Monitoring and Evaluation NEWS

A news service focusing on developments in monitoring and evaluation methods relevant to development programmes with social development objectives. Managed by Rick Davies, since 1997

Hierarchical Card Sorting (HCS)

A simple tool for qualitative research and inquiry, which can also be useful for planning and evaluation

Contents of this page

BACKGROUND

This is a tool I developed in 1993, while doing PhD field work with the Christian Development Commission of Bangladesh (CCDB), in Dhaka, Bangladesh. I subsequently wrote two how-to-do-it papers describing the method (both of which have been on the web for some time):

- Hierarchical Card Sorting: A Tool for Qualitative Research (1996)
- Tree Maps: A Tool for Structuring, Exploring and Summarising Qualitative Information (1998)

This page now integrates the contents of those two papers, and adds further content developed since then. An initial clarification may help. Hierarchical Card Sorting refers to the *process*. Treemaps refer to the *product* of that process.

WHAT IS HIERARCHICAL CARD SORTING (HCS)?

HCS is one of many types of card sorting methods (also known as pile sorting). Card sorting has been used in many contexts, from traditional ethnography to the modern day business of designing usable websites (See references below). In these contexts card sorting is typically used to elicit people's mental models: the categories they use, what belongs to these categories, and how the categories relate to each other.

WHAT USE IS IT?

In many organisations people accumulate a lot of knowledge, but often it is tacit and informal in nature. As such. it is not so easily shared. Yet sharing that

knowledge can make a difference, other people can make use of it, and they can help correct it and improve it. A HCS can help make people's knowledge more explicit and publicly available, contestable and usable.

HOW DO YOU DO IT?

The HCS method asks people about significant differences. About differences which are important to them and which have (or had) consequences. It has similarities in origin and approach with the Most Significant Change (MSC) technique. Central to the HCS is a question about the "most significant [static] difference", whereas MSC asks about the "most significant change". Both ask respondents to make observations and interpretations. The design of both tools was influenced by Gregory Bateson, especially his book "Mind and Nature: A Necessary Unity" (1979), in which he argues that information is "a difference that makes a difference". In turn, many people would argue that knowledge is structured information. The HCS is about eliciting and representing people's knowledge (i.e as a structured set of differences that make a difference).

Normally the HCS is used with one respondent. However, the process outlined below can easily be used with a small group. The steps:

1. **Identify the respondent's area of expertise** or knowledge that you want to explore. For example, the knowledge of animal diseases held by a paravet, or knowledge of local NGOs held by a INGO staff member working on NGO capacity building.
2. **Generate a list of actual *cases* which will be sorted**. If possible, try to select cases that represent a wide variety of examples. In many cases you will want to select cases that the respondent is responsible for in some way, and thus should be expected to know about them. For example, a list of cases treated by the paravet in the last month, or a list of NGO grantees managed by an INGO desk officer. Don't be too ambitious, especially to start with. Large numbers of cases (25+) will make the process more time consuming and will run the risk of boring the respondent and interviewer. Write the name of each case down on a separate card. Cases might be events (treatment provided) or entities (clients).

3. **Place all the cards in one pile** (see this as the trunk of a tree) and begin by asking the respondent to tell you about some of the differences between all these cases. The purpose of this initial question is simply to generate a shared awareness of the large number of differences that (inevitably) exist. It is a warm up exercise. If the respondent finds this difficult, randomly select two cards at a time, and ask them to compare them, with a view to finding interesting differences.
4. **Ask the respondent to sort all the cards into two piles *of any size*** (see these as the first two branches above the trunk of a tree), representing what they think is the most significant difference between all the cases represented on the cards. Emphasise that it is their opinion of "significant" which is important. If you want to direct their attention in a particular direction then use a prefix to the question, such as "*In your roles as*...what do you think is the most significant difference between...?. Or "*Considering the objectives of this organisation*...what do you think is the most significant difference between...?
5. **Emphasise that a distinction is significant if it makes a difference**. Because respondents may casually offer a difference simply to oblige the interviewer it is important to check its significance by asking "*What difference does this difference make ?"* If one can't be identified then suggest to the respondent that they consider if there are other differences which might be more significant.
6. **Keep a record of which cards are placed in which pile.** This is easier if all the cards have all been numbered beforehand. **And then write down a description of the reported difference between the two piles**. And what difference that difference makes.
7. **Take one of the two piles at at time and repeat stages 4 to 6 above.** Then repeat this process with second pile. There should now be four piles. Repeat the same process with these piles until there is only one card left in each pile. One way of keeping track of this process is to use a large piece of paper, to draw a tree whose branches spilt into smaller and smaller branches.
   - You can develop the tree structure in various ways: (a) breadth-first, or (b) or depth first, or (c) a mix of both. Your choice might best be guided by where the respondent seems to be most at ease.
8. In some cases there may be more than one example left in a pile but the respondent may not be able to identify an important difference between

them. **Don't force them** to do so, but simply note that no further difference could be identified.

9. **Document the results** of the exercise in a form that shows how the different groups and sub-groups or cards are nested, along with their perceived differences and the differences these make. See the next section here for examples.

EXAMPLES OF HCS RESULTS

The HCS process generated tree structures, which can be represented in a table or diagram form. Figure 1 is a tree in table form (using Excel), produced in 1993. Click on the image to read the text, and click again to enlarge. Read the results from left to right. Green shaded sections each refer to one of two binary types that the respondent would like to see funded more in the future. These preferences have then been converted to an overall ranking, as seen on the right. This example is incomplete because while it shows differences, it does not yet show the differences they make i.e. the expected or realised consequences.

