Quantitative versus qualitative user research.
Part 2: The paradox of affinity diagrams.

Maxwell J. Roberts
July 2024

Downloaded from:
www.tubemapcentral.com/writing/webarticles/Roberts_AffinityDiagrams_2024.pdf

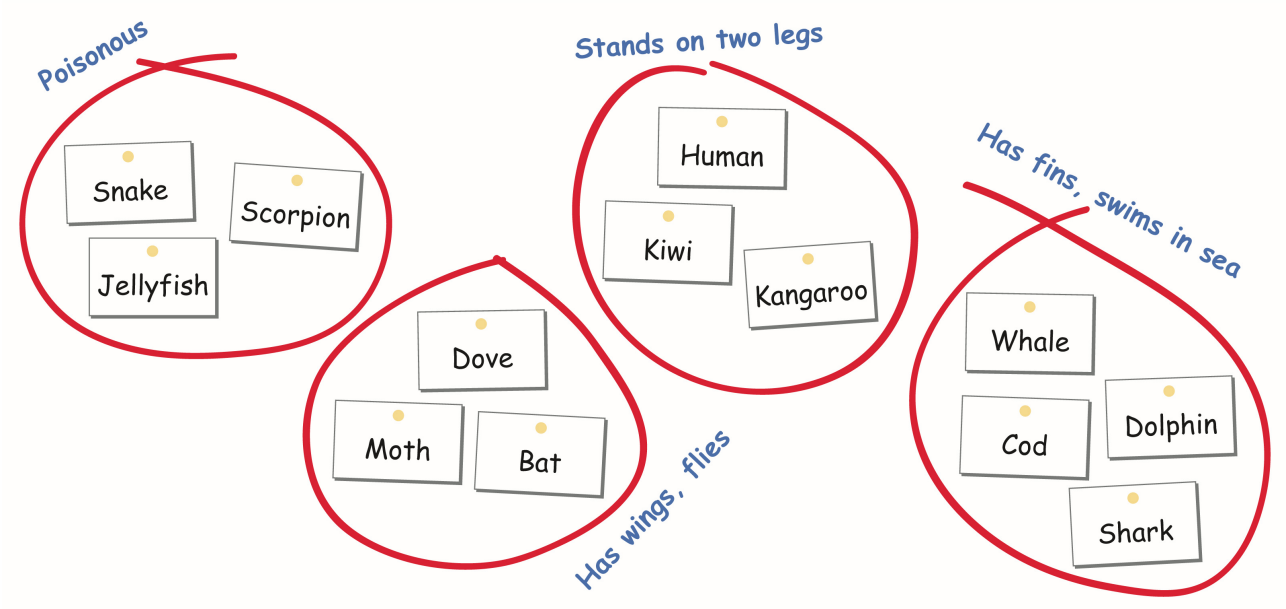
Also at LinkedIn:

https://www.linkedin.com/pulse/quantitative-versus-qualitative-user-research-part-2-paradox-roberts-brvve/

More articles at:

www.tubemapcentral.com/writing/writing.html

Affinity diagram analysis is a popular way of making sense of qualitative data. Unfortunately, the circumstances that necessitate its use also result in this technique offering of an inherently flawed method for making sense of such data. I investigate this problem from the point of view of conceptual frameworks, discuss the use of this methodology from the specific perspective of UX design, and suggest more appropriate uses of this methodology from a UX design perspective.



The outcome of an affinity diagram analysis will rarely be this incoherent. Nonetheless, there is a genuine risk that unsound conclusions will result.

In Part 1 of this series, I discussed qualitative and quantitative research methods, and suggested that the claimed advantages of qualitative methods might be overstated. One frequent assertion is that qualitative methods supply richer data than quantitative techniques. Unfortunately, qualitative data, by its very nature, is vast and unstructured, which causes problems for analysing and making sense of it. In Part 2, I will take a close look at one method for attempting this: *affinity diagram analysis*. I will be thinking about this methodology from the perspective of UX design because (1) this is related to my own research experience, as a psychologist, into transit map usability; (2) I have recently completed a course on UX design, run by industry professionals, in which I was taught that, when assessing usability, qualitative data collection in conjunction with affinity diagram analysis comprised a gold standard in terms of identifying and repairing usability issues; and (3) the difficulties of using affinity diagrams seem to apply particularly strongly to UX design.

