Groups and Indicators in Post-Industrial Society

Authors:
Dr S. Bell
Communication and Systems Department
Faculty of Mathematics
Computing and Technology
Open University
Milton Keynes MK7 6AA
UK
Email: s.g.bell@open.ac.uk

Professor S. Morse
Department of Geography
School of Human and Environmental Sciences,
The University of Reading,
Reading RG6 6AB
UK
Email: s.morse@reading.ac.uk

Abstract
Indicators of all kinds define our world. We are constantly measured and assessed. Perhaps the most important indicator in current use is Gross Domestic Product or GDP. It is the measure of a Nation's success and can be key to its ability to borrow money and appear internationally credible. This paper is set against the current debate on 'Beyond GDP' began in November 2007 with the conference hosted by the European Commission, European Parliament, Club of Rome, OECD and WWF. The initiative, with its five actions, recognises weaknesses in the ways in which indicators of all kinds are collected and presented and attempts to improve the indicator world but, is the answer to effective information for policy formulation hidden in the articulation of indicators? Maybe, indicator use is a function of the ways in which stakeholders are engaged in their use? Our conjecture is that indicator use is little understood and that this use dynamic can be better understood.
Following the articulation of some hypotheses concerning indicator use, the authors write from the perspective of their work undertaken in the European Union funded Framework 7 project 'Policy Influence of Indicators' (POINT; contract no 217207) which began in 2008. A major element of the project involved a number of group workshops designed to elicit viewpoints regarding the use of indicators (including sustainable development indicators) in sustainable development policy at EU and member-state levels.
The paper outlines some responses to the hypotheses, hints at how group approaches to indicators can be foreseen and some challenges for indicator use policy for the future. This may well impact on the way in which commonly applied indicators like GDP are produced and experienced.

Keywords: Beyond GDP, stakeholder engagement, Triple Task Method

"In early 2009, John Beddington, chief science advisor to the UK government , said the world was facing a "perfect storm" of food shortages, water scarcity, and costly oil by 2030 …. A week later, Jonathon Porritt … wrote in the Guardian that he agreed with Beddington's analysis but that the timing was off. He thinks the crisis will "will hit much closer to 2020 than 2030"." (Brown 2011)

The Indicator World

It is a truism to observe that contemporary citizens are surrounded by and, to some extent governed by, indicators. An indicator can be defined from the Latin as indicare: to proclaim or to say – "to point out or to direct to knowledge" (Webster Dictionary). This definition provides enormous room for interpretation and development and, over the last 100 years many agencies and powerful groups have made good use of this room. An indicator can be a quantitative signal as, for example, the value of gold bullion in the Nation's reserves as an indicator of national wealth or as qualitative as the outcome of a focus group exploring the use of a new kind of olive oil or the kind of authoritative quote set out from Brown's book above. Indicators are widely reported in this journal (for a recent selection see: Lyytimaki and Rosenstrom 2008; Coelho, Mascarenhas, Vaz, Dores and Ramos 2010; Hermansen 2010). Indicators 'direct us to knowledge' and this direction can be conscious and planned (as in the case of career defining exam results) or unconsciously absorbed (for example the tacit adoption of information about the results of football matches, road news or the weather). The indicator world is vast and constantly changing.

But it is also a truism of sorts that indicators have to be useful. This is not necessarily so, of course, as indicators can, at least in theory, serve no purpose whatsoever. They can be defined in terms of methodology and presentation yet exist as nothing more than part of a long list at the end of a report which no one, other than the author, looks at. But most would agree that indicators have to have a purpose; they have to fulfil a perceived need. Either to help monitor the effectiveness of an intervention, for example a policy, or perhaps even to help frame that intervention (Morse, 2004). This may sound obvious but therein rests a tale. After all, even the brief sentence above mask all sorts of nuances and indeed subjectivity. Who are these indicators useful for and what exactly is meant by 'useful'? The reader is invited to consider any intervention they are familiar with such as a development or research project, or indeed policy formulation and implementation, and the types of indicators (quantitative and qualitative) that had to be employed to monitor key aspects of achievement. Maybe the intervention was subject to a detailed 'blue print' such as a planning framework where the indicators are linked to performance targets to be reached by specific times. Whatever these are the reader is invited to consider alternative indicators that could have been chosen or ways in which the indicators could have been assessed differently. There are always alternatives and phrases such as the 'best indicators' or the 'most suitable' indicators are subject to a great deal of value judgement.

However, despite its importance the use or influence of indicators and how best to enhance this in any particular context has received almost no attention in the academic literature. There are some notable exceptions to this statement and the work of Hezri and his colleagues provides an example for Malaysia (Hezri, 2004, 2005; Hezri and Dovers, 2006).

In this paper we seek to explore the contemporary use of indicators, expound on some hypotheses and reflect back on research undertaken. The flow of the paper is shown in Figure 1.

<Insert Figure 1 about here>

Indicators in perspective

Based upon the work of Hezri and others and indeed experience gained over many years working in sustainable development projects spanning many countries it is possible to conceive a number of possible influences on the influence of indicators. For example, it seems reasonable to assume that the most 'used' indicators will be those that are relevant and hence match a need. All may not necessarily agree with that 'need', but if the indicator is in demand by a stakeholder then that indicator is more likely tom continue to exist. Indeed even if the indicator is very technical and difficult to estimate if the need is strong enough then it will continue to exist if demand is high. Those demanding the indicators may not necessarily know how it is calculated or its limitations but as long as the indicator helps them to encapsulate complexity for someone else to understand then they will be of use.

A good example of an indicator that has a high demand amongst at least some groups is the Gross Domestic Product (GDP). This indicator is at present very dominant in the planning and defending of economic policies. Simply put, a tendency has been for countries to try and maximise their GDP (as an indicator of wealth). However, even GDP is questioned in terms of its use. GDP is positively related to population size; countries with larger populations tend to have bigger GDPs relative to countries with smaller populations. This can be adjusted by dividing GDP by population size so as to generate a value of GDP/capita, but this in turn raises very big issues surrounding the equitable distribution of wealth. There is also the danger that the relative ease at which GDP can be measured along with its association with something as politically sensitive as national wealth can result in dominance of that indicator above many others. Indeed the debate about use of Indicators in the contemporary world raised in 2007 with the conference hosted by the European Commission, European Parliament, Club of Rome, OECD and WWF. Arising from this conference emerged the report (Commission_of_the_European_Communities 2009). The report underlined the vital importance of GDP but also its limitations:

"GDP has also come to be regarded as a proxy indicator for overall societal development and progress in general. However, by design and purpose, it cannot be relied upon to inform policy debates on all issues. Critically, GDP does not measure environmental sustainability or social inclusion and these limitations need to be taken into account when using it in policy analysis and debates."
(Commission_of_the_European_Communities 2009 page 2)

This quote from the Commission points to the way in which GDP has been used as a bell weather for society as a whole. It is applied as a means to tell not just policy makers, but citizens in general, how well a society is functioning. Crudely put,

Good GDP = Healthy Society.

