

CONDUCTING CRITICAL LITERATURE REVIEWS: A METHODOLOGICAL NOTE

ALEXANDRA GHEONDEA-ELADI

In this paper a general view over literature reviews is given, as well as a short description of the possibilities of two of the current bibliographical referencing softwares available, EndNote and Mendeley. I argue that unlike other types of literature reviews, a critical literature review cannot be undertaken by writing notes and using the search engines of bibliographical referencing softwares, but by focusing on a judgment structure which evaluates the theoretical background of the papers, their assumptions and one that provides a logical structure of the paper. Such a judgment structure is likely to be used across many fields in the social sciences. The proposed judgment structure has been provided based on its applications in two different fields from the social sciences.

Keywords: literature review, critical thinking, judgment structure.

There are several ways to conduct a literature review. Russell (1988: 127) describes three ways to proceed: “asking people, reading review articles and scouring the literature through use of bibliographic search tools”. As far as I have been able to see, the way in which researchers handle the amount of information is usually different from researcher to researcher. Some scholars extensively read various papers written on a chosen topic and, after a certain point, begin to structure the texts according to the trends in the literature or the scope of their research. Others, read extensively on the topic and choose several articles and books, considered to be the most important articles from the chosen pool of literature and focus on the way these articles have addressed the topic. Another possibility is to take notes from every article read, to make draft reviews in parallel with further reading and then to construct the structure of the information at hand by polishing and interpreting the notes. The last one, is perhaps the oldest of them all, since this was the only possibility a scholar had, when bibliography management software did not exist and making copies of a book was not technically available for everyone. Once mass copying became possible, scholars started to use post-its and notes on hands-on copies of the texts they read, in combination with indexing reading notes from texts which could not be copied. But managing vast amounts of information in such

Adresa de contact a autorului: Alexandra Gheondea-Eladi, Institutul de Cercetare a Calității Vieții al Academiei Române, Calea 13 Septembrie, nr. 13, sector 5, 050711, București, România, e-mail: alexandra.gheondea@googlemail.com.

a way can be difficult, since it all resumes to the researcher's capacity to 'keep in mind' at a certain moment in time a large amount of literature. When computer-based reference management software were created and available freely for everybody the manual indexing system was transposed into an electronic indexing system of bibliographical references. With programs, such as Mendeley, RefWorks, EndNote, Zotero and others, the researcher is no longer required to keep in mind indexes and codes and devise an efficient way to find them within stacks of folders.

Nevertheless, the literature review process goes beyond the simple indexing of text and merging indexes in a coherent story. It is about understanding the logic and meaning of scholarly publications. This perspective is based on the idea that literature reviews can be biased if performed in a way that is not structured and replicable. This is why the main rationale for undertaking structured and replicable literature reviews is to avoid cherry picking and biases that might lead researchers to consider as mainstream points of view which are only marginal in the literature. The dangers of such biased literature reviews is more obvious in medical research in which meta-analyses are employed to argue for or against the use of a certain drug or treatment. If such a review would be in any way biased, the dangers for those in need of treatment may be devastating. Similarly, in social policy, providing biased literature-based evidence that a certain group is vulnerable and should be included for social aid, for example, may lead to excluding another important vulnerable group from the benefits of social aid. In this article I propose a structure for critically evaluating literature on a given topic and argue that such a structure cannot be undertaken by using only bibliographical referencing software's facilities for notes or their automatic indexing and search engines.

WHAT IS A LITERATURE REVIEW?

Most scholars perform literature reviews as part of any scientific endeavour. More than this, a literature review is in itself a research endeavour with its own methodologies and specific inquiry styles. In performing a literature review it is important to know how to obtain relevant, accurate and as complete as possible information about the topic being researched. Literature review is above all a process by which a researcher surveys the literature available in a field of study at a certain point in time. The goal of a literature review is to provide a structured perspective over the current knowledge in a certain domain at a certain point in time.

This means that, on top of a general or more specific understanding of science, there should be a basic understanding of how scientific literature is stored and where it can be found (Russell, 1988). Around the world, there are various journal data-bases, like Scopus, EBSCO, Wiley, Blackwell, Sage and so on. The process by which literature review should be performed is well-documented in most social research manuals (Bryman, 2008). It is mainly comprised of two steps: searching

and analysing literature. The key to an effective and efficient literature search is **structure** and **replicability**. Bryman (2008) differentiates between a *narrative review* which is an exploratory endeavour aimed at understanding a new field of research and discovering its main concepts and ideas and a *structured review* which is a process in which the researcher investigates a field of study in a systematic and transparent manner in order to avoid biases that might lead to false conclusions about current knowledge in the field of study.

BIBLIOGRAPHICAL REFERENCING SOFTWARE

There is a wide variety of bibliographical referencing software available nowadays online, like Mendeley, RefWorks, EndNote and so on. These are programs that have several facilities in common: they store bibliographical references with or without source files, they are able to generate bibliographical references in most of the accepted formats, like American Psychological Association (APA), Harvard (in its many forms), and also are able to perform automatic indexing for efficient searches.

In order to understand better the difference between bibliographical referencing software, I will compare the use of two such software: Mendeley 1.13.4 (developed by Mendeley Ltd., freely available on the web) and EndNote 8 (developed by Thompson Reuters, available upon paid subscription). A detailed review of these software is beyond the scope of this paper. The comparison that will be undertaken here is meant only to provide evidence for their possibilities in performing literature reviews.

The main criteria used to evaluate the two software are: availability of referencing styles, the intuitiveness and usefulness of the display interface, importing and exporting options, notes options, pdf file indexing/coding options, portability, user accessibility. These criteria have been chosen to reflect the possibilities offered by the two software with respect to the two steps of a literature review – the search and the analysis. The two programs have been chosen to exemplify the use of one open source and of one subscription-based program. There are other software available for qualitative analysis of the literature, like MAXQDA, NVIVO, ATLASI/TI, ETHNOGRAPH and so on. Some researchers use Microsoft Excel for qualitative data analysis (Meyer and Avery, 2008). These represent software designed for qualitative data analysis, and not for bibliographic referencing.