Affinity diagram methodology

Affinity diagram methodology was devised in the 1960s by Japanese researcher Jiro Kawakita as a means of making sense of a large amount of qualitative ethnographic data collected in Nepal. The methodology is straightforward and has three steps.

1) *Element identification.* A team of people individually sift through the qualitative data: all of the notes, observations, transcripts, videos and so on. Any aspects of the data that appear to be important are noted onto cards. These comprise the data elements that will be analysed in the next step.
2) *Element grouping.* All the elements (cards) are brought together and the team now work collaboratively to sort them into groups in which the components of each group are related.
3) *Category identification.* Once the groups are finalised, these are then given informative names that represent the underlying similarities between the elements within each one.

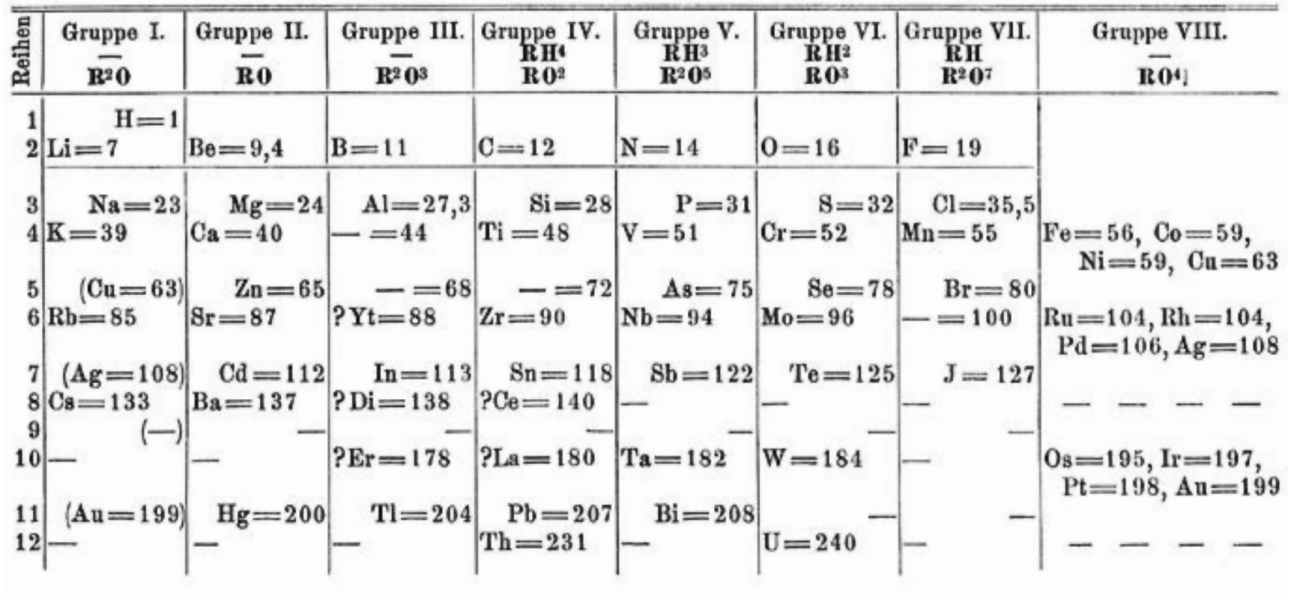
Affinity diagrams, therefore, are a method of organising data and identifying its key categories. From a UX design perspective, an affinity diagram analysis of qualitative usability data would commence with videos and transcripts that resulted from asking test subjects to perform pre-specified tasks while simultaneously vocalising: what they were trying to do; how they were going about doing it; and their expectations and feelings along the way. These would be supplemented with notes and observations from the testers. The team analysing the data would then identify, from all of these records, notable events and issues and write them on cards. Once the cards have been grouped and the categories identified, the completed analysis should enable the identification of the core strengths and weaknesses of a product being tested. For example, by identifying deal-breaking categories, further work can be prioritised to improve the product. Hence, the qualitative data becomes an action plan for the next part of the development cycle.