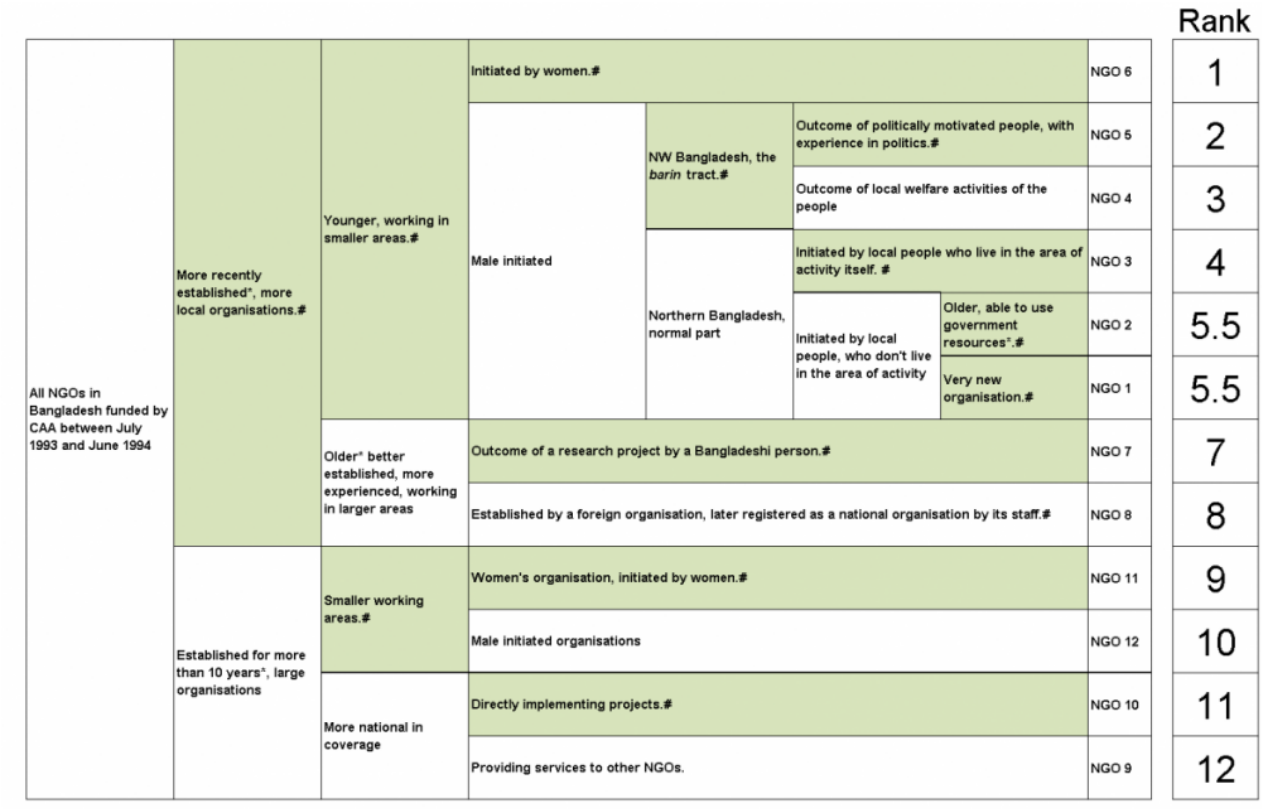

Figure 1: A classification of local organisations funded by donor NGO in Bangladesh in 1993

The same results can be shown as a more explicit tree structure (shown below without the text notes). Red lines = preferred types of partnerships in future

(discussed below). Any diagramming software can be used to do this. My favorite is the free version of yEd.

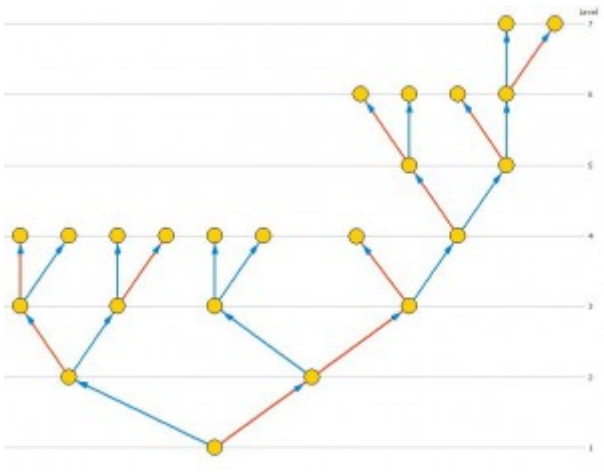

Figure 2: Pile sort results as a network structure

Another example is shown in Figure 3 below. This is the result of an interview with a rural community development worker, where I asked about the most significant differences between the villages they were working with. The wealth ranking on the right is based on binary comparisons made at each junction of the tree, by the respondent. After the tree had been constructed, I had asked them which of the two sub-groups of villages at each junction they thought were more versus less poor.

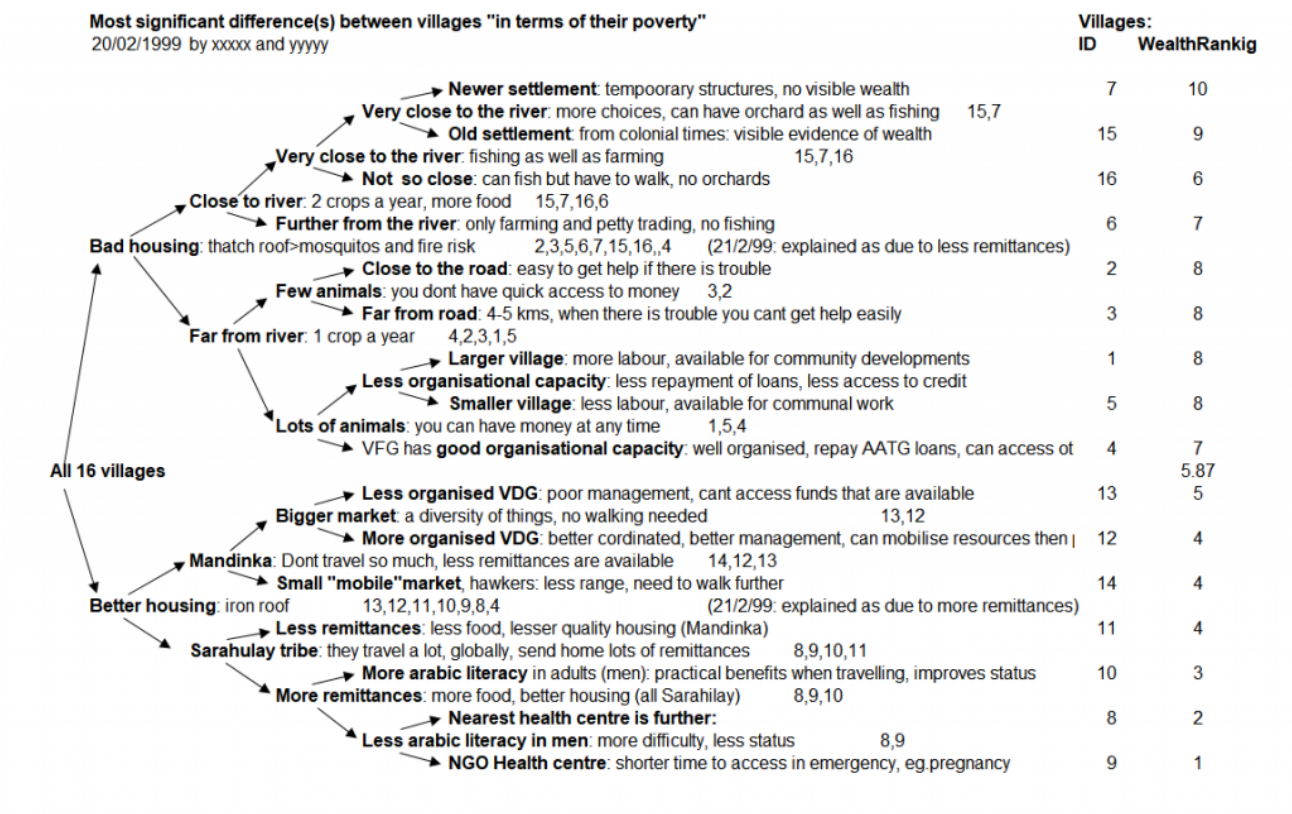

Figure 3: Pile sort results that include "the difference the difference makes" – based on an interview with a rural community development worker in west Africa