Reviewing EndNote 8 and Mendeley 1.13.4

Both softwares are able to use most of the referencing styles currently used, on an intuitive user-interface. Referencing styles differ in the way the references are edited and cited within text. For example, both APA and Harvard use «name, year» citations within the text, while the American Mathematical Society (AMS) style uses numbers for each bibliographical reference. Other citation styles place

extensive citations, in footnotes, across the text, in order to avoid fragmenting the reading. Usually, each journal uses a different citation style and thus authors have to face the time-consuming task of editing each citation within the text, as well as the entire bibliography for each journal. With bibliographic referencing software, such editing is no longer required, unless the citation style used by a journal is not a standard one.

Mendeley's Console view, with all relevant windows displayed at once optimizes the work-space automatically at no time-cost. This view is also well-tolerated on very small laptop screens. This happens because Mendeley uses a smaller font, which scores higher from the point of view of the optimization of space, but is likely to make the use of it rather difficult for those with serious sight impairments or those who simply feel more comfortable with larger fonts. Both Mendeley and EndNote can import files from most of the important journal databases. On the accuracy of importing, more errors occurred while importing with EndNote, than when using Mendeley (Basak, 2014). Export options for both of these softwares are available and easy to use, so none of the older files will be lost. Both of them export citations as text and both of them are able to export notes for the references. Nevertheless, Mendeley is easier to use with or without internet, since the web-based and desktop accounts are synchronized when-ever there is a network available, while EndNote requires work either on a web-based account or on a desktop account, depending upon the purchased subscription. The results of this comparison have been summarized in *Table no. 1*.

Table no. 1

Evaluation of EndNote and Mendeley Bibliographical Referencing Software

Criteria	EndNote	Mendeley
Multiple referencing styles	Yes	Yes
Display Intuitive Helpful	Yes The multiple windows need arrangement before efficient work can be carried out	Yes Yes
Subscription	Yes	No
Import options	On most journal databases	On most journal databases
Notes options	Text and some editing options	Text and some editing options
Export options	As text	As text
Pdf file indexing/ coding options	No	Only highlights and notes
Portability	Desktop version: upon subscription, but available only on the personal computer. Web-based version: upon subscription, but unavailable without internet.	Desktop account and Web-based account are synchronized regularly or on demand via the internet.
User accessibility	Font of display is quite big	Font of display is very small and cannot be modified

Performing searches

Perhaps the most important feature of a bibliographical referencing software is its search engine. The fields used for searching and the type of searching are crucial for retrieving information from a bibliography list. In both EndNote and Mendeley, searches are conducted within all fields or in specific fields of a bibliographical reference, including its notes (University of Queensland Australia, 2014; Mendeley Support n.d.). The searches described earlier are conducted within the bibliographical references that have been previously selected by the researcher. Both of these softwares perform boolean searches, meaning that logical operands, like OR and AND, are accepted (University of Queensland Australia, 2014; Mendeley Support n.d.).

Indexing for efficient searches is undertaken in several ways. First, the metadata of each file is indexed automatically. Secondly, the researcher is able to attach tags to each article. Thirdly, the content of the notes for each article is indexed automatically. In this way, when a search is performed, it is possible to access information not just from the title or the abstract of an article, but its notes, as well.

CRITICAL LITERATURE REVIEWS

There are many types of literature reviews that a researcher may perform. The most commonly used in the social sciences are: structured and narrative reviews (Bryman, 2008). Other types include meta-analyses, exhaustive literature reviews, bounded literature reviews, exploratory literature reviews. As it may become clear in the following lines, these types of literature review are not mutually exclusive. Structured literature reviews are reviews conducted in a systematic and transparent manner and recorded as such (Bryman, 2008). The main feature of such literature reviews is that they are replicable. The first step of a literature review, namely the search, should provide *search logs* for a pre-established period of time, with pre-determined key-words and all of their synonyms, antonyms and related terms. Meta-analyses are based on structured and critical literature reviews, but their goal is to statistically describe the results in a field of research, over a chosen period of time. Exhaustive literature reviews seek to describe as much of the literature in a field as possible. Very few of them exist nowadays, since the rapid increase of scholarly publications. On the other hand, the bounded literature review is meant to describe available knowledge in a field of study, in a relatively short-period of time, since the last literature review. Such an update can be bounded not only by time, but also by a new sub-field of a field of study. Last, but not least, exploratory (or narrative) literature reviews provide basic information in a field of study. The span of these reviews is quite large, since they cover a wider area of knowledge, together with its subfields. The point of an

exploratory literature review is to provide the researcher with general knowledge on a wider field in which specialization will later become possible.

One of the rarest types, nonetheless, is the critical literature review. I use the term *critical*, in the sense used by Karl Popper, in which the source of the information is irrelevant to the value of truth of the information (see Miller, 1985 for a selection of Popper's texts). *Critical*, for Karl Popper means also that by means of logic, the researcher is able to evaluate any argument, either by evaluating first-order implications, like "if A then B", or second-order implications "if A, performed in D way, then B, measured in E way". Furthermore, the term *critical* means the acknowledgement of situations when there is *not* enough information to make a judgment. Although, this may seem obvious, being able to recognize such cases is just as important as being able to provide an independent judgment based on relevant information.

At the same time, the term *critical* is intended to reflect the sense used by Brookfield (2012). In his book about teaching critical thinking he points out the importance of eliciting and understanding the underpinning assumptions of an argument or of a text in order to critically evaluate it. Assumptions are rarely explicit in the social sciences and they usually are based upon the theoretical perspective behind a piece of scientific literature.

A JUDGMENT STRUCTURE FOR CRITICAL LITERATURE REVIEWS

In providing this example, I do not intend to argue that this is the only possible way to perform critical literature reviews. I am, in fact, quite sure that there are many different ways to perform a critical literature review. Still, I think that some of the features of the judgment structure that will be the center of my exemplification will be common to most other critical thinking structures. But, before proceeding with this judgment structure, I would like to begin with an exemplification of a literature review *process*. The exemplification will be based on the literature review of a research conducted in the field of treatment decisions and patient autonomy.

