FQS FORUM: QUALITATIVE SOCIAL RESEARCH SOZIALFORSCHUNG

Volume 6, No. 1, Art. 27
January 2005

Empirics as Comparisons

Mike Metcalfe

Keywords: comparison, empirics, research quality

Abstract: As part of the *FQS* Debates on quality, LAUCKEN discusses "comparisons". This paper picks up on this topic with a particular focus on the empirical evidence presented in support of a knowledge claim. By "empirics" is meant evidence collected through someone's senses. It is believed that thinking about empirics in terms of "comparisons" provides common ground in the quality/validity debate between epistemologies. Therefore, the argument of this paper is that empirical evidence quality can be usefully thought of in terms of "comparisons" rather than the traditional epistemological grounds such as independence, measurement, repeatability or the identification of the conceptual frame. After discussing "comparison" as part of human thought, this paper will suggest how it can be used to design a range of empirical gathering practice.

Table of Contents

1. Problem Statement

Empirics are common across epistemologies even if their role in knowledge creation differs. In science, they reveal the truth (CHALMERS, 1982), in action research they provide an action to reflect against (ARGYRIS & SCHÖN, 1978), while in critical social theory, similarly, they provide the basis of a social critique (ALVESSON & SKÖLDBERG, 2000). They involve researchers in the act of the collection of "first hand" evidence through our own senses. In order for others to accept the quality of sensory evidence, it needs to be collected in a manner that will convince a knowledgeable audience. The problem then becomes one of asking what is required to make empirical evidence convincing. Historically each epistemology has had its own answer.

- A scientist might require precise measurement,
- a critical social theorist might require insight for the less privileged,
- a systems thinker might require demonstration of a novel and useful perspective,

- an action learner might require a series of repeated action and reflective loops,
- a JAMESian Pragmatist might require new evidence fit within the audiences' overall worldview, and
- an interpretive researcher might require the sense-making lens used to interpret the empirics to be made more explicit. [1]

Is there any common ground here? Is there some language that most of these epistemologies would accept as a convincing method of collecting and presenting empirical evidence? Finding some common ground may help at least pacify parts of the epistemological debate over the quality of this sensory evidence. This paper will present evidence to support the argument that the concept of comparison might provide this common ground. It therefore picks up on LAUCKEN's comments that:

> "There is nothing on earth, which cannot be compared to any other regarding any particular aspect. The assertion that two things cannot be compared can usually be disproved with some reflection. When taking a critical stance, incomparabilities should never be searched for. Instead, questioning whether comparisons using this or that measure is of any value, is essential. Do the comparisons at question unearth knowledge? Usually, comparisons are substantial sources of knowledge, but often they result in inanities and sometimes they even lead to false conclusions. It is really worth contemplating measures of comparison" (LAUCKEN, 2002, ¶4). [2]

2. Comparison and Thought

Measurement is a comparison. The physical attributes of some phenomenon under study are comparisons. Counting, as most readers will understand it, uses an international numbering system with a number base of 10. To number base 2, a computer chip may have a memory of 100MB; to number base 10 the chip has 4MB of memory. Reporting the chip's memory is unconvincing unless an agreed number base is used for the comparison. This numbering system underpins an international length, weights, monetary, and time measurement system. To measure the length of something is to compare it to rule; an internationally agreed standard. The ruler reading shows the length of something relative to all other measurable things. A length is a relationship. All metric rulers are calibrated from one agreed upon standard. Therefore, the basis of quality or validation in the sciences may be seen as not being measurement per se but of comparison against an agreed standard. [3]

The meter (m) is the Si unit of length and is defined as the length of the path travelled by light in vacuum during the time interval of 1/299,792,458 of a second (National Institute of Standards and Technology US). Put another way, measurement of length is the difference between a metre and that measured. Some strands of postmodern theory (LUHMANN, 1995) prefer the language of difference to that of comparison. While it is accepted that this alternative

language may draw out a different emphasis, for convenience, this paper will use the word "comparison" generically to include difference. [4]

The use of comparison for the creation of knowledge by human beings seems widespread to the point of being indistinguishable from the concept of thinking. For example, statistics are comparative; variables are compared in regressions, as trends, and against each other directly. IQ scores are comparative, where individual scores are compared to cultural or community norms. An equation is a comparison, one side of the equation needing to be equal to the other. Experiments are comparisons. Typically, they work by comparing an "if-then" statement with sensory inputs. If heat water is heated then it boils. If light passes through a change in medium then it is refracted. [5]

