interpreting results. Similar data, e.g. social and gender differences in library use, will be interpreted differently by a researcher relying on economic theories of rational choice compared to researchers utilising marxist perspectives or role theory. Theoretical perspectives will guide the concrete research questions and it will have a bearing on the choice of methodology.

Are the concrete research questions clearly formulated?

Let us take as an example a researcher interested in

1. The researcher can be interested in an in-depth understanding of the situation of adult learners and to analyse that situation in relation to certain variables, e.g. according to stages in the educational program, the different tasks with which the adult learner is confronted in his/her learning process, the private life situation of the adult learner etc.

2. The researcher is interested in investigating the extent to which differences in library use between adult learners can be explained by variables such as those mentioned under 1.

3. The researcher is interested in studying the effect of a given library program in stimulating and helping adult learners use library resources in their work.
4. The researcher is interested in studying adult learners use of libraries compared to other informational resources.

Are the links between the theoretical perspective and the concrete research questions clear. Is, so to speak, the concrete research question a valid operationalisation of the more general research perspective or researach question implicit in the theory? Are the basic concepts in the Concrete research question well defined?

Are the basic concepts operationalised in a valid way? In order to perform my research and measure my variables, I have to make my basic concepts operationable. That decides the validity of the research. Let us take as an example a researcher wanting to study public library use and base his/her research on Bordieus concept of habitus and way of life. How can « way of life » be operationalised? There are, obviously, several solutions:

• We know that way of life and cultural capital is related to profession and education. One solution, then, could be to categorise respondents to different ways of life according to profession and education. University teachers are relatively low in economic capital, high in intellectual capital and have a lifestyle where consumption of culture is important. We find them in the upper right quadrant below; real-estate brokers are relatively high in economic capital. They read and consume culture, but their cultural consumption tend to be conspicuous and oriented towards the broader side of cultural expressions: They attend musicals more than Wagner-operas and Mozart more than contemporary classical converts and advanced jazz. They are to be found in the upper left quadrant. Nurses and kindergarten-teachers, which can be placed in the lower right quadrant, are low in economic capital; they consume culture but read more women’s magazines like Femme and Elle than poetry and more Bridget Jones than Saramago. Some manual workers might be higher in economic capital, but their main cultural consumption consist of football and some crime. They are placed in the lower left quadrant. Socio-economic status, than, are used to operationalise the concept of way of life.

• But then one could say: This is not valid. People are not slaves of their socio-economic status. Therefore we have to ask people what they really do: If and what they read, if they use other cultural services and, if yes, which and to what extent; what they watch on TV, listen to on radio, which newspapers they read, what they read in the newspaper – sports, culture, financial news, gossip and so on. In such a way we can get a real picture of our respondents’ way of life.

• A third possibility could be to construct a number of people embodying different ways of life and ask our respondents who they tend to identify with.

The point is: Which way of operationalising the concept is the most valid. We are confronted with the same problem with literally speaking any concept: «role», «library use», «quality of library and information services», «status», just to mention a few examples.

5. Selecting methodology and research design

Is the selected methodology relevant with relation to chosen the research problem?

Is the selected methodology adapted to the research problem? Is it grounded in the theoretical perspective and adapted to the concrete research questions?
If the researcher wants to draw conclusions about a whole population, e.g. the information needs and information behaviour of ph.d-students in the social sciences at a given university or in a country or to compare the information behaviour of social science students with the information behavior of students in the humanist disciplines a quantitative approach using survey methodology is the proper choice. Then we have to ask if the sampling is done properly, if the researcher has followed standards when developing and testing out his questionnaire etc. The researcher has to account for what he/she has done in these respects and give us the material we need to look the researcher into his/her cards.

If the researcher wants to generate an in-depth understanding of information-needs or search strategies characterising ph-d-students in social science at different stages in their research, without aiming at statistical generalisations, qualitative methods, e.g. interviews and observation, seems to be a better choice. A survey questionnaire gives little information about the individual respondents but much information on an aggregated level, whereas qualitative methods give much and profound information about individual cases, but not on an aggregated level. They are, therefore, capable of answering different questions. Pharo (2002) was interested in identifying categories of importance when students are searching for material of relevance in their studies on the web in order to develop a model capable of analysing search strategies. Then a qualitative method – in this case observation – was used and turned out to be fruitful. A quantitative method would probably not have been capable of providing data rich enough in content for the purpose. But if he, in his next project, after having identified and developed the important variables in his model, would like to test out the extent to which a given strategy is followed by master students in a given subject in connection with their dissertation work, a quantitative study might be relevant.

If the researcher wants to test causal relationships, e.g. test an hypothesis explaining variations in library use, quantitative methods are again the right choice. Then we have to ask if the researcher applies relevant statistical tools opening up for multivariate causal analysis.

If the researcher wants to test out the effects of an intervention, a randomised controlled study is the best choice. Too many studies evaluating interventions are, in practice, valueless. I myself is an example in that respect. Approximately fifteen years ago I was engaged to undertake an evaluation of a programme undertaken in a suburb of Oslo by the local school and the local library in cooperation. The goal was to keep teen-agers between 13 and 16 years of age as readers and library users. A number of measures where introduced. Several times during the project period and at the end of the project I undertook evaluation studies using different methods, surveys being the most important. Of course the programme has effects. That is the well known Hawthorne-effect. If you, as one did in this project, literally woo a group for a long time, it will have short-time effects. And it is impossible to know which effects are resulting from the intervention and which are the result of other variables. What could I have done? One obvious solution in such evaluation studies is controlled randomised studies. In Norway such a study has been undertaken to test out a library programme aiming at stimulating public health physicians to use research-based documentation in their daily practice. (Forsetlund et.al, 2003). Another solution could be to wait with the research testing out the effects of the programme till the short-time Hawthorne-effects are weaker. Had I, for example, waited one year after the conclusion of the intervention with my research, I would probably have been in a better position to test out if the project had any lasting effects.
Is the chosen method properly designed and implemented?