But GDP does nothing more than measure how much money is being circulated in an economy. In order to do this people have to go out and measure how much money is flowing either at the point of earnings or – perhaps more usually – at the points of expenditure. Complications arise as a result of inflation (change in purchasing power of currency over time) and how to assess 'unrecorded' or hidden expenditures (via the 'black market' for example). Thus it measures a very narrow aspect of human existence and aspiration but even so, the use of this indicator is a reminder of how powerful they can be. GDP dominates the landscape of economic policy making at regional, national and even sub-national scales. Politicians typically manage the economy so as to maximise economic growth (typically assessed as a change in GDP over time) as this is assumed to bring employment and prosperity to at least some of the people who have an influence on the survival of those politicians. But such a narrow focus on economic growth can be detrimental as some parts of the world have witnessed with the recent recession and its continuing ramifications, especially in some of the smaller European economies which have adopted the Euro.. Of course it is not right to 'shoot the messenger' and GDP is primarily a means to provide important information for those meant to manage an economy, but the use of the message, its political importance and the way in which this imposes social orders on society is important to understand. After all it is people who both demand and use indicators. Nonetheless the fact that GDP tells nothing about physical or emotional health, environmental health, human happiness or moral and emotional balance (all features which people regard as being central to balanced development) appears lost in focus on an apparent economic imperative.

In order to address the perceived weakness of GDP the Commission suggested five measures:

Complementing GDP with environmental and social indicators
Provision of near real-time information for decision makers
More accurate reporting on distribution and inequalities
Developing a European Sustainable Development Scorecard
Extending National Accounts to environmental and social issues

The fact that this kind of activity is being undertaken by such a major agency as the European Commission adds weight to the impression that indicators, even the most important indicators, are not fulfilling the task expected of them – primarily to "point out or direct to" policy. But whose fault is that? Is it caused by demand or supply of indicators or perhaps a combination of the two? After all, alternatives to GDP already exist. For example, the Environmentally Adjusted Net Domestic Product (EDP) as part of a move towards more 'green accounting'. For a recent review of these 'green' economic indicators in the context of sustainable development please see Morse (2010). However, in contemporary 'green' politics the "Ecological Footprint" is

perhaps one of the most notable examples as it signifies a sense of consumption and can be used to imply greed or prudence (Siche et al., 2008;
Venetoulis and Talberth, 2008). However, it has to be noted that the EF is certainly not without its critics (Fiala, 2008). In measuring production indicators such as Maximum Sustainable Yield and Total Factor Productivity (also discussed in Morse, 2010) have become prominent but post World War 2 economics has highlighted the importance of GDP and its related family of similar indices.

Indicators do not comprehend nuance or discretion in their expression. As already suggested, they can intrude into parts of human life that are implicitly value-laden. Measuring immeasurable human qualities from populations and making generalisations about key concerns such as health, wealth, education and even happiness. Trying to measure something which is not universally constant in terms of meaning would seem to be highly problematic. Yet we now have indicators that assess 'quality of life' and 'wellbeing'. It is hard to identify filaments of human activity into which indicators have not intruded.

Indicators and understanding the 'human project'

It can be argued that indicators come into prominence when the personal and subjective knowledge of a 'thing' breaks down due to magnitude and the 'thing', whatever it may be; youth, an education system, a police force or an export drive for example, required to be known at some level of objective and generalised clarity (for a general discussion on this see: Verbruggen and Kuik 1991). In an era of evidence based policy[1] these issues often go hand-in-hand with a need to know whether the resources being used to improve something are effectively achieving impact. Put simply, is the spend worth the outcome? Has investment delivered desired change and if so then by how much? Given all of this, with objectivity being key, indicators and their related algorithms have emerged as an important way for 'us' to know 'it', but the key here is: who is 'us', what is 'it', and how do we 'know'? These three terms bring out the nexus of the issues around indicators.

In generalised and anonymous assessment made possible by indicators, the behaviour of more than 60 million people is condensed into a single measure of GDP for the UK. 'My' behaviour is lost within this; 'I' become but a very small part indeed of this aggregation. Also, most people have heard of GDP and perhaps have a somewhat vague idea as to what it is but relatively few would really understand what it is and be able to question how it is assessed. Those who do know are the owners of indicator and must, as a consequence, have a preferential facility and use of them. Here we may have a new and important and possibly worrying community. Suddenly, world 'facts'

[1] Heralded by papers such as that of Philip Davies, Government Chief Policy Social Researcher, UK Prime Minister's Strategy Unit, given in 2004: "Is Evidence Based Government Possible?".

become the province of the generalised indicator. It is in this post-industrial age that 'I', as a citizen of the UK, can see the GDP of Greece (with or, more likely, without any comprehension as to what this means) which allows the international community of policy makers in the European Union and the International Monetary Fund to impose economic austerity upon the population of that country. In this scenario, the wider population watches on, convinced by 'evidence' which may or may not be understood but possibly fearful of when the consequences of indicators like GDP may be cast in its direction.

Post-industrial society needs to understand itself in order to govern itself. Big society needs big tools, and the key tool used to-date is GDP. But and this is an important but, even with the best tools, the way they are applied – in short our understanding of who needs to know what in order to do something, is far from perfectly understood (eloquently aruged in: Chambers 1997). Indicators are a cultural artefact of our times. Lee may have intimated it in 'the Wired Nation' in 1972 (Lee 1972). Traber assessed myths and realities in 'the Myth of the Information Revolution' in 1986 (Traber 1986) and pointed to the need for localised communication to be the prize of post-industrial society rather than technological fizz in the northern economies. But there are ethical dimensions to indiscriminate indicator-hegemony. A significant problem is that indicators are a reflection of what a particular context deems to be important. They are intellectual constructs not Laws of Nature. Thus the indicators which measure oil spillage, volcanic ash density and fish stocks can all be contested. This means that the application of an indicator is the beginning and not the end of a process. The indicator tells us something about it. We then do something about it. These are easy words which have manifold consequences.