One literature review process

The project on which this review is based was meant to describe the structure of the decision-making process of patients diagnosed with a chronic illness (PwCh.I.D), regarding the choice of a treatment for their illness. It also wishes to describe the differences in terms of structure of the decision-making process for patients and their surrogate decision-makers. By selecting a sample from the population of patients diagnosed with chronic illness and their surrogates, the final aim of the study will be to generalize the findings of this research to this particular population. Therefore, patients (P.) – experimental units – who are under the

influence of the main independent variable called diagnosis of a chronic illness (Ch.I.D) will be observed with respect to the main dependent variable – or response – called structure of the decision-making process (SoD.M.P). The SoD.M.P. will be evaluated according to the way in which P. aim to both preserve their autonomy and increase their well-being, at the same time. Particular types of the chosen Ch.I.D represent the levels of the independent variable.

The exploratory review

The exploratory review was scheduled to take two weeks in which the library catalog was searched for general books on the sociology of health or health-behaviour. The point of this review was to identify bibliographical references which would provide general knowledge in the field of study, namely, the sociology of health and health-behaviour, such that at the end of this review a series of questions could be devised that would structure the forthcoming, more focused, review. The search log for this review looks like this:

17.12.2014
keyword: sociology of health
[University] Library Catalog
10 results
browsed: 10
saved: 10
refine search: health behaviour
14 results
browsed: 19
saved: 19
keyword: sociology of medicine
47 results
browsed: 47
saved: 16

keyword: medical decisions
0 results:

keyword: decision-making AND health
 . . .

The result of this exploratory review was a mapping of the main research questions, into sub-questions that required answers from the literature. More than this, the exploratory review on the sociology of health and medicine has been paralleled by some targeted methodological readings, into quasi-experimental and qualitative methodology. The document which resulted at the end of this review is the following:

The main research question of this project is:

How do patients who received a diagnostic of chronic illness and their surrogate decision-makers decide among multiple possible treatments available for their illness?

And in particular, the study focuses on

How do they pursue both autonomy and well-being?

To answer this question several other questions need to be answered:

Q1. Do patients decide or do others decide for them?

Q1.1. Identify characteristics of P. which cause them to decide for themselves

Q1.2. Identify characteristics of P. which cause them not to decide for themselves

Q1.3. Identify characteristics of the Ch. I or its D. which cause P. to decide for themselves

Q1.4. Identify characteristics of the Ch. I or its D. which cause P not to decide for themselves

Q2.1. Which SoD.M.P do PwCh.I.D who decide for themselves use:

I Decision under Uncertainty;

II Bounded Rationality;

III Rational Choice;

IV Rawls' Theory of Justice;

V Parallel Games;

VI Multi-Criteria Decision-making;

VII Other:

i Mix between the above;

ii Which other?

Q2.1.1. Operationalize each theory for PwCh.I.

Q2.2. Which SoD.M.P do Surrogates of PwCh.I use:

I Decision under Uncertainty;

II Bounded Rationality;

III Rational Choice;

IV Rawls' Theory of Justice;

V Parallel Games;

VI Multi-Criteria Decision-making;

VII Other:

i Mix between the above;

ii Which other?

Q2.2.1. Operationalize each theory for surrogates of PwCh.I.D.

Q3. What are the multiple treatments available for the Ch.I?

Q3.1. What makes something a treatment?

Q3.2. What available treatments are there for the chosen Ch.I.D?

Q3.3. What is the chosen Ch.I.D?

Q3.3.3. What were the criteria employed for choosing this particular Ch.I.D?

Q3.4. What is a Ch.I and a diagnostic?

According to Dean and Voss (1999), it is crucial to identify all sources of variation, like nuisance factors, covariates, and blocking factors. Nuisance factors (n.f) are independent variables which are likely to influence the response to the experimental treatment and which should either be kept constant or varied in a controlled fashion in order to differentiate the response to the experimental treatment from the response to these variables. Nuisance factors can either be covariates or blocking factors. An example would be the cultural norms which govern the patient-doctor relationship, or the very type of the relationship between the doctor and the patient. Covariates are independent variables which are likely to influence the response to the experimental treatment and which are characteristic to the experimental units, or intrinsic to them, such as, for example, the personal traits of a patient in our case, or their age, medical history, etc. Blocking factors are independent variables which are likely to influence the response to the experimental treatment, but are not of interest for the research study, and which should be varied in a controlled fashion or kept constant. Dean and Voss (1999: 10) suggest that these nuisance factors, covariates and blocking factors, do not need to be measured, but they have to be acknowledged. In our case, an example would be the place where the experiment is undertaken. Consequently, another question would be:

Q4. What are the nuisance factors, covariates and blocking factors of the SoD.M.P?

This type of analysis will also help differentiate between the various n.f and the treatment levels, thus leading to the next important question:

Q5. Which aspects of the Ch.I and its D. influence the selection of a SoDMP or another?

Q5.1. Which aspects influencing the SoDMP are mediated by the Ch.I.D.?

As Shadish et al. 2002 suggest, it would be important to evaluate the methodological assumptions of the research design. First, it is likely that such assumptions appear from the acknowledgment of the factors that are assumed not to influence the SoDMP, like the way the decision-making task is presented by the researcher and so on.

Q6. What is assumed not to influence the SoDMP?

Secondly, such assumptions appear from the evaluation of the reliability of the measurements.

Q7. Do the indicators employed measure the intended concepts?

The targeted review

Following the exploratory review, the resulting questions were pursued in the literature one by one, in a similar manner. The same type of review was undertaken for the methodological part of the project. The search log on the main data analysis technique employed in the study described earlier, verbal protocols, was as follows:

```

5.02.2015
Keyword: verbal protocol
Database: EBSCOhost
Results: 247056
Reviewed: 50
Saved: 14
    Keyword: verbal protocol AND patient decision
    Results: 82311
    Reviewed: 50
    Saved: 7
        Keyword: verbal protocol AND patient decision AND Hepatitis C
        Results: 60
        Reviewed: 60
        Saved: 0
    Keyword: verbal protocol AND patient decision AND surrogate decision
maker
    Results: 10409
    Reviewed: 11
    Saved: 6

```

By undertaking a targeted review on each of the questions and sub-questions from Section, the main concepts of the research have been explored and defined. The process was structured into two-weeks periods of reading and taking notes for each separate sub-question of the literature review and one week of writing down the review.