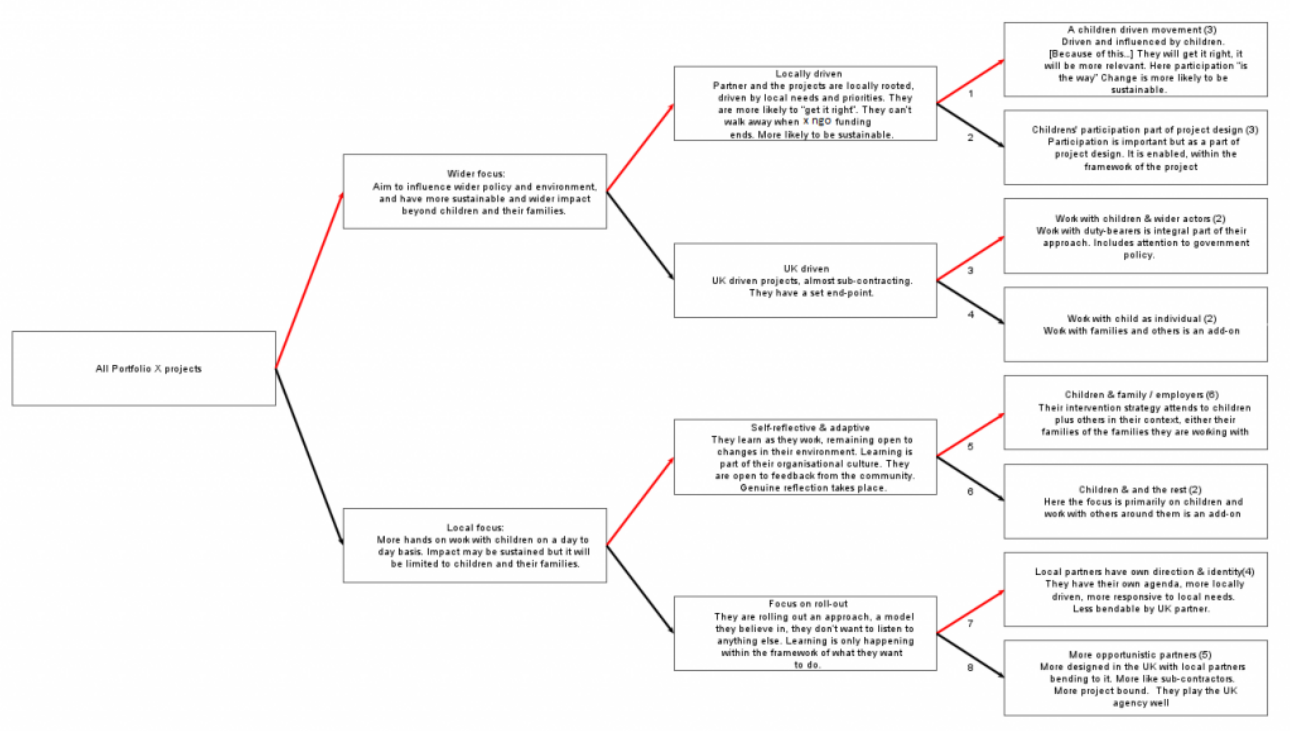

Figure 4: A donor NGO's view of the most significant difference between its grantees, in one thematic portfolio (click to enlarge)

Figure 5 was developed by a groups of 4 people, rather than by one single respondent, This was within a planning session, concerned with the design of a new rural livelihoods programme.

<table>
<tr><td rowspan="6">Direct Intervention (>greater possibility of effect on communities' lives)</td><td>Non-membership (>less…)</td><td colspan="3">WOFAN, CRDP-COCIN, Legal Watch</td></tr>
<tr><td rowspan="3">Membership (>greater accountability to the community, better responses to their priorities</td><td>Group membership (>reaches more, more diversity, more organic relationship with people)</td><td colspan="2">ZAPO, COWAN, TULA Women's Assoc.</td></tr>
<tr><td rowspan="2">Individual membership</td><td>Mixed RNR (>good but need to identify RNR activities possible)</td><td>Widow's Forum, Care for Katsina, Grassroots Development</td></tr>
<tr><td>Direct RNR</td><td>FADAMA Dev. Assoc, PFA, Miyetti Allah</td></tr>
<tr><td rowspan="3">Indirect Interventions (>lesser possibility…)</td><td rowspan="2">1-2 state coverage (>reach less people)</td><td>Less participatory institutional approach (>offers less…)</td><td colspan="2"></td></tr>
<tr><td>More participatory institutional approach (>offers opportunity for full utilisation of skills and potential)</td><td colspan="2">DEC-Bauchi</td></tr>
<tr><td>Wider geographical coverage (>reach a greater number of beneficiaries)</td><td>Non-network (>Less possibilities of information sharing and learning, spread effect of CBDD project)</td><td>Locally constituted mandate (>less constrained in…)</td><td>Non-RNR focus</td><td>Baobab</td></tr>
<tr><td colspan="2"></td><td>RNR focus</td><td>WIN, NEST, Nirado, CDTF, NCF</td></tr>
<tr><td colspan="2"></td><td>Externally constituted mandate (>constrained in ability to learn new issues from below)</td><td colspan="2">Technoserve, WWF, IITA</td></tr>
<tr><td colspan="2"></td><td>Network (>more…)</td><td colspan="2">NIPRANET, NINCOF, CDM Round Table</td></tr>
</table>

Tree map of the types of NGOs and CBOs that XXX may be working with. Produced by Groups A, May, 1998

Normal text = the most significant difference in the left side category of NGOs and CBOs
Bold = distinctions also made in analysis of NGO workshop results.
Italics = distinctions that generated much debate.
>.... = difference the significant distinction makes

Figure 5

TWO USER ACCOUNTS OF THE USE THE HCS

Bangladesh

In 1994 staff of the Research Unit of the Christian Commission for Development in Bangladesh (CCDB) used the cardless version of the same method in the process of exploring poor peoples conceptions of health, disease and medicine. The following quote describes the results of two applications of the method:

"In two somities [beneficiary groups] we also asked the members to come up with different medicaments they use when they are ill. We asked them to divide these medicaments into two groups. We asked for an explanation after each division. After dividing the group into two other groups we asked them to do the same for the two other sub-groups. In the first somity they divided the medicaments into allopathic and

herbal medicines. The allopathic medicines were divided into tablets and syrup. The herbal medicine into medicines used for cough and influenza, and medicines for influenza.

In the second somity they divided the medicaments into medicines given by the kobiraj and medicines given by the allopathic doctor. The medicines given by the kobiraj were divided into medicines for weakness and medicines for influenza and headache. The medicines of the allopathic doctor were divided into medicines for pain and medicines for gastric burns.

We used this exercise to get more information about people's concepts of medicines. When we asked about [the consequences of] the difference between herbal and allopathic medicines it was mentioned that allopathic medicines cure better and sooner but they are costly and can give weakness. Herbs are "softer" to the body, they don't harm the body, they are cheap , and although they work slowly they keep the body healthy." (People's Health and Nutrition, January 1995, CCDB)

Afghanistan

"*I thought I would let you know that I have tried your treemap method for the first time. Despite thinking it sounded rather unlikely to be useful when I first read it, it has provided a very nice alternative to classic wealth ranking for investigating the structure of villages. The people I am working with had the same unease as I do about launching into wealth analysis. We therefore tried your method as it seemed to offer a way round the problem. We ask the number of families in the village, then ask about differences between them with respect to livestock and well-being. Following the process as you describe, we rapidly get a good idea of the wealth structure of villages, which is far more disparate than I had imagined. From this exercise it is possible to find people to interview in more detail from each category. Repeating the exercise with paravets, we then asked them to indicate what proportion of their work was with which group. This worked well and provided an indication of who was benefiting most from the services of the programme.*"

PROBLEMS AND LIMITATIONS

1. **Some caution may be appropriate**. Like all participatory methods it requires some trust and confidence in the relationship between yourself and the person whose views you are seeking. Secondly, there is no guarantee that the views that are expressed will be stable over time. Peoples views of the world change, and the expression of their views is often very context dependent.