As a metaphor, the development of indicators can be compared to the development of the telescope by Galileo. He was developing a technology to understand the universe – ambitious and outrageous, or so it was thought at the time. Today policy makers are developing an indicator technology in an attempt to understand just about everything. The telescope of Galileo was crude and shown to need massive and large scale improvement. It showed a fraction of the universe and this in very poor quality. Our use of indicators, our under-standa-scope, may well be shown by history to be similarly crude. However, while the world of indicators may be messy it is nonetheless important and needs to be studied.

Understanding the 'human project' and indicators – some hypotheses

Indicators should be a tool to aid in evidence-based policy but do they simplify complexity too much and thus become misleading? What is the evidence that they have any influence at all amongst those meant to 'use' them? These are perhaps obvious questions but remarkably they have received very little attention amongst researchers. Instead the overwhelming focus has been upon technical issues of producing 'good' indicators and coming up with some means by which they can be presented. The need to better understand the ways in which indicators are produced, experienced and applied in the sense of influence policy and management was the underlying rationale behind the EU Framework 7 funded project called POINT (policy influence of indicators; contract no 217207) which began in 2008. The project was primarily interested in assessing how indicators are used by policy makers. What is important? Which indicators are commonly used? How are they applied? What is

their value? Thus the project set out to address the very questions and issues that have been alluded to in the earlier sections of the paper.

This work into understanding why and how indicators of all kinds are used by policy makers and members of the general public had some of its methodological roots in the earlier work of Bell and Morse in the emergence of indicators in coastal societies in the Mediterranean (Bell and Morse 2001; Bell and Morse 2002; Bell and Morse 2003; Bell and Morse 2003) and in still earlier work in the participatory development of sustainability indicators (Chambers 1992; Bell 1996; Levin 1997; Reckers 1997).

A major element of the project involved a number of participatory stakeholder workshops held across the EU, designed to elicit viewpoints regarding the use of indicators in sustainable development policy at EU and member-state levels. But, we purely exploring indicator use. We also wanted to explore more intimate social and psychological issues behind and informing use. To do this we developed a unique participatory methodology called the Triple Task.

Following an early team meeting, certain possible hypotheses emerged from group discussion. As set out earlier in this paper it is relatively straightforward to set a series of conjectures that one would expect to exist in terms of the use/influence of indicators. Distinguishing between use and influence in this context introduces much scope for diversity in interpretation. Here it is assumed that use implies a direct and presumably identifiable cause-effect of some form. Thus an economy is managed to maximise economic performance and the GDP should increase as a result. If is doesn't then something has gone wrong with the management and the assumptions that rested behind it. Influence is a much vaguer term but may be important nonetheless. It is not hard to conceive of indicators that don't have a direct use akin to GDP but may still have altered the thinking of those making an intervention happen. An example here may be the Ecological Footprint (EF) and its association with consumption of resources. The EF was not created to meet a demand from politicians but the concept has been influential nonetheless and these days indicators based upon carbon footprints are at least part of the rhetoric of many governments.

Given the above the hypotheses were intended to be tentative but also provocative for future research:

The use of an indicator decreases in proportion to its perceived lack of relevance
Indicators work if they match stakeholder requirement
Quantitative indicators can be just as attractive as qualitative indicators ... if assumption 2 is also true
Flashy presentation of an indicator is no guarantee of popularity
Dry presentation of an indicator is no guarantee of a lack of popular appeal
All people use indicators - most do not know that this is the case
People who need to use indicators often are ignorant as to their value
Few indicators are so powerful as to be able to 'find their way' without a degree of marketing
All indicators have a shelf life determined by assumption 2

These nine hypotheses pivot around the number 2 – the need to make sure that indicators match what is needed. This may sound obvious, but in our work we did not

assume that the stakeholder in indicator use is unambiguous, that the knowledge required by the stakeholder of the indicator is clear or that the needs for indictors to support the stakeholder group (who ever that is) was uncontested. Please note that the reader is encouraged to consider this list and to disagree if they so wish with any of the assumptions. Part of the intention here is to help encourage a discourse.

Predicting Indicator Use Dynamic –Triple Task Method

As already noted, a key innovation of Work package 6 of POINT was the development of an approach to assess how groups react to a variety of influences .. in the case of this research to issues related to indicator use. The method is called Triple Task and is variously referred to elsewhere (Bell and Morse 2009; Bell and Morse 2010).

What is Triple Task?

Triple Task (TT) is a form of participatory action research and Stakeholder analysis in the sense that it attempts to arrive at answers to research questions but also tries to understand what factors may have been at play in arriving at those answers. This attribute makes TT different to many other participatory techniques which are focussed on delivering outputs (e.g. representing an apparent 'consensus') and less concerned on the dynamic behind that 'consensus' and how the process may have influenced what was produced.

Without wishing to go into the details of TT it is nonetheless important for the reader to gain an understanding of the processes at play and why they are in place. TT involves three related tasks. The core of the methodology is referred to as 'Task 1' and this generates answers to research questions such as, in the case of POINT, 'what are the policy influences of indicators?' Participants are divided into groups of perhaps 5 or 6 people working for a day and a half to two days and each group generates a shared understanding as to its answers to the question. Task 1 takes the form of a systems approach by using elements of what is called 'soft systems'. The process follows a logical sequence of scoping → focus on key issues/tasks that need to be addressed → what is required for implementation. For convenience, Task 1 is subdivided into three main steps as set out below:

Scoping: Diagrams known as Rich pictures are employed as a means to capture 'stories' from participants. Participants are encouraged to draw out major tasks and issues which form a central concern to them. These are then organised in terms of precedent and priority. Groups of linked tasks and issues are 'clustered' into indicative systems of concern (Systems of Challenges; SoCs). This systemic process binds the group together, forges collective understanding and provides a legitimising process of further discovery.

Visions of Change: Moving from a shared understanding as to the challenges this step encourages the groups to explore what changes are required in order to address the SoCs. In other words, what needs to be done? Groups may derive a number of VoCs rather than only one, but the emphasis should be upon what the group deems to be more important and achievable.

Desired change: Groups encouraged to set out what practical steps are required to bring about their Vision of Change. This step is supplemented by activity planning and scenario setting: 'How might things look given certain kinds of change?' The

latter employs another Rich Picture; providing a sort of 'before' and 'after' story when placed next to the rich picture that arose out of Step 1. Participants not only enrich their own understanding of what is possible but act as vectors of change for colleagues.

By following this process to its conclusion Task 1 provides the insights with regard to the research questions (what has been done, by whom, why, how is this assessed in terms of effectiveness?). These insights can then be fed into a policy or managerial process.

How do the second and third aspects of Triple Task Work?