The critical literature review

The result of the first two steps of this process has been the creation of a general review of the literature that will underpin the entire research of the project. In any reviews there are crucial areas of knowledge in which it is not enough to simply review what others have written. Such areas of knowledge will inform operationalization of concepts, building up research instruments and the data analysis. For the main elements of a research a critical review is necessary, in the sense I have described earlier (see the opening of Section). This is the point in which a structure of judgment is required.

A structure of judgment for a literature review is simply a set of criteria based on which all references are evaluated. As mentioned before, it can take many shapes and have different criteria for evaluation. Still, the one that I am going to describe is likely to be applicable to a great number of papers in the social sciences and a wide range of domains from the social sciences: from decision-making to health-behaviour, volunteering, work-life balance and mathematical-sociology. To point this out, I will describe two variations of the same structure of judgment applied to two very different domains: game theoretical applications to the social sciences and work-life balance. Following these two examples, the general judgment structure will be obtained.

The first variation of the judgment structure has the following criteria:

1. Year	17. Nationality/ethnicity of participants
2. Authors	18. Number of participants
3. Title	19. Treatment
4. Abstract	20. Player1
5. Declared goal	21. Player2
6. Logical structure of paper	22. Alternatives in game 1
7. Theoretical background	23. Alternatives in game 2
8. Assumptions	24. Types of pay-offs
9. Dependent variable1	25. Number of rounds
10. Dependent variable2	26. If there is a preference elicitation session
11. Mediating variable	27. Type of knowledge
12. Independent variable1	28. Results
13. Independent variable2	29. Limits
14. Independent variable3	30. Observations
15. Independent variable4	31. Critiques
16. Method	32. How to use it

The first four items in the list above are identification criteria, while the last four are critical criteria (*Limits*, *Observations*, *Critiques* and *How to use it*), together with the sixth and the seventh. The rest of the criteria in the list above are specific to the research question investigated. At the end of the review a clear view over these criteria was expected across the entire range of the reviewed papers.

The first critical criterion of evaluation is given by the *Logical structure of the paper*. This particular perspective is meant to force the reader to summarize each paper as an *if ... then* sentence. The way in which I have found this to be very useful is not to place here the results of the paper, but to evaluate the meaning of the results. Most of the times I asked myself what was the meaning of the results obtained in this paper. What were the implications of obtaining the intended result, given the assumptions and the theoretical background of the paper?

The seventh criterion, *Theoretical Background*, was one that is likely to pass as unimportant, but for a critical analysis it plays a crucial role. Clearly, evaluating this criterion requires some previous knowledge on the major theories in the field of study and their characteristics. Sometimes, it happens that the author of the

paper being evaluated openly states this. Placing a paper within its theoretical background is likely to give contextual information on the general approach and the underlying assumptions.

Understanding the assumptions of a paper is likely to provide a critical perspective over the research being evaluated from the very beginning. The assumptions of a paper are usually not clearly stated within the paper. Sometimes, they emerge from the theoretical background that the research takes for granted, and which may not be very obvious for everybody. Assumptions usually enter the logical structure of the paper in an implicit way, which is why it is always important to try to elicit them in the *Assumptions* field.

The twenty-ninth criterion, *Limits*, is meant to capture the limits of the research being evaluated, both those that have been already mentioned in the paper and those that have not. Some research is unconventional and this is why an *Observations* field is usually necessary to state any outstanding, but important elements of the research.

The field on *Critiques* is meant to force the reader to think about criticizing the research being evaluated. Sometimes no critiques appear to mind, but most of the times they do and being able to fill out this field is more a thinking habit than something which speaks about the quality of the papers being reviewed. An interesting account on critical thinking as habit is given by Brookfield 2012.

The last, but by far one of the most important fields is the thirty-second one, named *How to use it*. The name of this field is ambiguous for a reason. The point is that it is not just the results of a paper that may be further used. Methodology, certain methodological details, certain questions, perspectives, definitions, concepts and so on might be of use for the reader's future research. The way in which I have been using this field is to write down why this article is important for my research and how should it be used further on.

The second example of judgment structure was used in a research on work-and life-balance. It comprises the following criteria:

1. Authors	14. Norms of Family Role
2. Year	15. Norms of Work Role
3. Nationality/cultural belonging of participants/ other characteristics of participants	16. Study method and data analysis
4. Abstract	17. Number of participants
5. Theoretical Background	18. Dependent Variable
6. Assumptions	19. Mediating Variable
7. Logical structure of the paper	20. Independent Variable1
8. Segmentation	21. Independent Variable2
9. Compensation	22. Independent Variable3
10. Fixed resources	23. Results
11. Spill-over	24. Observations
12. Researched Family Role	26. Critiques
13. Researched Work Role	26. How to use it

As with the previous example, the first four fields are needed for identification, while the *Theoretical Background*, the *Assumptions* and the *Logical structure of the paper* are important critical evaluation criteria, together with the last three of them, namely *Observations*, *Critiques* and *How to use it*. The rest of the fields are there because of the main research question that this particular example of literature review was meant to answer. The fields on the *Dependent variables*, *Mediating* and *Independent variables* are also common to the ones used in the previous example. A detailed exemplification of some of the articles evaluated by these criteria has been given in the Appendix.