At the level of theory, comparison is also extensively present. In psychology there is a well documented theory called Social Comparison Theory (FESTINGER, 1954), which suggests that people get a sense of self worth by comparing themselves with others. This act has been found to have a significant impact on our well being. Meanwhile, BATSON (1979), the biologist, reminds us the brain and senses are difference detection devices, comparing millions of input senses to some norm, only becoming explicit to their owner when a significant difference is detected. [6]

Comparative analysis is also a well-established method in literary criticism, which uses comparisons and contrasts to help us appreciate one piece of work relative to another. There is also a large literature on comparative religion where religions are compared and contrasted with each other (SMITH, 1993). Comparison is also used in the wider community in order to make important decisions about each other's lives. In a murder trial, the jury is asked to compare numerous things in order to determine what they believe. They can compare what the defendant is claiming with what the prosecutor is saying. An expert witness may compare the DNA of blood found at the scene with the defendant's blood. People who know the defendant may be asked to compare what they know of his or her past behaviour to what is now alleged. Relevant timings of how long various actions take to complete may be compared in re-enactments, perhaps timed. Eye witness reports may be compared to each other in much the same way the forensics are a comparison of international standards of measurement. The Law, as written down, also needs to be compared with the defendant's behaviour. Indeed, our justice system is designed around comparison, justice as comparison. [7]

LATOUR (1986) argues that the progress of science has much to do with precise recordings. Writing, mathematics, pictures and logic can all be seen as means of capturing thought with some precision. Once in a reasonably precise form, then small improvements can be made so as to slowly ratchet-up the detail through precision. Without writing, thoughts are condemned to be always rather vague, having to be reconstructed from first principles. This means that they can get lost or fail to ratchet up to some insightful conclusion, especially when group debate is involved. Edits are compared to the original that is, more and more, picking apart (DEWEY, 1910). Increasing detail becomes possible from precision drawings or

carefully recorded logical steps. Engineering drawings and mapmaking are good examples of thought being precisely recorded to a level not possible in language and thought. The first surveyor will record certain detail onto a diagram or map. Successive surveyors can add to more and more accuracy to the original measurements as better methods become available. Particular disputed readings can be carefully remeasured using tighter and tighter learning loops. The result is a precise record of millions of inter*connections* of landscape that one person could not hope to repeatedly communicate to others, especially after their death. [8]

Comparison may be made more convincing by using multiple comparisons. Rather than comparing two eyewitnesses, three may be used or ten. Eyewitnesses using with very differing vantage points can be more convincing. In positive research, this use of multiple comparisons is sometimes called triangulation after the surveying or navigational practice of using the inter-section of the compass bearing of three or more reference points to locate a different geographical point. For example, the bearing of three hilltops can be used to plot your own position on a map. Triangulation, the use of multiple approaches to data, observations, research methods and theories, is advocated as means to locate a one objective truth (CHAMBERLIN, 1965; OPPERMANN, 2000). The attractiveness of triangulation is that it addresses the problem of the incompleteness of any one particular piece of evidence. Individual pieces of evidence may be open to suspicion but then reinforced with supporting evidence then the overall weight of the evidence becomes convincing. It instils an underlying assumption of research as being measurement and discovery, suggesting research in like an explorer looking down on a map. It therefore is located very much within the objective epistemology including that it is possible to have a detached, overall, "helicopter's eye" view of a research problem. [9]

The triangulation metaphor to visualise multiple *comparison* measurement becomes less attractive to the constructionist epistemologies. The ironic epistemology (HATCH, 1997; RORTY, 1989) for example, does not believe in the "helicopter's eye" view of research problems, that we can ever stand outside of our situation and see the terrain objectively. Rather, it believes us only capable of viewpoints, interpretations, like crossing torch-beams in the darkness. These insights can come from metaphors. Metaphors are *comparisons* (LAKOFF, 1993; MORGAN, 1986). This epistemology would see more comparisons as providing more contradictory, often paradoxical, reasonable, but irreconcilable interpretations of the physical world. For example, NEWTON used the root metaphor (MORGAN, 1986; VAN DE VEN & POOL, 1995) of attraction; EINSTEIN used relative speed; quantum mechanics uses probability distributions as its torch-beam to provide an interpretation. All these interpretations are reasonable; all provide insight and yet at the same time create a certain level of blindness to some issues. Relativity is blind to sub-atomic particles, quantum mechanics to the forces of gravity on astronomical bodies. An irony arises because astronomical bodies are made up of sub-atomic particles. LATOUR (1986) sees the dialectic tension, *comparison*, between measurement and re-visualization (perhaps from metaphors) as being important for knowledge creation. What metaphor lacks in detail it makes up for by shining a new light-