Tasks 2 and 3 run in parallel with Task 1 and are designed to explore the ways in which the groups function (as social and individual entities) and how this influences their use of indicators both in terms of what emerges under Task 1 but also in terms of the variation one might see between members of the group and how they are able to influence the dynamic. Thus Tasks 2 and 3 interlock with Task 1 and attempt to provide at least a partial explanation as to why groups have arrived at their conclusions. From the perspective of participants they only experience Task 1; Tasks 2 and 3 are largely invisible to them. They don't see the results of Tasks 2 and 3 as these are analyses for the researchers, and neither do they need to know what the results of Tasks 2 and 3 are for them to complete Task 1. Tasks 2 and 3 are summarised as follows:

Task 2: This is an 'outside in' review of the group dynamic. In effect it is the researcher's assessment of the group process using a matrix approach originally developed at the Open University and known as BECM (used in, for example, the Open University Course: 'Managing Complexity: a systems approach'.). BECM stands for Being, Engaging, Contextualising and Management. BECM can be used as a form of Socio-Analysis and is related to the psychoanalytic tradition (Bell and Morse 2011).

Task 3: 'inside out' review of the group dynamic – stakeholders' assessment of their group process employing the Symlog methodology (Keyton and Wall 1989; Hurley 1991; Blumberg 2006).

Thus Tasks 2 and 3 are intended to be complimentary. One is based upon the facilitators perspective of the groups by observing their behaviour and 'body language' while the other is the groups view of itself, and is more nuanced in the sense that it is also founded upon discussions within the group which an outsider would not have access to. Thus Task 2 would be expected to have only a partial correlation with Task 3, but a correlation would be expected nonetheless.

How is Triple Task Assessed?

TT is an holistic tool with each of the Tasks relating to the other (from a researcher perspective). In the case described in this paper, a group is asked: "What are your thoughts on indicator use". They are then asked to explore this question by use of the stages of Task 1, their group dynamic is examined by the Researchers by means of Task 2 and their own reflections on their personal experience and their experience as

members of a group is captured in Task 3. Thus TT provides for an internal analysis of group outputs and dynamics which can be explored further by looking for patterns relating to other influences such as international, national, sectoral, institutional contexts. One means of looking for such patterns is to plot the results of the three tasks onto what is called a TT Field Diagram. To do this we essentially quantified the various outcomes.

Outcomes and insights from Triple Task and indicator use

POINT included participatory workshops with a number of groups of people having an interest in indicator use from Malta, Slovakia, Finland, Denmark, UK and Belgium. and Finland The participants were selected locally, we had no say in the representation within each group however, we provided each organiser within country with guidelines on the balance of groups (ensuring that groups contained a mixed number of technical experts, practitioners of various degrees of expertise and interested lay people). Groups represented the range of domains under focus in POINT; i.e. agriculture, transport, sustainable development and energy. For the purposes of this paper we will only present the results from the groups who focussed on indicator use in sustainable development; two from Malta, three each from Slovakia and Finland. The specific sectors of transport, energy and agriculture, all of which are important in sustainable development of course, have been omitted as they are subject to a specific policy context. The results of the sustainable development indicator workshops can be summarised as a TT field diagram (Figure 2) which has a number of elements. In each case (Figure 2a) the horizontal axis represents the quality of the outputs generated by each group (the quality of the analysis they arrived at in Task 1), while the vertical axis represents the quality of group function as gleaned from Task 3 (Symlog; high functioning groups at the top, lesser functioning groups at bottom). The groups from the POINT workshops are plotted as circles in the field diagram as shown in Figure 2b (Malta – groups A and B; Slovakia – groups C, D and E; Finland – groups I, J and K). A larger and shaded circle shows a higher group functioning as assessed with Task 2 (BECM), while smaller and unshaded circles imply lesser group functioning. One interpretation of the spaces within the field diagram is provided in Figure 2c. These are, in effect, 'metaphor' labels based upon the characteristics one may expect of groups inhabiting that space within the field diagram. Figure 2d provides another interpretation of the space. In this case the groups that are towards the centre of the diagram are pragmatists. Presumably the majority of groups would be expected to appear here as this is the space that suggests a reasonably 'text book' type of analysis between indicators and use. Those outside of that centre can be considered to be 'mavericks' in the sense that there analysis and group function have been out of the ordinary. Maybe the group was especially dysfunction or indeed functional, or maybe the analysis was unusually poor or good. The remaining elements of Figure 2 (e,f and g) provide a range of possible linkages between group function and the quality of analysis. Perhaps the most logical assumption is shown as Figure 2e which suggests that as group function gets better then so should the quality of the outputs (analysis) they arrive at. An alternative possibility is provided by Figure 2f where the assumption is that the quality of inputs improves with worsening group function. This might sound odd but it is quite possible that groups with a greater degree of friction could produce the most novel insights; the friction may be a reflection of intensive discussion. The final hypothesis is presented as Figure 2g and is in effect a mixture of hypotheses (e) and (f). Increase in output

quality with group function followed by a 'switch-over' when output quality improves with worsening group function. It is possible that comfortable groups may exist which shown good group dynamic in the sense that everyone agrees on the analysis but this is only because everyone is repeating what the textbooks say. At the other extreme there may be groups having the characteristics suggested by Figure 2f.

Insert Figure 2 about here

Therefore it has to be admitted that the pattern of group plots one can expect to see within the TT field diagram (if indeed there are any) is a matter of conjecture. The simplest position is to expect an increase in quality of output with 'better' group function as shown in Figure 2e but does this assumption hold true? It certainly does not seem to be the case for plots of the POINT groups but why is so?

One of the obvious points to make about Figure 2 is that while all the groups were asked to dissect and explore indicator use in sustainable development they obviously differed in being from 3 countries. Finland has relatively extensive experience of trying to make sustainable development a reality and indicators have long been seen as a part of this. Hence those involved in the Finland workshop (I, J and K) were able to draw upon this experience with indicators and provide critical insights. The varied experience of those involved in the workshop also aided constructive and indeed contested dissection of issues within groups. Thus while their group functioning may not have necessarily been all that good (points for groups I and K are at the bottom end of the vertical scale) their outputs were excellent in terms of the richness of insights. By way of contrast the groups from Malta (A, B) and Slovakia (C, D and E), where sustainable development indicators are far less ingrained in national policy and management, provided more mechanical analyses (text book answers as to what indicators should be and what they should do) and also tended to talk far more of indicators coming from outside – indeed there was a sense of them being imposed by the EU. The results for the latter tend to be less exciting. But why should this be reflected in group behaviour, or at least group behaviour as perceived by the members of the groups via Symlog? It was the case that two of the groups from Finland showed a lot of animation and individuals were vociferous. Group members clearly had stories to tell from their own experience rather than just conjecture and much emotion was on show; there was tension and friction. In Malta and Slovakia this was - mostly – less evident with groups agreeing far more about the need for indicators and how the EU drives this process. Maybe this reflects the relative lack of personal experience of indicator use within these groups, with the result that participants fell back on what they perceived as accepted wisdom and the need to do what the EU says they should. Indeed with these groups there was much talk about why SD indicators had not had the influence they should have had, including the dominance of economic indicators. The one exception was group E which itself was dominated by an individual who had much experience within sustainable development and indicator use. The dynamic of group E was not good but it did arrive at some insightful analyses. So maybe the dynamics at play are at least in part a function of familiarity which in turn leads to a confidence in being able to critique indicators and their influence. Alternatively, less familiarity leads to a resort to more text book enactments and less disagreement as to what indicators are and what they should do. In effect this is a call not so much for the