It is by now easy to note the common features of the two judgment structures provided. Still, the main criteria that I suggest should not comprised of only the common features of both, but also some of the features that, may prove to be useful in later reviews and should probably have been used in both of these structures. Its main criteria are

Author	Method
Year	Number of participants
Abstract	Dependent variables
Nationality/cultural belonging of participants/other characteristics of participants	Mediating variables
Declared goal	Independent variables
Theoretical background	Results
Assumptions	Observations
Logical structure of the paper	Critiques
Research specific criteria	How to use it

CONCLUSIONS

Literature reviews are very common tasks for most researchers and they are mandatory in almost any scientific field. In the social sciences, literature reviews are particularly extensive and require some indexing, searching and synthesis abilities. Yet, the structure of a literature review varies based on the purpose of the review. A critical perspective over the reviewed literature can provide a common ground for evaluating scientific literature. Just like in many other fields the ability to think critically is likely to be mastered by providing structure to the reviewing process and the habit of thinking critically. Also, a critical review of the literature is likely to be aided by the use of judgment structures. By regarding critical thinking from the point of view of Karl Popper (Miller, 1985) and Brookfield (2012), this paper provided a common judgment structure which takes into consideration the theoretical background of the papers, their assumptions and the logical structure of the paper, on top of the research specific variables.

The way in which a certain procedure leads to a better review than a random selection of papers is nevertheless not a trivial issue. Some readers may argue that

the points presented here are not backed up by evidence that such structure may enhance the quality of the review. The answer to this question is based upon the idea that literature reviews are research endeavors in their own right. Let us be reminded of how research quality is assessed in science. First, there is the criterion of replicability, which means that an experiment can be undertaken several times, in the same way that will lead to the same results. A random selection of literature will not lead to the same results if performed by other researchers unless, the task is performed a great deal of times and some mean can be computed from all the means of the random samples. This is not possible because ideas, are not quantifiable in this way. Secondly, literature review results should be amenable to the general population of literature in that field or sub-field. This property is usually called generalizability and it opens the discussion on how the sampling of the literature has been performed. Thirdly, the results of the literature review should reveal reliability, meaning that its results should accurately reveal the state of the art in a certain discipline, as close as possible to the entire literature in the field. For this reason, each piece of literature should be analyzed from the point of view of the same criteria used for the others. This is where the judgment structure provided in this paper will be useful. It provides a common framework for review of each article and it makes the criteria transparent for everybody to see.

Acknowledgments

This paper is made and published under the aegis of the Research Institute for Quality of Life, Romanian Academy as a part of programme co-funded by the European Union within the Operational Sectorial Programme for Human Resources Development through the project for *Pluri- and inter-disciplinary in doctoral and post-doctoral programmes* Project Code: POSDRU/159/1.5/S/141086. Thanks to the anonymous reviewers from the Quality of Life Journal for their comments and suggestions that have improved the presentation of this paper.

BIBLIOGRAPHY

- Basak, S. K., *A Comparison of Researcher's Reference Management Software: Refworks, Mendeley, and EndNote*, in "Journal of Economics and Behavioral Studies", vol. 6, no. 7, 2014, pp. 561–568, available online at: <http://search.proquest.com/docview/1560669492?accountid=15533>.
- Brookfield, S. D., *Teaching for Critical Thinking. Tools and Techniques to Help Students Question their Assumptions*, San Francisco, Jossey-Bass, 2012.
- Bryman, A., *Qualitative Data Analysis*, in "Social Research Methods", Oxford, Oxford University Press, 2008, pp. 537–563.
- Bryman, A., *Social Research Methods*, 2nd ed., New York, Oxford University Press, 2008.
- Dean, A., Voss, D., *Design and Analysis of Experiments*, New York, Springer, 1999.
- Mendeley Support, *What advanced search features are there in Mendeley Desktop*, available online at <http://support.mendeley.com/customer/portal/articles/227955-what-advanced-search-features-are-there-in-mendeley-desktop->.
- Meyer, D. Z., Avery, L. M., *Excel as a Qualitative Data Analysis Tool*, in "Field Methods", vol. 21, no. 1, 2008, pp. 91–112.
- Miller, D., ed., *Popper Selections*, 2nd ed., New Jersey, Princeton University Press, 1985.
- Russell, H. B., *Research Methods in Cultural Anthropology*, Newbury Park, CA: SAGE Publications, Inc., 1988.
- Shadish, W. R., Cook, T. D., Campbell, D. T., Generalized causal inference: a grounded theory, in *Experimental and Quasi-Experimental Designs for Generalized Causal Inference*, Belmont, CA: Wadsworth, Cengage Learning, 2002, pp. 341–373.
- University of Queensland Australia, Searching an EndNote library, 2014, available online at <https://www.library.uq.edu.au/research-support/searching-endnote-library>.

În acest articol este prezentată o perspectivă generală asupra analizei literaturii științifice, precum și facilitățile oferite de două programe actuale de gestiune a referințelor bibliografice, EndNote și Mendeley. Argumentul principal este că spre deosebire de alte tipuri de analize de literatură științifică, analizele critice nu pot fi realizate doar prin intermediul scrierii de fișe de lectură și căutărilor asistate de motoarele de căutare ale programelor pentru referințe bibliografice, ci prin utilizarea unei structuri de judecată care evaluează perspectiva teoretică din spatele articolelor evaluate, presupunerile acestora și structura logică a articolului. O astfel de structură de judecată poate fi folosită în multe domenii de științele sociale. Structura de judecată propusă a fost obținută pe baza aplicațiilor sale în două domenii diferite din științele sociale.

Cuvinte-cheie: analiza literaturii științifice, gândire critică, structură de gândire.