2. Some people react at some stage to the exercise by saying "*There is no difference between these*". Here I have cautiously tried to **give many examples of possible differences**, while being careful not to lead in any particular direction. I have emphasised that differences can be found even between objects that look identical, the question is which of these is most important from their point of view. I also emphasise that we are looking for relative rather than absolute significance. But if the person is really struggling, especially after having already identified previous differences between the earlier bigger piles, I do not push them.

3. Another problem is almost the opposite in nature. People can approach the task in what appears to be an un-engaged manner, **blithely tossing off distinctions** which don't seem too significant. When this happens I have tried asking "*In what way is that significant, what difference does that make* ?, checking to see that the respondent can articulate the significance, and if not checking to see if they really understand the exercise.

4. Another problem relates to respondents who are almost too helpful. As can be seen from the tables below it is common for some **respondents to report more than one difference**. When well organised I have dutifully noted these down and then asked the respondent, after reading them back, "*and which of these……are the most significant* ?" Failing to do this has meant I have been the one that ends up speculating on their relative importance to the respondent.

5. **Many respondents find it easier to identify differences, than it is to identify the differences these differences will make**. Often both respondent and interviewer assume that this is self-evident and fail to document this part of the exercise. Yet these beliefs can contain important assumptions or hypotheses about the way things work, which can benefit by being openly described, scrutinised and tested.

MAKING USE OF HCS RESULTS

Value can be obtained from tree diagrams at two stages:(a) during the creation of the tree diagrams, and (b) through comparisons made between parts of the structure once it has been created.

Ethnographic use

During creation of a tree diagrams the main use is as a ethnographic tool: a means of understanding people's view of the world. There are three types of use:

- **Identifying the distinctions that people see as important**. This is evident in the contents of the differences reported. It is also evident in how early in the exercise they are reported, and how often they are reported (on different branches).
- **Identifying the limits to people's knowledge**: When respondents cannot identify differences between two or more entities the limits to their knowledge seem to have been reached. Knowing what people do not know about can be important, especially when they might be expected to, or claim to have, expertise in that area.
- **Identifying the direction of learning:** It is also worth noting where there is more versus less differentiation of knowledge, visible respectively where where branches end up with a single case rather than multiple cases, which have not yet been differentiated.

Planning and evaluation use

Exploring

After creation, the tree structure can be used as the basis for a series of judgements, about the future and the past. At each junction judgements can be made about the available binary choices i.e. between the two alterative branches. The process of soliciting and documenting these judgements can proceed in either of two directions: from trunk to leaves, or from leaves to trunk.

Some examples of questions that can be asked at each junction are:

- "How will your work in the next six months with this group be different, compared to this group?"
- "Which of these two groups do you plan to be spending more time with ?",
- "Which of these two groups do you expect will present the most problems?"
- "Which of these two groups do you want to scale up / expand / fund more of, in the future?"
- "Which of these two groups do you think *will be* most successful in terms of ... over the next x period of time?"
- "Which of these two groups do you think *has been* most successful in terms of ... over the past x period of time?"
- "What is the most significant difference in the kind of evidence available regarding the success of this group, compared to the other?"

The answers, and associated rationales, can be added as a further annotation to the tree diagram, at the relevant junction.

Reiteration and testing

A HCS can be carried out at multiple points in time, prior to, during and after an intervention. In order to: (a) identify current status of the people/organisations being compared, (b) to compare these with past assessments, and (c) to detail plans for the upcoming future. This kind of application would be most useful in situations where there was a diversity of contexts and needs, and the need for variations in the specifications of interventions.

Creating a Theory of Change with testable predictions

Each branch of a tree diagram can potentially be seen as a causal configuration, i.e.a set of conditions associated with an outcome seen in the entities that have been sorted . A complete set of branches can then be seen as a particular kind of Theory of Change, one that is notably different in at least two ways.

- Firstly, because there will be multiple branches – it will be capable to representing equifinality i.e the reality that there are often multiple alternate means of reaching the same outcome. And asymmetrical causal processes i.e. that an outcome can be absent not because of the absence of its usual causes, but become of the presence of other influences.
- Secondly, contra to more conventional representations, each segment in a branch is not a consecutive event, forming a chain of events. Rather, each segment is an additional kind of difference that is expected to make a difference to the outcome observable in the entity at the end of each branch.

As can be seen in Figures 1 and 3 above, it is possible to convert a set of binary predictions about expected relative success into a ranking. This ranking data can then be compared to independent measures of success, also converted to rank value, to test the validity of the elicited predictive model. Figure 6 below shows the results of one such comparison, summarised in the form of a Confusion Matrix. In this analysis each set of ranking values has been dichotomised into higher versus lower rank values (initial at the median value). This table data can then be analysed using different performance measure. One common measure is Accuracy (= True Positives + True Negatives/All cases). In this instance the HCS predictions had an overall Accuracy of 59%

		Observed in project assessments		
		More successful	Less successful	
Predicted in HCS+ model	More successful	TP:12	FP: 4	16
	Less successful	FN: 7	TN: 4	11
		19	8	

Figure 6: Confusion Matrix summary of HCS predictions of project success versus independent measures of project success

Experiments with different cut-off points can often identify a better performing prediction model. In this example, lowering the cut-off point for the predicted success status to less than 8 (the lowest ranking available) increased the performance to 74%.

Aggregating Most Significant Change (MSC) stories

In the section immediately above the focus was on outcomes that are measurable in some way, at least using some for of ranking. But outcomes are often diverse in form and thus not easily measurable. For example in large scale decentralised programmes, and in programmes that emphasises peoples participation in the planning and management of programme resources. These are the circumstances where the **Most Significant Change (MSC)** technique is typically useful.

When it comes to aggregating a set of MSC stories the default method is "summary by selection". For example, by using the hierarchical structures found in many large organisations. Here is an example graphic from tools4dev.org

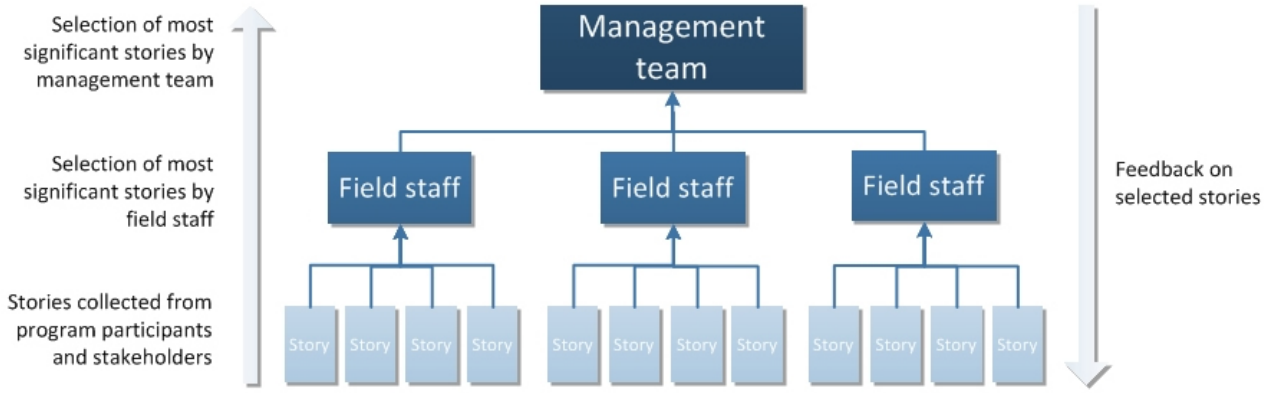

Figure 7: tools4dev.org

A tree structure generated by a HCS exercise provides a similar basis for summary by selection. The tree structure can be seen as something like a tennis tournament structure. At the "branch ends" pairs of MSC stories are like entry level tennis players , who can be compared, and the most significant of these then promoted "up" to the next level. There they meet and are compared with another MSC story that has been similarly prompted. And so on, up the hierarchy.