model in Figure 2e but that in Figure 2f – better quality outputs with worsening group function! Perhaps it is not a straight line at all but a curve which combines elements of Figure 2e and 2f. But all of this could be a function of many factors besides national experience with sustainable development and indicator use.

Some preliminary reflections

Behind the Triple Task Field Diagram as presented in Figure 2 is the key issue of indicator use. Most of the groups directly addressed the indicator use theme of the workshop head on in Task 1. High performing Task 1 shows a high degree of engagement with the indicator issue. Across our groups however, five emergent themes appear from the research findings:

Indicators are not widely known and understood (more so in countries of long standing in the EU, less so in new accession countries).
Indicator use in policy is vaguely grasped generally
Issues of 'right' and 'wrong' indicators abound
The value of indicator use in policy formulation is questionable.
Known and dominant, universal indicators like GDP, even if not well understood, dominate the indicator landscape.

The TT Field Diagram, as it captures the 'process stories' of the various groups, tells us how some of the groups both used indicators and how they worked as groups. We see that there is no clear link between a 'good group' and 'good work'. Rather there is nuance on what makes a good group (for example, conflictual groups can produce very good results – Groups I and K of Finland) and a good result (remaining questions and lack of closure can indicate a good result just as much as a clearly closed case – Group A from Malta).

Discussion

Indicators are an evident sense making tool available for the governance of post-industrial society. They have been produced to support almost every realm of service provision and global evaluation. They have a position of explicit authority in decision making spheres, however, theory espoused (e.g.: indicators like GDP have explicit authority in decision making spheres) and theory in use (e.g. 'damn the indicators .. we need to follow policy!') can be at variance. The methods used in our project are attempts to understand both the espoused theory of indicator use .. and more importantly the theory in use.

The workshops which we ran show that hidden indicators and indicator use are in evidence, but are under-represented in formal thinking. The indicators in demand by policy makers (such as GDP) do have a strong influence and policy makers engage

and drive indicator evolution in directions that they wish to see. Alternative indicators relating to multiple perspectives – beyond GDP, the multiversal[2] indicators, are not in such demand so by definition the consumers do not engage so much and evolution does not work so strongly. The result is limited or no demand for multiversal indicators. This gives little pressure for change which in turn does not enhance demand and influence of such indicators. A different 'spin' on this, which explains the outcome just as well is that there is a dominant 'neo- classical economic' world of discourse at work. This tells the conventional "story" and provides little or no room for other Worlds of Discourse (WoD) to compete. The multiverse of indicators can be seen as being hostile to the dominant world view of economic determinism which is strongly represented by the GDP narrative. In this analysis, by conspiracy or complacency the multiverse of indicators fails to succeed in production, use or influence.

So how does this connection become established? In our research, groups often mentioned the need for better education and communication on the part of those promoting the indicators multiverse. There is also the interesting comment that democracy is perhaps 'too big'. What the workshop participants meant by this was that democracy forces an indicators universe which reflects the day-to-day concerns of the voters – keeping their jobs, increasing their pay, lowering the cost of living and taxes etc. In this argument democracy is bad for the indicators multiverse – it works against the involvement of policy makers and other consumers such as the media in indicators which voters do not rank high on their concerns at election time (not necessarily the same as what they would say between elections) which in turn restricts the natural selection that would make these indicators 'better' and more responsive to need. Hence the calls by some groups for better education of the electorate to encourage them to consider the multiverse, but that is far easier said than achieved in the current world . However, if our conspiracy model were correct, this would see this as a wasted endeavour – educating the citizenry is pointless if the decision makers are already disinclined to see any reality other than that which is drawn by an economic uni-verse of GDP.

Our Original Hypotheses Revisited

To return to our original hypotheses, we may now be able to suggest some tentative insights:

H1: The use of an indicator decreases in proportion to its perceived lack of relevance

[2] We use the term Multiverse here to mean indicators which represent and accommodate multiple perspectives, the local as well as the international, the non-literate as well as the literate, the un-empowered as well as those in authority.

The responses of all 17 groups would indicate that this would seem to be true. Critically the issue of indicator relevance is clearly demonstrated to be located in the interests and value system of the user.

H2: Indicators work if they match stakeholder requirement

Again this would appear to be tentatively true. No requirement for knowledge = no point to the indicator. This came out time and again in our workshops. The key point here is: who is the stakeholder and what do they think that they need to know?

H3: Quantitative indicators can be just as attractive as qualitative indicators ... if assumption 2 is also true

The evidence of the rise of GDP itself would tend to support this hypothesis. Our research indicates that quantification, even if not personally understood by the user, is seen as a symbol of resilience in the indicator.

H4: Flashy presentation of an indicator is no guarantee of popularity

Certainly true. We found no evidence that indicator presentation led in any way to indicator uptake.

H5: Dry presentation of an indicator is no guarantee of a lack of popular appeal

Ditto the point above.

H6: All people use indicators - most do not know that this is the case

Yes. This hypothesis holds true overwhelmingly. It is often the language of indicators which distracts us from the understanding that indicators of all kinds are in daily and passionate use at all levels of society.

H7: People who need to use indicators often are ignorant as to their value

Yes. This is part of the indicator issue. Very often the information those engaged in our workshops needed was available but sometimes the location of the indicator was not clear, and sometimes the need for the indicator by the stakeholder was not known.

H8: Few indicators are so powerful as to be able to 'find their way' without a degree of marketing

Partially true. Some indicators seem to ‘break through’ and have a life of their own (e.g. GDP now, Ecological Footprint a few years ago).

H9: All indicators have a shelf life determined by assumption 2

This would seem to be very true. The indicator has value for as long as stakeholders find it a useful source of knowledge. We presume that when automobiles began to take over our streets from the previous hegemony of horse drawn carriages, indicators on straw and hay availability at wayside inns would have become redundant.