Appendix

Table no. 1

Judgment structure of work-life balance review. Part 1

Authors	Year	Nationality/cultural belonging of participants/other characteristics of participants	Abstract
Wicker, A.W., Burley, K.A.	1991	Southern California, USA	This article is very important because it shows how 2 norms associated with the family-couple relationship (equity and hierarchy) interact in the decisions (which can hardly be called decisions, according to the article) about the division of labour. These 2 norms actually give a combination rule for the influence of the partners in work or family areas depending on the gender-type for the business. The resulting combinations are summarized at pg. 89, par. 3.
Wallis, J.J.	2011		This paper argues that the real trade-off that individuals face is between the organizations they belong to in contrast to social rules, and asks when do individuals find it in their interests to act in the interests of their organizations and when do they support impersonal rules? The answer involves a distinction between anonymous relationships between individuals who do not know each other personally, but know the organizations that the other belong to, and impersonal relationships in which all individuals are treated the same. (48)
Bagger, J., Andrew, L., Gutek, B.A.	2008	USA, employees of one national architectural firm	Using identity theory and the gender role framework, this study examined the interactive effects of family identity salience, family-interference-with-work (FIW), and gender on two outcome variables: job satisfaction and job distress. Results from a sample of 163 employees support the proposed buffering hypothesis for job distress and job satisfaction, such that individuals who experienced a high level of FIW reported more job distress and less job satisfaction only when they were low in family identity salience. Additionally, we found support for a three-way interaction, such that the two-way interactive effects of family identity salience and FIW on job satisfaction were stronger for women than for men. Theoretical and practical implications of the results, as well as directions for future research, are discussed.
Brito, R., Waldzus, S., Sekerdej, M., Scubert, T.	2011	Lisbon, Portugal, people in metropolitan area	A study tested hypotheses derived from relational models theory on how four models – communal sharing, authority ranking, equality matching, and market pricing – are used to structure relationships in different types of social groupings. Portuguese participants rated 10 relationships on the Modes of Relationships Questionnaire and reported their shared group memberships. The four-factor structure of the models was confirmed, and a bottom-up approach identified families, friendships, and organizations as the main distinct types of social groupings. The main results showed that relations within families combine mostly communal relations with authority between age groups, whereas relations in organizations combine mostly market relations with authority between age groups as well as professions. Friendships feature a mix of all models, except authority.

Table no. 2

Judgment structure of work-life balance review. Part 2

Authors	Theoretical Background	Assumptions	Logical structure of the paper
Wicker, A.W., Burley, K.A.	Social norms theory	Conflicts may be reduced by normative differentiation, that is, by adhering to different norms in different spheres of activity	If the assumption is correct, then the spouses have different powers to influence decisions about the division of labour among them, according to the norms which dominate each domain (family and work).
Wallis, J.J.	Neo-institutionalism	The decision to support an organization depends on the tastes, preferences, beliefs, the choices available and the relative prices associated with the choices; organizations are 'aligned' 'interests of the individual members' created to 'sustain coordinated and cooperative behaviour'; organizations generally propose personal rules that apply to members, not impersonal rules; Stryker 'mutually supporting organizations' were created as a response to 'the threat of violence, to sustain a modicum of cooperation [between elites] and to enable third-party enforcement for organizational arrangements' (49)	If the assumptions are correct, then, in obeying rules, people choose between the interests of the organizations they belong to and obeying impersonal rules, rather than choosing between the interest of the group and the interests of the individual, as proposed by typical economic theory.
Bagger, J., Andrew, L., Gutek, B. A.	Identity theory and gender theory	'(Stryker, 1980; 1987) once an individual instils meaning into any of his/her social roles, that role becomes part of his/her self-concept, or identity. Identity performs two important functions.: gives 'meaning to life' and 'provides normative guidance for behaviour'	If the assumption is correct, then identity salience moderates converse proportionally the relationships between: job satisfaction and family-interference-with-work and between job distress and FIW and this moderating effects are different for women than for men.
Brito, R., Waldzus, S., Sekerdej, M., Schubert, T.	Relational models theory	Relationships can be identified by each person and they may be described by four basic relational models: community sharing, equality matching, authority ranking and market pricing; 'relationships tend to develop within any type of groups or social units that are bounded social systems or relational networks' (411)	Check if assumptions are correct (reliability of relational models, identity groups that define the distinct social context), then 'a large part of relational variation can be explained by the type of social unit in which a relationship develops. In addition, membership in social categories, such as age and gender, could be relevant for interpersonal relationships within those systems or networks.' (411)

Table no. 3

Judgment structure for work-life balance review. Part 3

Authors	Results	Critiques	How can I use it
Wicker, A.W., Burley, K.A.	‘when wives worked fewer hours in the business and when the business was male typed, they had less influence at work and more influence at home. When wives and husbands worked equal hours and when the business was gender-neutral, the two partners had about equal influence in both settings. But when wives worked more hours in the business and when the business was female-typed, they had only about the same amount of influence as their husbands in both settings.’(89)	Somehow the equity norm was left aside at the end, despite stating clearly at the beginning that this norm is also one of the interests of this research; in fact, the results suggest that the equity norm is present (the influence of wives is different depending on the number of hours worked, rather than the equality, which I find hard to notice in this research. The interview guide is missing, so I find it difficult to understand how each concept was actually measured.	This article problematises the discursive utility by saying that some issues remain undiscussed, but are nevertheless solved or decided upon, somehow. One explanation, which Wicker and Burley do not give is that there is another norm that relates to the very task of division of labour: the woman should take care of the house chores, while the man takes care of whatever is outside the house. In this article, the norms they discuss are meta-norms (as Axelrod, ... defines them), while the norm regarding the division of labour is a ‘direct’ norm.
Wallis, J.J.	Framing the problem as a choice between individual interests and social rules has not led us to satisfactory answer’, therefore the choice should be framed between the interests of the organizations they belong to and the impersonal rules. ‘Impersonal rules are consistent with individual rationality, but only through the intermediation of human organizations’.	Sometimes his arguments are obscured by the form he decides to put them into and by their abstract/theoretical nature. And moreover, the examples he gives seem always a little bit out of the theoretical point he’s trying to make, or they seem forced.	This article points out the need to differentiate between impersonal rules and anonymous rules. By doing so, it emphasizes the conflict between impersonal rules and interpersonal relationships, which is a distinction I should also make in my thesis. Also, read North 1990 and North, Wallis and Weingast 2009 for the original framework that Wallis talks about. This will be useful for the part on rules.