Unlike the larger scale participatory process described by Figure 7, this process of pair comparison then selective promotion can be undertaken an individual or a small team. The net results is a set of stories with N different levels of significance, where N = the number of levels in the tree structure. For example, in Figure 1 there are seven different levels. As with any MSC exercise, it is important not only to *describe* which MSC is seen as most significant within a pair comparison, but also to *explain* why it is seen as such – what have beenor will be the consequences of that change.

Aggregation of multiple HCS results

If multiple individuals are asked the the same HCS question then their results can be aggregated, and then analysed, with at least two different purposes in mind:

1. Identification of *similarities and differences* between different people's views of the world
2. Identification of a *synthesised* model, one that best predicts an outcome of interest, based on the views of all the participants in the HCS exercises

A Qualitative method

The first method is qualitative, and involves what could be called a "Macro-HCS". Each HCS sort result is in effect treated as a new type of "card" or case. Participants are then asked to identify the differences between these different HCS results. Firstly to generate a list of such differences, in order to reflect on them. Then to select "the most significant", on a reiterated basis, generating a tree diagram, of the kind already seen above. Figure 8 is an example of a Macro-HCS, using each of the previously described HCS results as "cards"

<table>
<tr><td rowspan="3">How donors see NGOs</td><td rowspan="2">View of established relationships</td><td>Theme based partnerships.</td><td>Figure 4</td></tr>
<tr><td>Country based partnerships.</td><td>Figure 1</td></tr>
<tr><td colspan="2">View of potential relationships</td><td>Figure 5</td></tr>
<tr><td colspan="3">How NGOs see communities</td><td>Figure 3</td></tr>
</table>

Figure 8

A variant of this approach could be called a "Meta-HCS". In Figure 8 the "cards" (i.e. HCS exercises) have been differentiated on the basis of the different types of cases they each analyse (i.e. NGOs and villages). But if all the HCS exercises were using the same set of cards (e.g., same set of NGOs), then the Figure 8 Meta-HCS would have to focus on *differences between the kinds of differences* documented in each HCS.

Quantitative methods

Step 1: The raw material for this kind of analyses are data matrices, one per respondent. In each matrix the rows represent the names of the cards sorted, the columns represent the differences identified between the cards, and the cell values of 1 mean the column description applies to the row card, and 0 that it does not. So, the data matrix for Figure 1 would look that shown in Figure 8 below.

	Differences																				
NGO	MRE	EFM	YWI	OBE	SWA	MNI	IBW	MI	OOA	EBA	WOI	MIO	DIP	PST	NWB	NBN	OOP	OOL	IBL	IBL2	OAT	VNO
6	1	0	1	0			1	0			0	0										
5	1	0	1	0			0	1			0	0			1	0	1	0				
4	1	0	1	0			0	1			0	0			1	0	0	1				
3	1	0	1	0			0	1			0	0			0	1			1	0		
2	1	0	1	0			0	1			0	0			0	1			0	1	1	0
1	1	0	1	0			0	1			0	0			0	1			0	1	0	1
7	1	0	0	1					1	0												
8	1	0	0	1					0	1												
11	0	1			1	0					1	0										
12	0	1			1	0					0	1										
10	0	1			0	1							1	0								
9	0	1			0	1							0	1								

Figure 9: Initial data set generated by a HCS exercise

Step 2: Note that there are many gaps in the matrix, where the respondent has not yet told us about the presence or absence of an attribute that describes a difference, in relation to some of the cases. This is because the information that has been provided during the HCS exercise about each difference has always only been about a sub-set of cases, starting from the second-from-left column in Figure 3 and moving rightwards. This missing data, from each participant's matrix, could subsequently gathered using a standardised survey instrument, because the differences with missing data points have already been identified.

Step3: Figure 9 is called an affiliation matrix, Social Network Analysis (SNA) terminology. This now needs to be converted into an adjacency matrix, where both the rows and columns represent the cases (i.e. the entities sorted), and the cell values indicate the number of attributes (differences) shared by the particular combination of row and column cases. This conversion, shown below, has been done using Ucinet.

ID	1	2	3	4	5	6	7	8	9	10	11	12
1	7	6	5	4	4	3	2	2	1	1	1	1
2	6	7	5	4	4	3	2	2	1	1	1	1
3	5	5	6	4	4	3	2	2	1	1	1	1
4	4	4	4	6	5	3	2	2	1	1	1	1
5	4	4	4	5	6	3	2	2	1	1	1	1
6	3	3	3	3	3	4	2	2	1	1	1	1
7	2	2	2	2	2	2	4	3	1	1	1	1
8	2	2	2	2	2	2	3	4	1	1	1	1
9	1	1	1	1	1	1	1	1	4	3	2	2
10	1	1	1	1	1	1	1	1	3	4	2	2
11	1	1	1	1	1	1	1	1	2	2	4	3
12	1	1	1	1	1	1	1	1	2	2	3	4

Figure 10

Step 4: Ucinet can then be used to carry out two types of matrix operations. The first is to aggregate the adjacency matrices of all the respondents.Using the associated network visualisation software NetDraw this data can then be visualises as a network structure. A nested hierarchy of (linked together) attributes can then be identified using a successive filtering of links, from weakest to strongest. This is a form of cluster analysis.

Step 5: Ucinet can also be used to calculate the correlation between the different respondents adjacency matrices. That correlation matrix can then be visualised as a network structure, with high correlation valued links connecting more similar respondents and low correlation valued links indicating more dissimilar respondents

Synthesising using prediction modelling

Another available option is to use supervised machine learning algorithms, to search and find the combination of attributes that have been identified by *all* participants, which is the most accurate predictor of the presence of the outcome of interest, or its absence. This can be done using Rapid Miner Studio, or EvalC3. QCA software can also be for the same purpose.

To do this it is first necessary to include information about the outcome status of each case, as an extra column on the right side of the data set, as shown in Figure 11, already available in Figure 1. According to whatever outcome was of interest. And to ensure inclusion of missing values, as already noted above.