Conclusion and Recommendations.

Working from the hypotheses and to draw inference and make some really positive suggestions to policy makers could be a successful strategy. Arising from our preliminary hypotheses, suggestions to policy makers could be framed as follows in five premises:

For it to work at a practical and theoretic level, the EU needs to make information provocative, relevant and innovative yet conform to understood ‘rules’. This is because people are more likely to want information which is in their ‘world’ and which conforms to items in their World of Discourse (WoD). If information takes people on a journey of discovery and makes it possible for them to feel that they are really doing break-through thinking in their WoD the investment in the exercise will be rewarded.

Display these rules clearly and show that they can be knocked down by premise 1 so long as the rule is no longer relevant to the WoD in question.
Always leave gaps for people to fill in themselves, thereby leaving their creation in their WoD in place
Bring information together in unexpected combination WoDs thereby allowing WoD planes to collide and be ready to pick up the exciting new WoD outcomes.
Encourage exploration and innovation in all WoD experience relevant to information/ indicator use. In other words, seek to enhance the process of natural selection by encouraging a wide range of indicators consumers.

By empowering citizens and respecting the authenticity of their world view we may be able to provide indicators which reflect this multiverse and, maybe, point to a measurement and policy framework which points far beyond GDP as our primary measure of value.

References

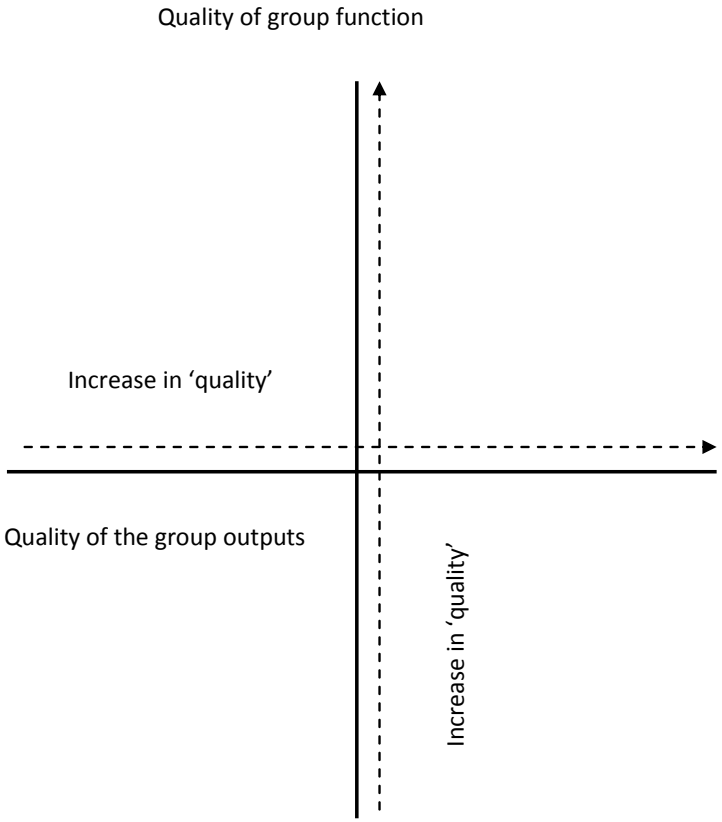
Argyris, C. (1982). Reasoning, Learning and Action. San Francisco., Jossey Bass.
Bell, S. (1996). "Approaches to Participatory Monitoring and Evaluation in Dir District, North West Frontier Province, Pakistan." Systems Practice **Vol. 9**(No. 2): pp. 129-150.
Bell, S. and S. Morse (2001). "Breaking Through the Glass Ceiling: Who really cares about sustainability indicators?" Local Environment **6**(3): 291-309.
Bell, S. and S. Morse (2003). Measuring Sustainability: Learning from Doing. London, Earthscan.
Bell, S. and S. Morse (2010). "Triple Task Method: Systemic, reflective action research." Systemic Practice and Action Research **23**(6): 443 - 452.
Bell, S. and S. Morse (2011). "Being, Engaging, Contextualising and Managing: BECM Matrix – a means to assess group dynamics?" Systems Research and Behavioural Science **Wiley Online**.
Blumberg, H. H. (2006). "A Simplified Version of the SYMLOG (R) Trait Rating Form." Psychological Reports **99**(1): 46 - 50.
Brown, L. (2011). World on the Edge. New York, W W Norton and Co. .
Chambers, R. (1992). Rural Appraisal: rapid, relaxed and participatory. Brighton, Institute of Development Studies.
Chambers, R. (1997). Whose Reality Counts? Putting the first last. London, Intermediate Technology Publications.
Checkland, P. B. and J. Scholes (1990). Soft Systems Methodology in Action. Chichester, Wiley.
Coelho, P., A. Mascarenhas, P. Vaz, A. Dores and T. Ramos (2010). "A framework for regional sustainability assessment: developing indicators for a Portuguese region." Sustainable Development **18**(4): 211 - 219.
Commission_of_the_European_Communities (2009). Communication from the Commission to the Council and the European Parliament: GDP and Beyond - measuring progress in a changing world. E. Commission. Brussels, Commission of the European Communities.
Fiala, N. (2008). Measuring sustainability: Why the ecological footprint is bad economics and bad environmental science. Ecological Economics. 67(4), 519-525