For maximum effect, it is often recommended that a diverse group of people, not just the UX researchers, takes part in this exercise. Stakeholders should be involved as well, both to ensure that the process is as democratic and unbiased as possible, and also for political reasons: making the people who commissioned the work an integral part of the evaluation process. This should make them more likely to accept the eventual findings.

Affinity diagrams are spectacularly ubiquitous for two reasons, I suspect. First, having made the commitment to collecting qualitative data, this is just about the only analysis that is viable within a constrained budget and timescale. Second, the methodology is guaranteed to yield findings of some sort which can then be turned into a plan of action. The question is: did the affinity diagrams yield sound findings and the best possible plan of action? To answer this, I need to take a step back and talk about conceptual frameworks.

What conceptual frameworks are, and why they are important

The creation of a conceptual framework turns an intellectual pastime from a hobby into a science. It is the identification of (1) the fundamental categories of a domain; (2) how these categories relate to each other; (3) the nature of the elements within the categories; and (4) why the elements belong in their particular categories. Good examples of conceptual frameworks are the periodic table of chemical elements, and the biological taxonomy of organisms. Developing both of these was a painstaking process but, once systems were achieved that obtained scientific acceptance, the respective disciplines were transformed. When a discipline attains a workable, valid conceptual framework this will have a unifying effect on researchers, directing their focus towards fleshing it out, understanding why the fundamental categories exist and how their component elements are similar and different to each other as a function of their categories. Difficult-to-categorise elements will receive particular attention and, as a result of this, the framework itself will be refined or expanded to accommodate them. When Mendeleev announced his periodic table of chemical elements in 1869, he audaciously left spaces in it for as-yet undiscovered elements, and predicted their properties with a high degree of success. Twenty-eight years later the atom began to yield its inner secrets and chemists were finally able to begin to relate position on the periodic table directly to the structure of the atom.



Reihen	Gruppe I. $\frac{}{R^2O}$	Gruppe II. $\frac{}{RO}$	Gruppe III. $\frac{}{R^2O^3}$	Gruppe IV. $\frac{RH^4}{RO^2}$	Gruppe V. $\frac{RH^3}{R^2O^5}$	Gruppe VI. $\frac{RH^2}{RO^3}$	Gruppe VII. $\frac{RH}{R^2O^7}$	Gruppe VIII. $\frac{}{RO^4}$
1	H=1							
2	Li=7	Be=9,4	B=11	C=12	N=14	O=16	F=19	
3	Na=23	Mg=24	Al=27,3	Si=28	P=31	S=32	Cl=35,5	
4	K=39	Ca=40	—=44	Ti=48	V=51	Cr=52	Mn=55	Fe=56, Co=59, Ni=59, Cu=63
5	(Cu=63)	Zn=65	—=68	—=72	As=75	Se=78	Br=80	
6	Rb=85	Sr=87	?Yt=88	Zr=90	Nb=94	Mo=96	—=100	Ru=104, Rh=104, Pd=106, Ag=108
7	(Ag=108)	Cd=112	In=113	Sn=118	Sb=122	Te=125	J=127	
8	Cs=133	Ba=137	?Di=138	?Ce=140	—	—	—	— — — —
9	(—)	—	—	—	—	—	—	
10	—	—	?Er=178	?La=180	Ta=182	W=184	—	Os=195, Ir=197, Pt=198, Au=199
11	(Au=199)	Hg=200	Tl=204	Pb=207	Bi=208	—	—	
12	—	—	—	Th=231	—	U=240	—	— — — —

Mendeleev's 1871 periodic table showing blank spaces for predicted discoveries of elements in the future. The original 1869 version is in portrait format.