	Differences																					
NGO	MRE	EFM	YWI	OBE	SWA	MNI	IBW	MI	OOA	EBA	WOI	MIO	DIP	PST	NWB	NBN	OOP	OOL	IBL	IBL2	OAT	VNO	Outcome
6	1	0	1	0			1	0			0	0											1
5	1	0	1	0			0	1			0	0			1	0	1	0					2
4	1	0	1	0			0	1			0	0			1	0	0	1					3
3	1	0	1	0			0	1			0	0			0	1			1	0			4
2	1	0	1	0			0	1			0	0			0	1			0	1	1	0	5.5
1	1	0	1	0			0	1			0	0			0	1			0	1	0	1	5.5
7	1	0	0	1					1	0													7
8	1	0	0	1					0	1													8
11	0	1			1	0					1	0											9
12	0	1			1	0					0	1											10
10	0	1			0	1							1	0									11
9	0	1			0	1							0	1									12

Figure 11: Additional development of a HCS data set, prior to a QCA or predictive modelling analysis

Heterarchies – keeping multiple perspectives in view

There are various definitions of what a heterarchy is (Cumming, 2016). This is my interpretation: a set of overlapping hierarchies, as in this example, borrowed from Hearn et al, 2014:

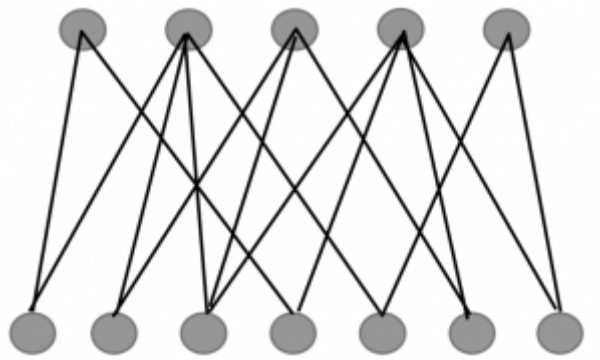

Figure 12

If you ask different people, or groups of people, to each do a card sort of the same set of entities,this will generate a heterarchy. Here is one representing the result of card sorts of the same set of organisations, by two groups of people, in a project planning session in Nigeria. Figure 5 above is one of the two sorts.

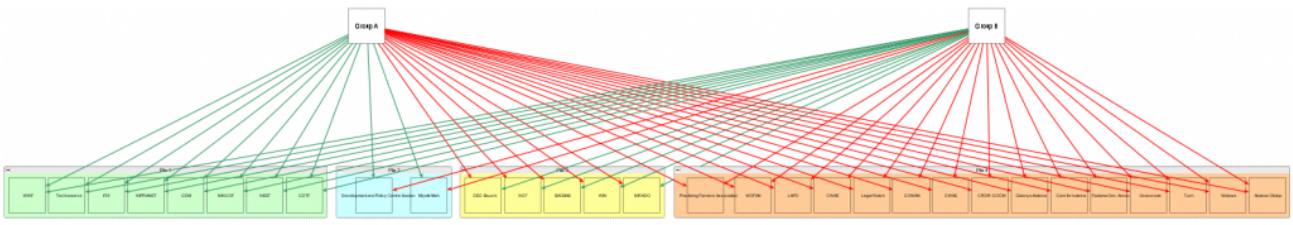

Figure 13: Red and green links represent the initial piles 1 and 2 as identified by each group. The four colored groups of nodes represent the four possible combinations of these sorts

There are 4 different piles of cards, representing the four possible combinations of the first two sets of differences identified by each group in their respective HCS exercises.

But how could this be practically useful? Imagine each group that carried out the HCS was a sub-section within an aid organisation, responsible for different facets or themes within the portfolio of projects that were funded by the organisation. Understanding the relationship between the different groups' views of the most significant differences between the projects could be useful, because different combinations of differences could have different consequences.

There are relatively simple ways of doing such a mapping. Online card sorting services, such as Optimal Sort, can be used to solicit and aggregate different teams' views. Or the same exercise can be done face to face. With both approaches there are risks of a "combinatorial explosion", which may be problematic. With multiple levels of a HCS level the number of possible combinations grows very quickly. Likewise, as the number of participants in parallel sorting exercises increases. Or, if free sorting is used, the number of piles participants choose to use.

An example of the later: In Indonesia four staff members an aid organisation each carried out a card sort of 24 different districts where their programme was operating. In those free sorting exercises exercises staff used between 4 and 8 piles. The net result was 22 different piles with unique combinations of attributes. Working out an appropriate strategy for responding to such a complex mix of locations could be quite challenging.

Another example. 11 participants collaboratively developed 11 different storylines about the future, using a web app called ParEvo. They were then each asked to identify what they thought was the most significant difference between these storylines. The net result was 11 different piles, each with a unique combination of attributes. But some piles were more similar to others, than others. In Figure 14 the most similar piles are in the centre of the network, and the most dissimilar are on the periphery. Different strategies might be developed for working with the core versus periphery piles – a common one for those in the centre and individualised ones for those on the periphery, because they are each quite different from the other.

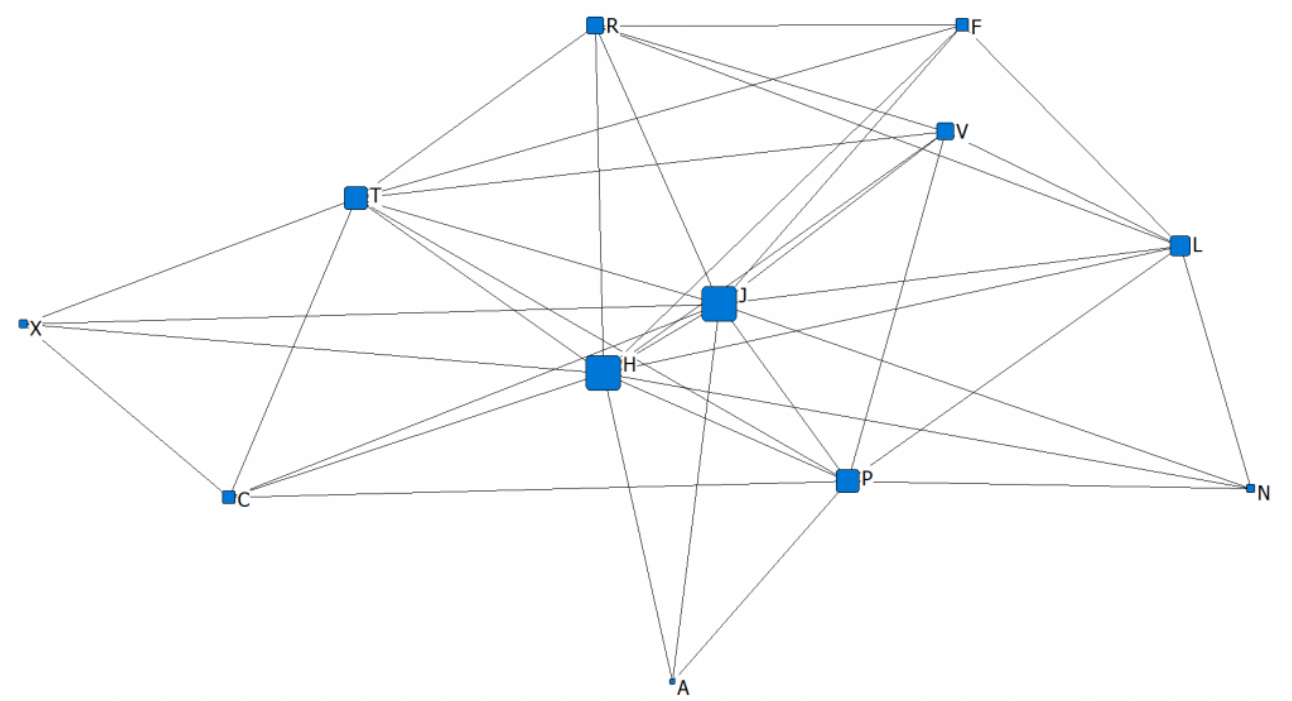

Figure 14: 11 piles of storylines, with links showing above average similarities between storylines. larger node size = higher "closeness" a SNA measure of similarity