Hermansen, J. E. (2010). "Mediation of tropical forest interests through empowerment to locals by means of ecological indicators." Sustainable Development **18**(5): 271 - 281.
Hezri AA (2004). Sustainability indicator system and policy processes in Malaysia: a framework for utilisation and learning. Journal of Environmental Management 73, 357-371
Hezri AA (2005). Utilisation of sustainability indicators and impact through policy learning in the Malaysian policy processes. Journal of Environmental Assessment Policy and Management 7(4), 575-595
Hezri, A.A. and Dovers, S.R. (2006). Sustainability indicators, policy and governance: Issues for ecological economics. Ecological Economics. 60(1), 86-99
Hurley, J. R. (1991). "Self-Acceptance, Acceptance of Others, and SYMLOG: Equivalent measures of the two central interpersonal dimensions?" Journal of Clinical Psychology **47**(4): 576 - 582.
Keyton, J. and V. D. J. Wall (1989). "Symlog. Theory and Method of Measuring Group and Organisational Communication." Management Communication Quarterly **2**(4): 544 - 567.
Lee, R. (1972). The Wired Nation: Cable TV: the electronic commuincaitons highway. London, Harper and Row.
Levin, M. (1997). The Quest for Quality in Participatory Inquiry. Forum 2: Action research and Critical Systems Thinking, Hull, Centre for Systems Studies, University of Hull.
Lyytimaki, J. and U. Rosenstrom (2008). "Skeletons out of the closet: effectiveness of conceptual frameworks for communicating sustainable development indicators." Sustainable Development **16**(5): 301 - 313.
McQueen, K. (2003). "Improving Team Teaching in the Earth Sciences through an Action Learning Approach." Planet(10): 24-29.
Morse, S. (2004). Indices and indicators in development. An unhealthy obsession with numbers. Earthscan, London.
Morse, S. (2010). Sustainability: A biological perspective. Cambridge University Press, Cambridge, UK.
Paton, G. (2001). "A Systemic Action Learning Cycle as the Key Element of an Ongoing Spiral of Analyses." Systemic Practice and Action Research **14**(1): 95-112.
Reckers, U. (1997). "Participatory Project Evaluation: Letting local people have their say." Development in Practice **7**(3): 298-300.
Siche, J.R., Agostinho, F., Ortega, E. and Romeiro, A. (2008). Sustainability of nations by indices: Comparative study between environmental sustainability index, ecological footprint and the emergy performance indices. Ecological Economics. 66(4), 628-637
Traber, M. (1986). The Myth of the Information Revolution: social and ethical implications of communication technology. London, Sage Communications in Society Series.
Venetoulis, J. and Talberth, J. (2008). Refining the ecological footprint. Environment Development and Sustainability. 10(4), 441-469
Verbruggen, H. and O. Kuik (1991). Indicators of Sustainable Development: an overview. In Search of Indicators of Sustainable Development. O. Kuik and H. Verbruggen. Dordrecht, Kluwer Academic Publishers**:** 1-6.

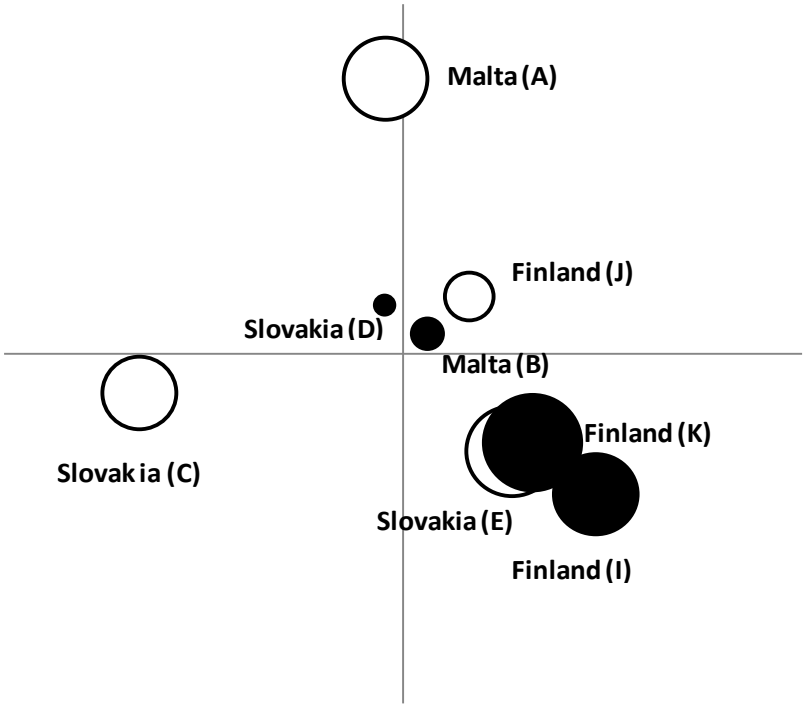
Figure 1. Conceptual Model of the Paper Structure

Figure 2. The Triple Task field diagram and its interpretation.

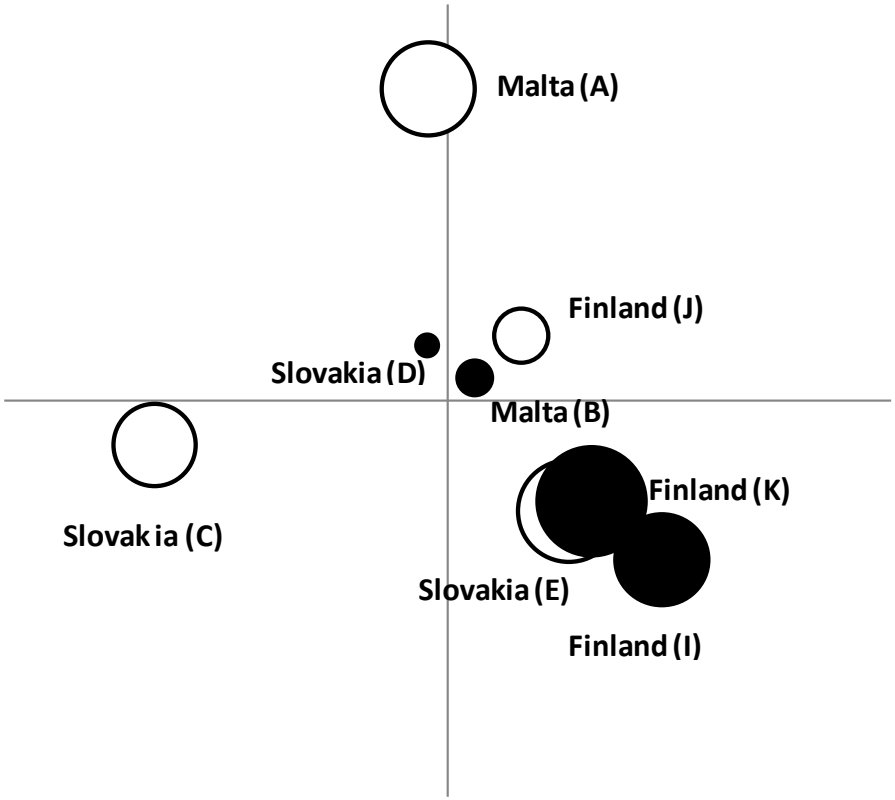
(a) Interpretation of the field diagram axes



(b) The field diagram and location of the groups from Malta, Slovakia and Finland.