The other useful property of conceptual frameworks is that, once established, they become a tool for making sense of new data. Hence, biologists visiting unexplored islands can determine that they have discovered new species of beetles, ants, birds and mushrooms. Occasionally they may discover something harder to categorise, such as a marsupial, but then a debate will ensue and the conceptual framework will be revised and made more sophisticated to accommodate this.

Affinity diagrams are a method of turning data into a conceptual framework

Casting our minds back to 1960s Nepal, the natural inference is that Jiro Kawakita realised that his research and the (qualitative) data that he was collecting were so novel that he simply did not have any straightforward way to make sense of it. In other words, he lacked a conceptual framework. There is absolutely nothing wrong with this. Every discipline has to start somewhere. Hence, he developed affinity diagrams as a method for bootstrapping an initial stab at the key categories of his research topic from his data. One would hope that, as his research developed, he continually refined and developed his conceptual framework, with his later versions being far more sophisticated than his initial ones. Once the fundamental categories were finally identified there should be no further need for any affinity diagram methodology. New data and findings would slot neatly into pre-identified categories, and the framework itself could be modified to accommodate unexpected findings.

Once affinity diagrams are seen as a way of creating conceptual frameworks, there are three important points to note.

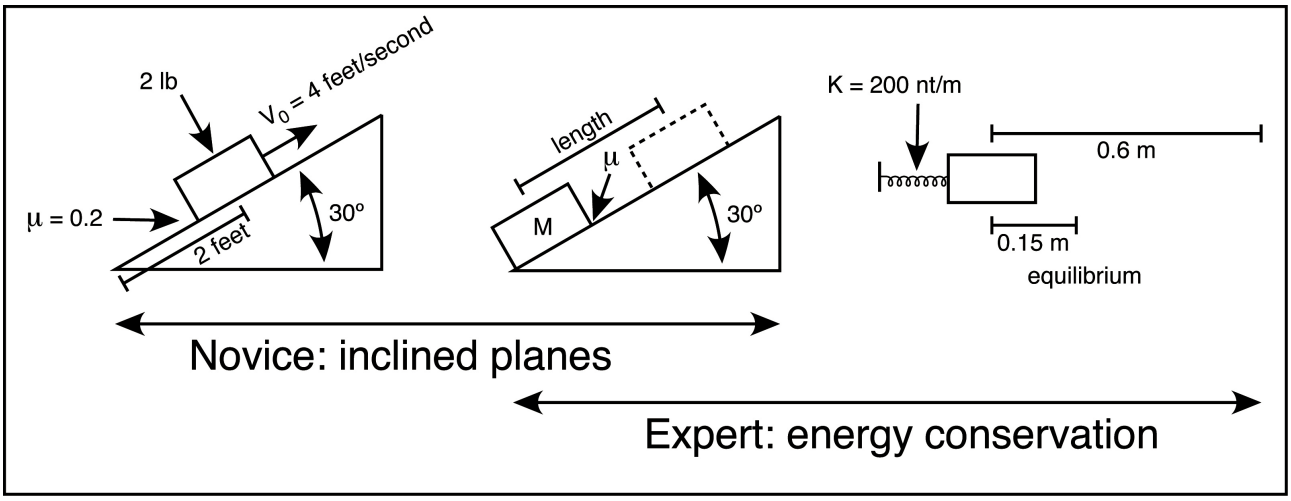
1) You don't need affinity diagrams to formulate a conceptual framework.

This first point should be obvious because the two major conceptual frameworks discussed earlier both predate affinity diagram methodology by more than a century. In general, creating a conceptual framework is a collaborative and iterative exercise amongst researchers, each version is built up incrementally from a combination of top-down and bottom-up thinking, based on hunches, observations, data (both qualitative and quantitative), and a desire to identify orderly structures underlying the disorder of life. My own conceptual framework for transit map design was devised without any affinity diagrams in sight (see http://www.tubemapcentral.com/showcase/d1_transit_map_design.html). It evolved, in part, as an extrapolation of my knowledge of the cognitive psychology of information interpretation and also from my personal creations of maps, varying the design rules and seeing the outcomes. My explorations became more systematic and exhaustive as my ideas developed.