beam of very different viewpoints, where precision can again start to be built up. Re-visualising through metaphor is analogous to DEWEY's (1910) distinction between analysis and synthesis. Analysis is precision picking apart, while synthesis is like the use of metaphor to use images from elsewhere to think about some phenomenon under study. Measurement, precision, and analysis, perhaps through triangulation, are but part of the dialectic *comparison*. The other is new perspective, metaphor, and synthesis. [10]

Another example of comparison being used in different epistemologies is in argumentation theory (REHG, 2004). This is an epistemology where knowledge is created and the quality judged by competitive reasoning. While history has determined that reasoning evidence is usually distinguished from the definition of empirics, reasoning can be seen as a sense. Our brain, therefore our reasoning, significantly influences what our sensory organs are sensing, a influence that is now well acknowledged in the cognitive science literature (BATESON, 1979). An interpretation of the world can be considered an argument, a comparison, between our reasoning and our sensory inputs. MISSIMER (1995) takes this to a strategic level by pointing out that the great discoveries in science are usually achieved in an atmosphere of a new discovery needing to disprove old views. The obvious example is GALILEO's efforts to argue to a sceptical audience that the earth was hurtling through space at a high speed so as to circumnavigate the sun once a year. The common sense view is that the earth is not moving. GALILEO knew people would compare what he was saying to their common sense. Generally, suggestions of new knowledge are likely to be compared to old knowledge. This aligns with the JAMESian (1907/1910) pragmatic definition of "truth". For someone to accept something as true he or she needs to be able to compare it satisfactorily within the context of their overall interpretation of the world. [11]

The action learning epistemology of knowledge creation through learning loops (ARGYRIS & SCHÖN, 1978) provides at least two further examples of comparison. It sets up learning or knowledge creation to be motivated through the identification of problems as the comparison between a desired outcome and the actual one. Then each learning loop can be compared with the previous one. WEICK and ROBERTS (1993) provide an excellent example with their description of a crisis involving a night landing on an aircraft carrier. Here thousands of integrated learning loops have been acted out over a period of 60 years training, involving thousands of sailors attempting to get aircraft airborne and landed at sea under differing conditions. Those involved in the learning experiences also have to pass their knowledge on to the next generation of sailors in a very coordinated way. It is a self-reproducing and learning system, one that inherently holds a large body of knowledge. Repetitions, comparing one learning attempt with another again and again, ratchet-up the knowledge creation detail. Written procedure manuals, and instrumentation allow precision against which sailors can compare and make explicit their experiences and ideas. [12]

Comparison, betweenness, or connection is an epistemology in systems theory. Phenomena are known by their connections, their comparisons to other

phenomena. A green chair is known by its connection to colour through green, it is connected to wood through its frame and to comfort through sitting. VICKERS (1983) provides a more elaborate example. To an architect a university school is known as a building designed to accommodate scholarly activities. To a planner it is known as a service to a community. To an educator it is known as a collection of scholars and students. To a sanitation engineer it is known as one source of sewerage in a network of sources and outlets. LUHMANN's (1995) systems based social theory focuses almost exclusively on the communication between people. BATSON (1979) describes attributes such as pride and aggression as relationships, meaningless unless people are seen to be operating within some system. He sees comparison that identifies a difference as information. Information is defined as a difference that makes a difference. Our sense organs are comparators, designed to respond to differences. BATSON goes on to make the point about the importance of difference by saying the difference between ink and paper is the signal. A stone cannot respond to difference. ACKOFF (2000) often uses the story of the carburettor of a Rolls Royce being no use in a Mini, the relationships are wrong. In the system's epistemology to understand something is to compare it with something else:

> "the ideas which relative words stand for are often clearer and more distinct than those substances to which they belong. The notion we have of a father or brother is a great deal clearer and more distinct than that we have of a man" (An Essay Concerning Human Understanding, John LOCKE, 1690). [13]

3. Empirical Methods

Having argued that differing epistemologies and human inquiry activities seem to make use of the concept of comparison, the idea that this concept may provide common ground will now be extended to the design of research methods or techniques such as interviews, and observation. [14]

3.1 Interviews

Interviews have a long history of being used in a wide variety of situations from the media, to jobs, to police, to court, to research (GUBRIUM & HOLSTEIN, 2000). They come into their own as empirical evidence when the perceptions of participants are sought or when the researcher cannot experience first hand what the interviewee has experienced. Rightly, as with any evidence, there are concerns about the quality. Interviews can be seen to be second hand evidence, filtered through language, which provides evidence that is not easily open to repetition and experiment. Those predisposed to the idea that the researcher is somehow more intelligent, unbiased or more appropriately primed to interpret first hand experiences, are often suspicious of interviewees as their research instrument. [15]

Comparison can and has been used to provide some quality in interviews. There would seem to be several options depending on the situation and audience. Some comparisons are expected to be more convincing than others. A starting

place for thinking about comparisons with interviews may be to try to compare what is said by the interviewee with the some notion of "truth". While clearly this is an epistemological issue, there is a "common sense" approach of believing that a physical world exists. Interviewees as witness should be able to accurately or correctly report physical events. For example an eyewitness to some event can be asked "what, where, when, who and how often" questions and their answers can be judged as to whether they are "correct". They will appear correct if they *compare* to some standard; a reliable source. With varying degrees of power to convince, the options include:

1. Comparison to an instrument reading, such as a video recording, when care has been taken that the instrument is reliable.
2. Similar reports of other eyewitnesses not believed to be involved in any form of conspiracy; a majority report.
3. A logical comparison of the eye witness reports to the remaining physical evidence. For example, an interviewee reporting that a glass fell would be logical given that a broken glass was later found on the floor.
4. If the eye witness interviewee had little reason to lie or was known for his or her honesty. [16]

The problem of comparisons becomes a little more complicated when the interviewee is being used as an expert witness, being asked to evaluate or interpret "why" an event occurred. For example, a manager may be asked in his or her experience why some human system is not working and what might be done about it. The comparisons above may be thought to be insufficient to be convincing. Additional comparisons may be required such as:

1. Checking that the interviewee's interpretation remains stable over a period of time.
2. Checking that the interviewee's actions are consistent with what they say.
3. Comparing alternative explanations and asking interviewees to counter these.
4. Checking if the interviewee's what? is consistent across analogous situations.
5. Checking to see if their statement is internally consistent and well justified. [17]

The exception to these may be if the purpose of the interview is to seek a novel insight or some imaginative response. Here the comparison comes in the form of how different what is said to what others say. Does their interpretation provide a novel view, rather than restate a well-known interpretation? [18]

There is the a "reverse-side" to this issue of whether the interviewer's report of what was said by the interviewees is considered to have been reported accurately and/or interpreted with insight. The interviewer needs to make sense of the interview in a convincing way. Recordings of the interview, and the use of two interviewers, are common ways to use comparison to convince someone of quality. Another is to use an explicit a priori frame (lens, perspective, theory or viewpoint) to "sense make" (WEICK, 1995) what is taken form an interview. For

example, a study of women's issues that uses interviews may use feminist theory to interpret, to compare, what is said in that interview. [19]

3.2 Observation

Observational empirics, evidence through the eyes, are the mainstay of science. In social science, purely observational evidence usually involves the researcher personally observing peoples' actions, behaviours or artefacts through their own eyes or through some instrument that helps their eye in the way that a microscope or video does. Measurement can be involved but this sort of study usually excludes actually talking with those under study. The comparisons that might make observational evidence more convincing include the use of:

1. Recording devices. Care would have to be taken that they had not been edited.
2. "Before and after" comparisons of some event.
3. Measurement (compare actions to standardised weights and measures).
4. An explicit upfront conceptual frame against which predictions of actions are compared with actual activity.
5. Comparisons of what is seen with what others say happened.
6. The provision of relevant context for what is observed. [20]