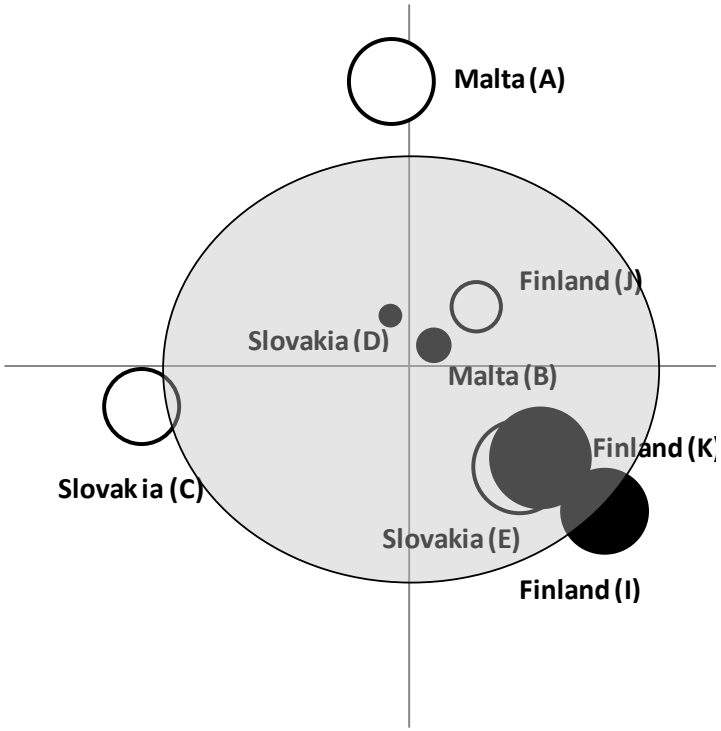


(c) Some textual interpretation of the four quadrants of the TT field diagram



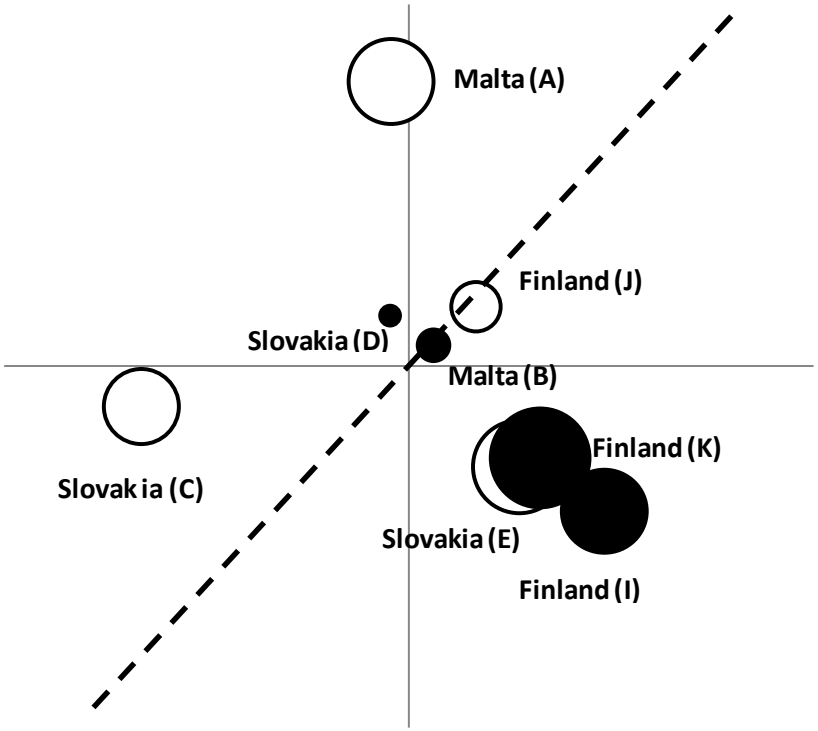
Quadrant 1 Functional Instrumentalists High group function but low quality output Is the group interested in what it is doing? Does it value the process? Possibly switched off from the process? (Metaphor: 'Disinterested')	**Quadrant 2 Functional Enthusiasts** High group function and high quality outputs A well organised and engaged group of people who overcome any initial problems of the group makeup and work well on the task suggested. (Metaphor: 'Well-oiled machine')
Quadrant 3 Dysfunctional Instrumentalists Low group function and low quality outputs The group does not rise over any issues which it has as a divergent set of individuals. They do not engage well in the task and cannot function as the process would expect.	**Quadrant 4 Dysfunctional Enthusiasts** Low group function and yet high quality output. Very interesting group which performs well on the task despite possible conflict and issues over group membership. Here we have high output arising in part as a consequence of the problems which

(d) Pragmatists and mavericks in the TT field diagram

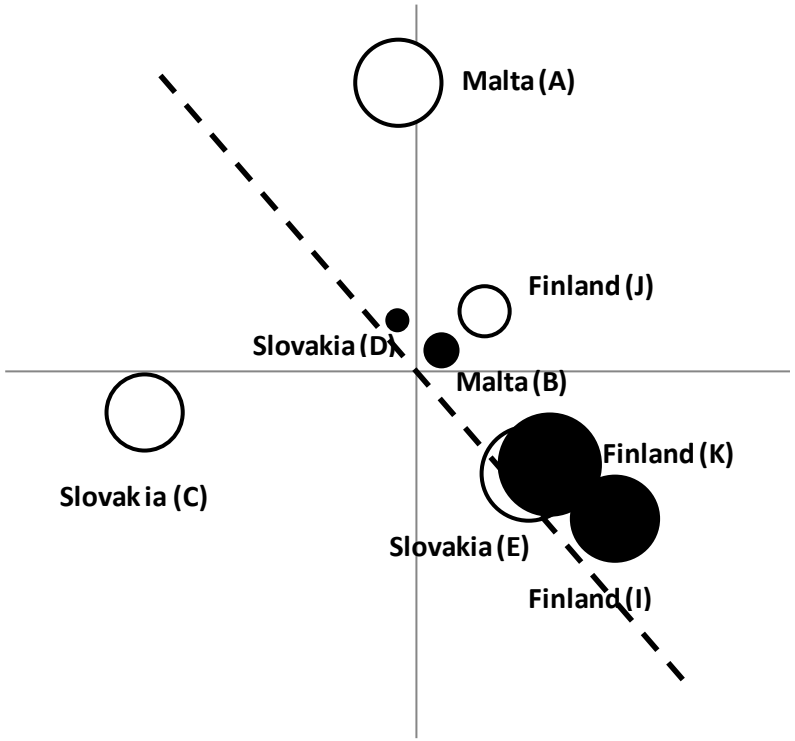


Inside ring: Pragmatists
Outside ring: Mavericks

(e) Hypothesised (logical) relationship between group function and quality of the outputs which groups generate



(f) Another hypothesised relationship between group function and quality of the outputs which groups generate. Here the quality of outputs increases as group function declines.



(g) Third hypothesised relationship between group function and quality of the outputs which groups generate. This is a mix of (e) and (f).