Table no. 3 (continual)

Bagger, J., Andrew, L., Gutek, B.A.	Family identity salience and family-interference-with-work interacted to predict two outcome variables: job satisfaction and job distress, such that the effects of family-interference-with-work on outcome variables were stronger when family identity was less salient. Additionally, we found a three-way interaction involving gender, such that the aforementioned two-way interaction on job satisfaction was stronger for women than for men' (205)	It appears strange that for measuring the family identity salience, the author uses a scale which has been designed to 'equally applicable to men and women... [as well as] to individuals at various stages of role anticipation and/or implementation' (Amatea et al., 1986: 832). I might be getting the meaning of 'equally applicable to men and women', but operationally, I think this means that it will not yield different results for men and women. While there is no significant correlation between Fis and gender, (p. 197), the results show that there are differences in men and women, so the gender bias must necessarily come from the other variable. Secondly, the author says: 'since the idea is to assess the salience of an identity, a person need not currently be in a role to identify with that role' (p. 194) and I think that this assumption was not necessary. Moreover, I think it was used to increase the number of people in the sample and this might bring a huge bias in the results, because of the future projected, hypothetical aspect brought in. In a way, it seems fair to assume that the salience of a role may be salient before the necessary arrangements are made, but from my own experience I know that if you are rather young, your family role salience is rather weak, compared to the family role salience, once you actually have a family, whatever that means to you. Again, I think this bias is not a legitimate risk to take.	This is a very nice example of a well-written article. Look up to it for the Theoretical implications part where the implications of the current findings are discussed in relation to other research. Also, this is an example of research which emphasizes a moderating effect, as Voydanoff (2011) says. Furthermore, this article makes clear what a moderation effect is and how it is measured. Also, it contains a nice presentation of identity theory and unlike other articles using this theory, it actually gives a reference for this theory. You should check this reference up, as well as the Bielby and Bielby 1988 article which seems to be quoted in many articles on this topic.
Brito, R., Waldzus, S., Sekerdej, M., Schubert, T.	The four-dimensional model fits the data better than the two dimensional one. The social contexts identified are: family, friends and organizations. 'The CS was most associated to within-family relationships. EM to both families and friendships, MP to organizations, and AR to both families and organizations. Age difference enhanced AR in organizations as well as in families, but more so in the latter than in the former.' (422)	The discussion on the generalisability of the results is not convincing and the argument seems forced.	This article is part of the group on Fiske's relational models theory. It brings me to wonder if the relational models cannot be related to the classes of games in my model. It also has implications for the article on cross-sex friendships at work by giving the percentage of people who have friends from the opposite gender (p. 417). The authors bring an important idea about groups as functional frames for relationships (422). They make an interesting distinction between groups and social categories and loose associations (411). While intimacy groups and task groups were 'perceived as social contexts for relationships', 'social categories and loose associations were not'. (411)

Judgment structure for game theory review. Part 1

Year	Authors	Title	Abstract
2007	Jenna Bednar and Scott Page	Can game(s) theory explain culture? The emergence of cultural behavior within Multiple games	The hallmarks of cultural behaviour include consistency within and across individuals, variance between populations, behavioural stickiness, and possibly suboptimal performance. In this article, we build a formal framework within which these behavioural attributes emerge from the interactions of purposive agents. We then derive mathematical results showing these behaviours are optimal given our assumptions. Our framework rests on two primary assumptions: (1) agents play ensembles of games, not just single games as is traditionally the case in evolutionary game theory models; and (2) agents have finite cognitive capacity. Our analysis combines agent-based techniques and mathematics, enabling us to explore dynamics and to prove when the behaviours produced by the agents are equilibriums. Our results provide game theoretic foundations for cultural diversity and agent-based support for how cultural behaviour might emerge.
2010	Bednar, Bramson, Jones-Rooy, Page	Emergent cultural signatures and persistent diversity: A model of conformity and consistency	Empirical evidence demonstrates that cultures exist, they differ from one another, they're coherent and yet diversity persists within them. In this paper, we describe a multi-dimensional model of cultural formation that produces all of these properties. Our model includes two forces: an internal desire to be consistent and social pressure to conform. When both forces operate, the society converges to a coordinated behaviour that is consistent across the attributes. We find that convergence in the two-force model is slower than a pure conformity model and that a preponderance of one force over the other slows convergence, rather than hastening it. We further find that the two forces amplify small errors in individual behaviour and prove capable of producing substantial persistent diversity.
2011	Choi and Mengrajani	Can group discussion promote cooperation in ultimatum games?	The influence of discussion-induced shared cognition on bargaining behaviour was examined. Three studies tested the hypothesis that shared cognition regarding the best method for reaching a bargaining agreement would decrease the frequency of selfish offers. Consistent with this prediction, participants who engaged in such a group discussion made less selfish offers than those who did not discuss (all studies) or those who engaged in a group discussion regarding commonalities that they shared (Study 2). Study 3 showed that the discussion effect was mediated by shared cognition developed through the discussion. Thus, discussion regarding how best to reach bargaining settlements may develop shared cognition that assists in cooperative bargaining. Implications and limitations of the studies are discussed.

Table no. 5

Judgment structure for game theory review. Part 2

Year	Authors	Declared Goal
2007	Jenna Bednar and Scott Page	‘to explain the existence, and better yet, the emergence, of these characteristics in a society of rational agents. To date, the formal literature has pursued a different objective when incorporating culture into its models, and therefore has not been able to explain how these five behavioural characteristics of culture might arise. Most of the formal literature that includes culture does so in order to type agents, to demonstrate the effect of different utility functions or belief systems. Relatedly, culture can be appealed to as an external factor that refines equilibriums by making some more focal than others. Two excellent examples of this work are Greif (1994) and Fearon and Laitin (1996). The approach helps us to understand the effects of culture, but it cannot explain its emergence. In the context of the above wish list, these models explain what happens if traits (2)–(5) are true, rather than how these cultural aspects would appear.’ (p. 69)
2010	Bednar, Bramson, Jones-Rooy, Page	Can people be relied upon to be nice to each other? Thomas Hobbes famously did not think so, but his view that rational cooperation does not require that people be nice has never been popular. The debate has continued to simmer since Joseph Butler took up the Hobbist gauntlet in 1725. This article defends the modern version of Hobbism derived largely from game theory against a new school of Butlerians who call themselves behavioural economists. It is agreed that the experimental evidence supports the claim that most people will often make small sacrifices on behalf of others and that a few will sometimes make big sacrifices, but that the larger claims made by contemporary Butlerians lack genuine support.
2011	Choi and Mengrajani	How do fairness concerns or standards of fairness regulate bargaining decisions?