3.3 Applications

BARLEY (1986) compares the change in organisational routines before and after the introduction of a new X-ray machine in two hospitals. He presents a convincing argument that the different personnel at the two hospitals evolve different organisational routines around their personal characteristics despite the technology being identical. In one hospital the radiographers were keen to be hands-on involved with the machine. In the other, they were not. The "before-after" comparison is not as revealing as the between-hospital comparison. This sort of research also implicitly draws on the experience of the audience, to compare their own work-related experiences to those presented by BARLEY. Most experienced managers have noticed the impact on work routines due to variations in staffing. [21]

MARKUS (1994) presents a convincing argument that email technology merely enables interpersonal preferences to be expressed. It is not inherently a rude means of communication. She does this by comparing interviewees' reflections against each other and against technological determinism arguments. She compares what interviewees say with the complications revealed by setting up the two extreme views that either invented technology is forced on an unwilling society or that technology is developed in response to a social need. The result is a convincing finding that reveals the complexity of the situation. [22]

It may be useful to look in more detail at a short empirical research report. HUNTER and CHANDLER (1999) present the results of a qualitative study into "resilience" in adolescents in an inner city vocational high school. They report use of a standard psychological Resiliency Scale where more is better; "resilience is a healthy and positive state". The authors test this assumption using a qualitative research method. They set their research question as "What does resilience mean to adolescents?" and first administered the Scale questionnaire. They then set the 51 participating tenth- and eleventh-grade high-risk students a writing exercise. This consisted of spending ten minutes writing down their responses to nine daily stimuli, with the intent of getting participants to tell their stories about surviving and overcoming adversity in their life. After individual writing, the participants read out and discussed their stories with other participants. The researchers kept daily journals to record their perceptions of the experience. Their research report conclusion says: "Resilience is adolescences in inner city vocational high schools was a process of defence using tactics such as insulation, isolation, disconnecting, denial, and aggression or was as process of process of survival using responses such as violence" (HUNTER & CHANDLER, 1999, p.247). [23]

They also comment that the writing and storytelling were therapeutic for the children, it made them reflect on their situation and response behaviour. [24]

As expected, there is a lot of use of comparison in this research. The scale is compared with the written comments. Qualitative is compared to quantitative. The participants are compared to other schoolchildren in terms of age, school location, social risk, and location. The researchers are compared in terms of their dairy entries. The whole research report is to be compared to other findings on resilience. They found different results from previous researchers. Their results are ironic because resilience reflects both vulnerability and survivability. [25]

This paper has been arguing that empirical research might be judged in terms of what comparisons it includes. The Resilience Scale research may be considered convincing because its design captures numbers thus allowing individuals to be compared against sample distributions. It is unconvincing because it collects simplistic comparisons; simple answers from simple questions. It is also unconvincing because it does not compare different interpretations of what resilience means to different people; it assumes the researcher has the helicopter view of resilience. The qualitative research is convincing because it allows the researchers to learn about the relevant issues from those with the experience. It is unconvincing, to some, because averages cannot be calculated and individuals compared. [26]

4. In Summation

This paper has presented evidence in support of the argument that empirics can be perceived in terms of comparison. Measurement can be seen as a comparison with international standards of weights and measures, interviews can be seen as a comparison against alternative experiences, and observation as a comparison with a theory or other observations. Comparison seems to provide an alternative means of thinking about how to improve the quality of qualitative research, one that might be recognised across epistemologies. Quality becomes an issue of the credibility of the comparison. Qualitative researchers, operating under this comparison view of research, would need to design and present their empirics so as to maximise the quality and explicitness of the comparisons in a manner that is convincing to a knowledgeable audience. While not every epistemology will accept comparison as objective or critical, hopefully there is enough common ground in the idea for it to be increasingly considered as a viewpoint from which to think about research quality. [27]

References

Ackoff, Russell (2000). *Making a Difference, Systems Thinking—Systems Change*, http://www.judgelink.org/Presentations/GirlsLink/ (Accessed December 17, 2004 [restricted access, FQS, August 2005]).

Alvesson, Mats & Sköldberg, Kaj (2000). *Reflexive Methodology*. California: Sage.

Argyris, Chris & Schön, Donald (1978). *Theory In Practice*. San Francisco: Jossey-Bass.

Barley, Stephen (1986). Technology as an Occasion for Structuring. *Administrative Science Quarterly, 31*, 78-108.