Other Specific Applications

1. **Assessing capacity building activities**: When supporting capacity building work with individuals or whole organisations, we might expect that this assistance, either in the short or long term, would make a difference to the person or organisations relationships with their clients. One relevant attribute of that relationship is responsiveness. For example, the service provider might be more sensitive to the differences between client's needs. Or, they may also be more up to date in their knowledge about their various clients' needs. Or, the differences they see between their clients (that they think are significant) may be more reflective of their clients concerns, and not just their own. Much of this information is available, in the first instance, in the form of knowledge the service provider has (or does not have) about its clients. Their knowledge is in effect a proxy indicator or responsiveness. Such information about their knowledge and its likely consequences can be verified by independent observation and follow-up contacts with clients. Are services visibly differentiated (rather than homogenous)? How frequently have they been modified?
2. **Doing a stakeholder analysis**. It should be possible to use tree diagram as a means of doing a stakeholder analysis in a development project. This could initially be from the perspective from one observer, possibly an individual

stakeholder. Firstly, a list of cases reflecting the maximum possible variety of stakeholders would be identified. The process would then start at the trunk, with the respondent being asked to identify "the most significant difference between all the stakeholders in the project". Perhaps after a prefix saying "Bearing in mind the objectives of this project..." , Then each of the two initial categories of stakeholder could be progressively differentiated until all cases were located as a leaf of their own. Information would be generated not only about the different types of stakeholders, but also about the consequences, past, present or future of those differences.

3. **Scenario planning**: This is a practice widely associated with Shell. It involves developing a number of alternate views of the future and then identifying how the organisation concerned would react differently to each different scenario. Tree diagrams of different scenarios could be developed by starting off with a question, such as: "*What is the most important difference in people's views within this organisation/group about what is likely to happen to the oil market over the next 10 years* ?" Within each new branch created by the answer, the question could be repeated, but re-phrased in terms of "*Within this view of the future...*". Once the tree diagram is constructed follow-up planning or evaluation questions could be asked, in the ways described above. PS 2021: A web based participatory scenario planning app, called ParEvo, has recently been developed which makes use of a branching tree structure and whose design was also informed by the same PhD research in Bangladesh.
4. 5. **Analysis of qualitative data**: An NGO may be using the Most Significant Change technique to identify and document stories of change. Normally a participatory process is used "summarise-by-selection". But in some situations a single person may need to make sense of many MSC stories, on their own. In these circumstances a HCS exercise would be one way of doing so, using the different MSC stories as the entities to be sorted.

FUTURE DIRECTIONS

Online accessible software will be developed that enables individuals and groups to carry out multiple HCS exercises and to explore the knowledge structures generated by those exercises. Figure 15 outlines the proposed overall workflow. If

you are interested to be part of the development and early testing of this software, then email: rick.davies@gmail.com

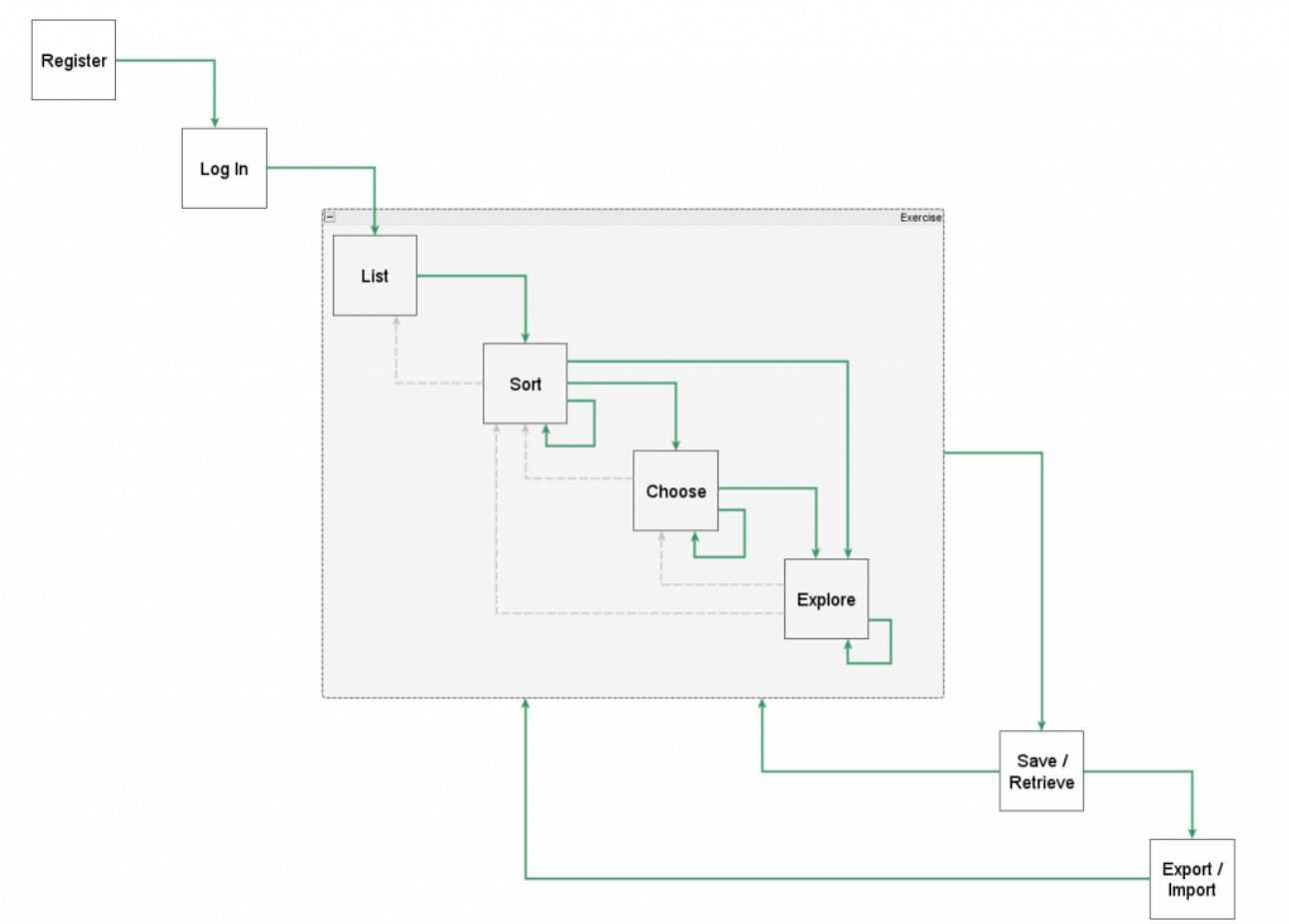

Figure 15: Workflow for HCS Online

PS 2021 08 13: it looks like there may be an easy shortcut to achieve a workable online collaborative platform for HCS exercises. I am currently exploring the features available in Miro: An Online Whiteboard & Visual Collaboration Platform. More details here shortly.