Table no. 6

Judgment structure for game theory review. Part 3

Year	Authors	Logical structure of the paper
2007	Jenna Bednar and Scott Page	If we can improve the theoretical model employed for cultural variation we can predict better the dissemination of cultural traits which have 5 properties.
2010	Bednar, Bramson, Jones-Rooy, Page	If we model agents by an n-dimensional vector (each dimension is a characteristic of the agent), constrained by two forces: consistency of action (actually consistency of attributes because the agent chooses two random distinct attributes and changes the value of the first attribute to match the value of the second p. 418.) and conformity (The first paired agent randomly chooses an attribute and sets the value of that attribute equal to the value that the other agent assigns to that attribute p. 419). , and if culture manifests itself across different domains, we can predict the behaviour of a combined (conformity and consistency) society and estimate the time taken by a combined population of agents to reach convergence.
2011	Choi and Mengrajani	If we can show that people use “shared cognitions” to cooperate in UG then we can explain why so many people usually cooperate.

Judgment structure for game theory review. Part 4

Year	Authors	Assumptions	Limits
2007	Jenna Bednar and Scott Page	Rational agents with cognitive constraints, but actors 'possess sufficient capacity to play any one of the games optimally' (p. 71) 'We first establish that there exists a utilitarian social welfare maximizing equilibrium in two -state automata for each of the six individual games.' (p. 78) This claim is proved in Appendix.	'while these claims are instructive [contextual effects and suboptimal behaviour], they can only derive the best possible outcomes for our agents; they are not capable of demonstrating what is most likely to emerge from a dynamic model of interacting agents.' (p. 78)
2010	Bednar, Bramson, Jones-Rooy, Page		
2011	Choi and Mengrajani	Shared cognitions appear in open communication	Communication was framed as: "After the participants were acquainted with the bargaining rules, the experimental manipulation involving group discussion was initiated. About half of the participants were in this discussion condition. The experimenter asked them to sit facing each other as a group, and converse with each other about the best possible way to approach and reach an agreement in the ultimatum bargaining. The experimenter told them that there was no correct answer, and that reaching a group consensus was the goal of the discussion. A 10-minute time limit was given for the discussion. It should be noted that the experimenter did not explicitly tell them yet that they would later participate in the ultimatum bargaining with each other, although the participants might have considered this as a possibility (p. 384) for the 1st study (while the control group did not engage in any communication) Aiming to reach consensus is similar to aiming to reach cooperation in the game. All should agree and all should be happy. ..."

Judgment structure for game theory review. Part 5

Year	Authors	Observations	Critiques
2007	Jenna Bednar and Scott Page	Suboptimal behaviour is seen as a part of culture: 'Suboptimal behaviour – The strategy employed by individuals within a community may be suboptimal, where individuals could benefit by acting in a different way. Formally, the behaviours are not equilibrium strategies in the repeated game or if they are equilibrium strategies, the resulting equilibrium does not belong to the set of Pareto efficient equilibriums.' (p.69) (comment: it seems that they only use two-player games)	I cannot see the data. There is not statistic for the data... it all seems very much only told, not shown. The computational algorithms are not given, the name of the program used for the simulation is not stated... this makes the article just a report of results, not something one may use further.
2010	Bednar, Bramson, Jones-Rooy, Page	Because conformity and consistency can pull in different directions, the time required to attain an equilibrium can be greatly increased (p. 422)'	This article is written in a completely non-scientific vocabulary and takes the Hobbesian part in the debate between Butlerians and Hobbesians.
2011	Choi and Mengrajani	All participants played the proposer and the responder roles. They were first asked to give the highest offer they were willing to give to a responder and then the lowest offer they were willing to accept from a proposer. The point of this was to earn as many points as possible because the number of points earned would determine the number of chances they have to win in a lottery with a 10\$ prize at the end.	One problem is in the categories of selfish, altruistic and equity offers. In equitable offers the interval 10, 10.5, 11 (from 21) are considered, but this does not seem to reflect the real behaviour of participants since competition is based on the number of points and thus even one point of half a point might determine who wins and who loses. Another critique refers to the following judgment: "If increased sense of ease or attraction for one another was the main cause behind the mean difference in offer levels, then it should have affected the minimum acceptance level too; it also should have been significantly lower (reflecting more generosity) in the main-discussion condition than in the other conditions" (p. 390) I do not find justified the assumption that the level of the offers and the minimum accepted level are not independent events. The level of the offer and the minimum accepted level are independent events if the individual is assumed to be selfish. If we depart from this assumption then we should have altruistic offers and levels to a greater extent, but this does not happen. More than this, the experiment is not preceded by a preference elicitation session and without this, the selfishness assumption is not grounded.

Table no. 9

Judgment structure for game theory review. Part 6

Year	Authors	How to use it
2007	Jenna Bednar and Scott Page	This article uses a short review of UG puzzles in the literature which might prove useful. Secondly it builds a framework for multiple games, which resembles my idea of parallel games
2010	Bednar, Bramson, Jones-Rooy, Page	The conclusions that emerge from such a model are very interesting indeed for the study of culture and possible ways to predict it's development, but this study cannot be used in my thesis.
2011	Choi and Mengrajani	This article may be used as a nice example of good data analysis report. Furthermore, the structure of the experiments is similar to mine so, it would be good to follow its structure of data analysis reporting. Secondly, this article studies the influence of communication on the behaviour in UG, it should be included in the literature review. Thirdly, the idea that the belief in the equality of power between game partners leads to equal offers, while the perception of an unequal power or effort between them leads to unequal offers may be confirmed by some observations during my pre-test: "Once bargainers expect their opponents to base their decisions on the same script that they have (shared cognition), they themselves are likely to make decisions based on the same script improving coordination. Namely, if bargainers believe that they share the same equality or altruistic script, then making equality or altruistic offers would be deemed more feasible because they would be more favourably accepted. (p. 394)"