2) Your first attempt at a conceptual framework will almost certainly be wrong; there are some really good empirical and psychological reasons why this will be the case.

Returning to chemistry, Mendeleev's periodic table didn't appear out of the blue, it represented the pinnacle of over 50 years of attempts to organise and categorise the chemical elements. Earlier versions pointed the way forward but all of them were, at best, simplistic. This difficulty in organising the elements occurred for a relatively simple hard science with just ninety-four naturally occurring components to make sense of (although many of these were undiscovered in 1869, making the task of identifying structure from the known elements much harder). Compared with chemistry, the disciplines of ethnography, psychology and human-computer interaction are orders of magnitude more complex. It took great chemists over 50 years to understand their problem and identify a solution. It took even longer for biologists to develop the modern taxonomy of organisms. One would hope, therefore, that a bunch of social scientists working together in a room for an afternoon might show appropriate humility in asserting the soundness of their own findings.

The psychological evidence pinpointing the difficulty of identifying sound conceptual frameworks comes from extensive research into the differences between experts and novices in how they represent and organise information. Numerous domains of expertise have been investigated, all producing similar findings. The classic study is by Chi, Feltovich and Glaser, published in 1981. Expert and novice physicists were asked to sort and categorise textbook problems. Novices organised these according to superficial surface properties, in other words their visual appearance. Experts categorised problems according to their underlying physical properties, which were concealed by the surface features.



Examples of physics problems investigated by Chi, Feltovich and Glaser. Novice physicists lacked a suitable conceptual framework and categorised the two leftmost problems together because these have a similar visual appearance. Expert physicists were able to identify the underlying principles of the two rightmost problems: energy conservation. Even relatively simple concepts can be very different from what might be implied by their surface features. An affinity diagram created by physics novices would be hopelessly inappropriate compared with experts.

The conclusion from expertise research by psychologists (there are many studies with similar findings) is simple and stark. If you are genuinely using affinity diagram analysis because you have no conceptual framework from which to make sense of your data, then your basic categories will almost certainly be shallow and wrong. And, because you have no conceptual framework, you will not be in a position to recognise this, nor understand why your categories are inappropriate. Furthermore, because element identification itself is part of the analysis, this will also impact the results. If, during this stage, an event that has hidden underlying significance is neglected and doesn't end up on a card, then it won't influence the outcome no matter how important it is.

The most cogent response to my argument would be that, in the contexts in which affinity diagrams are most commonly used, the goal is rarely truth or correctness, it is merely shared narrative. The methodology is, therefore, a technique that is ideal for getting people to work together on a problem and come up with a solution. This is fine is as much that "*just do something, anything*" might be a valid strategy in certain contexts. However, looking at this from the perspective of UX design – a discipline that seems to take pride in

its evidence-basis – concepts such 'easy to use' and 'hard to use' should have some sort of objective basis in reality and, hence, so should the methodology used to attain satisfactory usability. Failure to create objectively usable products will have genuine consequences such as product failure and bankruptcy.

3) If you are an experienced professional then you should have a sophisticated conceptual framework already. If you have a conceptual framework then you shouldn't need to be starting from first principles, you should be using/embellishing the framework that you already have. So, do you genuinely need affinity diagrams at all?

Imagine that you have some health concerns and you visit the family doctor (the GP in the UK). You describe your symptoms in the hope that the GP will be able to fit these (qualitative) data into the relevant conceptual framework and, from this, make a correct diagnosis and suggest treatment options. If the symptoms are unusual and the GP is not quite sure, then you would expect to be referred to a specialist, who will have a highly developed conceptual framework relevant to your issues. You ***do not*** expect the GPs in the practice to take the transcript of your conversation, add it to a stack of transcripts that comprise all the difficult cases that month, and perform an affinity diagram analysis on them. If other stakeholders, such as the practice manager and the receptionists are asked to join in with this analysis then, understandably, you might start to get very nervous indeed. Medical conceptual frameworks are not perfect, but they do the job reasonably well, and are being revised and improved as further research results are disseminated.