Bateson, Gregory (1979). *Mind and Nature*. New York: E.P. Dutton.

Chalmers, Alan (1982). *What is this Thing Called Science*. Brisbane: University of Queensland Press.

Chamberlin, Thomas (1965). The Method of Multiple Hypotheses. *Science, 148*, 754-759.

Dewey, John (1910). *How We Think.* NY: Dover.

Festinger, Leon (1954). A theory of social comparison processes. *Human Relations*, *7*(1),17-40.

Gubrium, Jaber & Holstein, James (Eds.) (2000). *Handbook of Interview Research: Context & Method*. California: Sage.

Hatch, Mary Jo (1997). Irony and the Social Construction of Contradiction. *Organization Science*, *8*(3), 275-288.

Hunter, Anita & Chandler Genevieve (1999). Resilience was not necessarily a healthy concept. *Journal of Nursing Scholarship*, *31*(3), 243-247.

James William (1907/1910). *Pragmatism*. Cleveland: World Publishing (Meridian).

Lakoff, George (1993). The Contemporary Theory of Metaphor. In Ortony Andrew (Ed.), *Metaphor and Thought* (pp.202-251). Cambridge: Cambridge University Press.

Latour, Bruno (1986). Visualization and Cognition. *Knowledge and Society,* 6, 1-40.

Laucken, Uwe (2004). Quality Criteria as Instruments for Political Control of Sciences [83 paragraphs]. Forum Qualitative Sozialforschung / *Forum: Qualitative Social Research* [On-line Journal], *3*(1), Art. 6, http://www.qualitative-research.net/fqs-texte/1-02/1-02laucken-e.htm (Accessed December 17, 2004).

Luhmann, Niklas (1995). *Social Systems*. Stanford, CA: Stanford University Press.

Markus, Lynne (1994). Finding a Happy Medium: Explaining the Negative Effects of Electronic Communication on Social Life at Work. *ACM Transactions on Information Systems*, *12*(2), 119-149.

Missimer, Connie (1995). The Case that Alternative Argumentation Drives the Growth of Knowledge —Some Preliminary Evidence. *Informal Logic*, *17*(2), 201-211.

Morgan, Gareth (1986). *Images of Organisations*. California: Sage.

Oppermann, Martin (2000). Triangulation—A Methodological Discussion. *International Journal of Tourism Research*, *2*, 141-146.

Rehg, William (2004). *Assessing the Cogency of Arguments: Three Kinds of Merits*. Chicago: National Communication Association.

Rorty, Richard (1989). *Contingency, Irony and Solidarity*. Cambridge: Cambridge University Press.

Van de Ven, Andrew & Poole Marshall (1995). Explaining Development and Change in Organizations. *Academy of Management Review*, *20*(3), 510-540.

Vickers, Gregory (1983). Human Systems Are Different. *Journal of Applied Systems Analysis*, *10*, 3-13.

Weick, Karl (1995). *Sensemaking in Organizations*. Thousand Oaks CA: Sage.

Weick, Karl & Roberts, Kalene (1993). Collective Mind in Organizations: Heedful Interrelating on Flight Decks. *Administrative Science Quarterly*, *38*, 357-381.

Author

Mike METCALFE is an associate research professor at the University of South Australia, where his main duties are as PhD thesis adviser. His own PhD is from Adelaide University, on group problem solving. He has published 5 books and over 50 lead-author refereed academic articles on different aspects of perspectival systems thinking and argumentative inquiry in journals that include *Systems Research and Behavioral Science*, *IT & People*, *Informal Logic* and the *European Journal of Information Systems*. He has lived in over 6 countries including 17 years in Australasia, and has worked in the merchant navy, the army reserves, system design, university lecturing and consulting, including a term as advisor to the Deputy Premier and Treasurer of South Australia.

Contact:

Mike Metcalfe

School Of Management, University of South Australia (City West)
Adelaide, South Australia, 5000

E-mail: mike.metcalfe@unisa.edu.au

Citation

Metcalfe, Mike (2004). Empirics as Comparisons [27 paragraphs]. *Forum Qualitative Sozialforschung / Forum: Qualitative Social Research*, *6*(1), Art. 27, http://nbn-resolving.de/urn:nbn:de:0114-fqs0501270.

Revised 6/2008