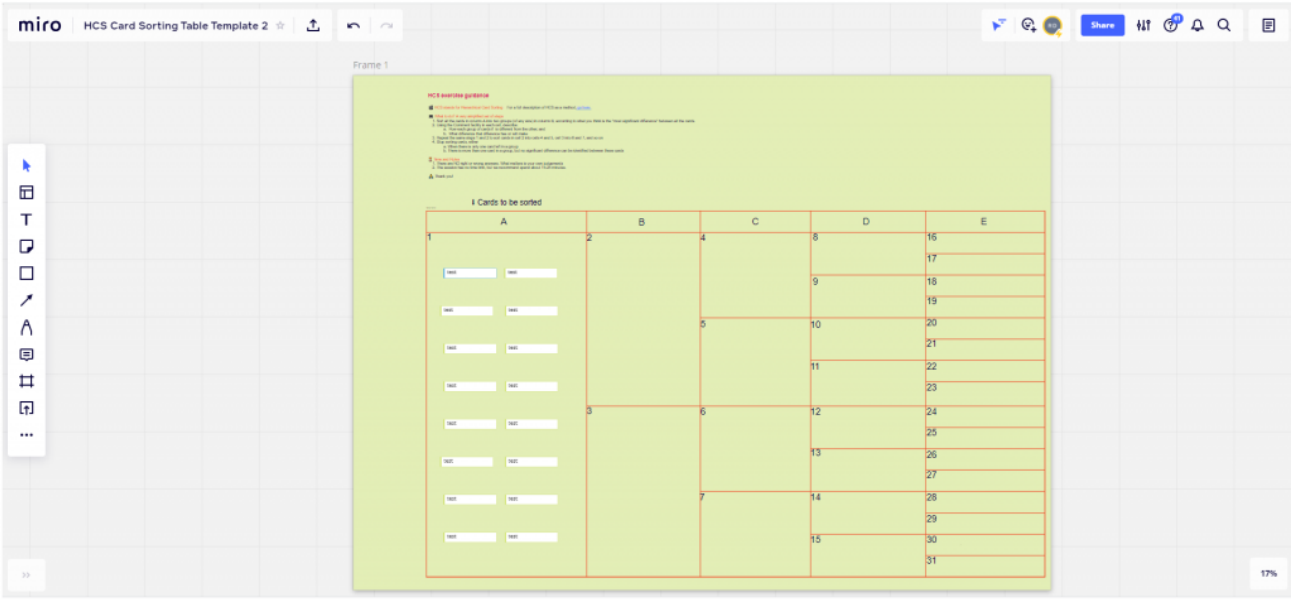

Figure 16: Example of a HCS structure on an online whiteboard. View online here

SIMILAR METHODS

Divisive hierarchical card sorting

Courtesy of Joachim Harloff, I have come across *"Sorting Data: Collection and Analysis"* by A.P.M. Coxon, published by Sage in 1999. In this useful book Coxon (page 26) refers to "divisive hierarchical card sorting" and the fact that it was first documented by James S Boster, a cognitive anthropologist, in his 1986 journal article " ~~Can Individuals Recapitulate the Evolutionary Development of Color Lexicons?~~ *Ethnology* 25(1):61-74. There he was looking for lack of difference in color perception across disparate communities as possible evidence of non-cultural determinants of that perception. The question Boster asked respondents was slightly different to the HCS question described above, it focused on similarities rather than differences: "*I would like you to do is sort these colors into two groups on the basis of which colors you think are most similar to each other*" (page 64). Another more important difference was that there was no follow up question asking "why do you think that difference was most significant". Perhaps for two reasons. Firstly, because there was no need, the respondents were being asked about differences in color, independent of any context that could give one color more significance than another. Secondly, there was no theory of information (such as Bateson's) imbedded in the method. Boster subsequently wrote a paper on "the successive pile sorts" in 1994, as did others in the same period (Wong, 1991), but the method described there appears to be a more

complex process of both agglomeration and differentiation of a variable number of starter piles.

There are many different ways of doing sorting exercises. To find out more see also the more recent "How to Sort" by Harloff and Coxon, 2007, and their Method of Sorting website

Q-sort

in 2021 Bob Williams asked me " ... how does it differ from Q Sort?"

Q-sort, or Q-methodology is described by Wikipedia as "*... a research method used in psychology and in social sciences to study people's "subjectivity"—that is, their viewpoint. Q was developed by psychologist William Stephenson. It has been used both in clinical settings for assessing a patient's progress over time (intra-rater comparison), as well as in research settings to examine how people think about a topic (inter-rater comparisons)*"

The differences as I see them:

1. Q-sort works with statements, whereas HCS can work with just about any entity. I have asked people to do HCS with names of people, organisations, geographic locations, villages, project activities, etc. I even did one years ago in Burkina Faso, where I asked a small group of vegetable growers about the most significant differences between the different parts of a large vegetable garden they collectively managed. One thing I remember was that they distinguished between old and new areas of ground in the garden, because the older areas had been cultivated for longer and were now easier to dig compared to the newer areas. Being a desk-wallah, the relevance of that distinction would never have occurred to me

2. Q-sort asks people to rank items i.e. statements, whereas HCS just ask people to sort things into 2 piles. I think the latter is less cognitively demanding (but can still be very informative).

3. Q-sort asks for rankings that adhere to a particular type of distribution (normal, I think), whereas with HCS, the two piles can be of any size.

4. It appears that in a Q-sort ranking the facilitator determines the broad criteria for ranking. Wikipedia: "a subject might be given statements like "He is a deeply religious man" and "He is a liar," and asked to sort them from "most like how I think about this celebrity" to "least like how I think about this celebrity." whereas with HCS no guidance is given on the type of differences that people should focus on. Though they could be, I suppose.

5. Q-sort involves (but may not require?) factor analysis of the sort results, says Wikipedia. But there is no statistical analysis associated with HCS. (Though you can do simple analysis of the tree structures that are developed and where items are within that structure)

6. The tree structure generated by a HCS can be used by participants for simple planning and evaluation purposes (as described above, on this web page) , whereas I don't think the Q-sort results are so readily usable by the participants

Treejack

Treejack is the name of an online tool developed by Optimal Sort, which allows web designers to test the usability of website navigation structures, which are usually in the form of a tree structure. Respondents are asked to find different kinds of items by navigating their way through a labelled tree structure. The data generated by multiple respondents's search behavior is then made visible, using various data visualisation methods.

Why mention it here? Because it could be a useful way of testing how people make use of a HCS tree structure once developed. The only difference between website and HCS tree structures is that junctions in the former often offer more than two alternative branches.

The Treejack metrics include:

- First click: The first branch they followed

- Last click: Their chosen locations
- Path descriptions: Exactly as followed by each respondent
- Success: Did the respondent correctly identify the location of the mentioned item?
- Direct: Did the respondent need to backtrack at any stage in order to identify the correct location
- Time taken: To find the correct location

EXTRA REFERENCES

Some extra resources on card/pile sorting and related methods:

- Harloff, J., & Coxon, A. P. M. (2007). How To Sort; A short guide on sorting investigations.
- Borgatti, S. P. (Ed.). (1999). Elicitation Techniques for Cultural Domain Analysis. In Designing and Conducting Ethnographic Research (Ethnographer's Toolkit).
- Coxon, A. P. M. (1999). Sorting data: Collection and analysis. SAGE.
- Gladwin, C. H. (1989). *Ethnographic decision tree modeling*. Sage.
  To do any type of sorting exercise a list of cards to be sorted is necessary. Such lists can be identified/generated using unstructured interviews, or by using more structured ethnographic methods such as freelisting. Some suggested reading:
  - Borgatti, S. (1998). Elicitation Techniques for Cultural Domain Analysis. In The Ethnographer's Toolkit (Vol. 3). Altimira Press.
  - Quinlan, M. (2019). *The Freelisting Method* (pp. 1431–1446). https://doi.org/10.1007/978-981-10-2779-6_12-2

Like this:

Loading...

One thought on “Hierarchical Card Sorting (HCS)”

Pingback: Dumbledore’s pensieve – is there any other way to get knowledge out of an expert? | The sum is greater