In quantitative studies there are relatively clear rules for sampling procedures in order to avoid biases and in order to be able to infer from your sample to the population in question. These procedures are based on randomisation. Some of the important questions we have to ask are:

Is the sampling procedure really based upon randomisation? In a recently published book on knowledge management (Holsapple, 2003), a piece of research trying to identify what knowledge managers do is presented. Tables with figures and percentages are presented and used to discuss the content of KM-work. This research bases itself upon a questionnaire partly distributed to participants at a conference, partly to friends of the researchers working within the field of KM. The number of respondents is relatively low – 41 all together. One might ask: Can we, based on the findings and figures presented, draw any inferences whatsoever to the universe of knowledge managers and the content of their work? Has this researcher a sample which only can be analysed in a qualitative way and had it, therefore, been better to interview these respondents qualitatively in order to identify and get an understanding of categories that the content of work of knowledge managers consist of?

One tradition in library research bases itself upon data-collection among library users, for example distributing questionnaires to visitors of the library. When evaluating research using such methods, we should ask:

- Has the researcher designed the project in a way making it possible to avoid systematic biases due to self-selection? If we distribute our questionnaire throughout the library and it is up to the users to fill them in or not, self selection might be a problem. If the librarian or the researcher approach the possible respondents, biases due to our built-in tendency to contact people we find sympathetic and non-threatening, i.e. like ourselves, is a problem.
- Has the researcher taken into consideration the variation in visits throughout the opening hours? If we are to infer from sample to population (the population can for example be every library visitor in week 25 in the public library of Buenos Aires) we have to make use of a randomised process where all members of the population have a known and, usually but not necessarily, equal chance of being selected. If we do not take into consideration fluctuations in library visits throughout the day and throughout the week, that vital condition is not fulfilled. That means that we have to do a thoroughful and tedious work in order to identify such fluctuations before we embark on collecting our data.

If research is undertaken without taking questions such as these into consideration, it is more or less without value. No matter the size of our sample – 100, 200, 300 – our findings are only valid for these 200 or 300 persons. We cannot say anything about the universe of library users. Again, we have a sample yielding data that open up for qualitative interpretation (which, of course, might be both useful and fruitful if we are aware of the limitation), not matter their quantitatitive nature. Probably qualitative methods would have produced more useful results than tables and figures and percentages resulting from research with such methodological flaws. It is highly relevant, therefore, to demand that the researcher in his/her report accounts for measures taken in order to meet the conditions referred to above. That decides, not necessarily the quality of the research, but the criteria according to which it must be evaluated. Are we dealing with quantitative data that allows for qualitative interpretations and generalisations, or can we generalise statistically from sample to universe.
We must also be aware of the limitations of the generalisations we can make. If a user survey in a library is correctly conducted, the universe to which we can make generalisations, is dependent upon how the universe of users is defined: It may be defined as all those holding a library card. Then, if we have drawn our sample properly, we can make inferences to all holders of library cards in the given library – but not to those who make use of library services without having a library card or members of other libraries; it may be defined as all visitors to the library in a given period – a day, a week, a fortnight etc. Then we can make inferences to that group. It can be defined as all inhabitants in a community or all members of an organisation who answer that they have used the library above a minimum level in a given period (ever, at least once over the last 12 months, regular users). Other inferences are, so to speak, qualitative in nature, even though they might be useful and reasonable.

If qualitative methodology is used, is it particularly important that the researcher accounts in detail for every step. It is easier to hide yourself when doing qualitative research – the reader do not have access to the audi-tapes or the observations and what the researcher has done cannot be repeated and tested in the same way as the statistics in a quantitative project. The researcher, therefore, has to account metriciously for how he moves from raw data, e.g. his interview data or observations, to patterns and conclusions. Some important questions to ask, are:

**Is a qualitative approach sensible?** Is the aim of the research to understand, interpret or describe phenomenon and subjective experiences?

**How were the respondents selected?** Is the sample of respondents selected in a way that is relevant if one takes the problem, question and the vital variables of the project into regard?

**How were data collected?** Which methods were selected? Interviewing? Observation and field studies? Focus groups? Document analysis? Is the method of data-collection well grounded in the problem and theoretical perspective of the research?

**Has the researcher accounted for his own perspective?**

How was the analysis performed? How did the researcher move from data to patterns, findings and conclusions? Does this process seem to be firmly grounded?

**Does the results, findings and interpretations help me understand the context in which I work?** Even though qualitative research does not aim at statistical generalisation, it is important to ask to what extent it can help us understand similar situations.

6. **Conclusion**

By asking questions such as these to research we are confronted with instead of having a more opportunistic attitude, we will contribute in developing research as a powerful tool and in developing librarianship as a research-based profession. A research field oriented towards refining practice but with a critical distance to that field combined with a field of practice which is research based but at the same time with a critical distance to the field of research – that will create the dynamic dialectic we need.
Literature


(2nd revised version)