Next, let us revisit our biologists exploring strange, new environments. If these are unexplored islands on planet Earth then their Earth-derived taxonomic system should be adequate for categorising their new discoveries, apart from a few oddballs, which will advance their knowledge and make their framework more sophisticated. But, what if the unexplored environment is a new planet, far from Earth? To a certain extent life is an engineering solution, so the same problems as those on Earth might produce the same solutions elsewhere but, crucially, there are also chance elements to evolution. There is every reason to expect that the categories of organisms on another planet will be very different. In this situation, an affinity diagram analysis might be an excellent way of trying to make sense of the new lifeforms that have been discovered. Even so, this would be best left to the biologists. The ship's captain, security officer and chef would have no useful contribution to make to this particular task. Only the biologists would have a deep conceptual understanding of how organisms are constructed in general, which will focus their attention towards the potential categorisation features of the newly discovered ones. This isn't an assault on creativity, it's basic common sense.

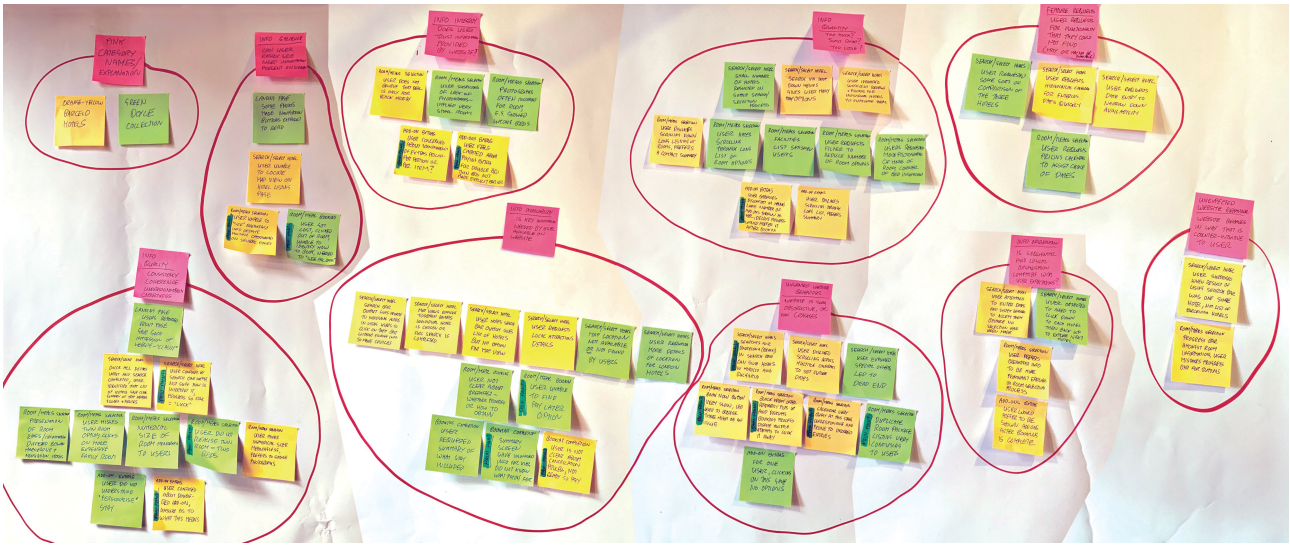
Returning to earth, suppose we have a UX design team who have a lot of expertise at online banking apps, but now they have been asked to help a company set up a hotel booking website. The UX designers are in unexplored territory. The question is, are they on another island, or another planet?

If you are experienced in a domain, such as interface design, and you have read around, then you should already have some sort of understanding of the underlying concepts of your discipline. You will have seen many of the problems before, know why they happen and their solutions. You should be able to build an interface that is basically competent, recognise the defects in interfaces designed by non-experts, and know how to fix them. Occasionally, you might be asked to design an interface for something that has not been attempted before but, even so, you won't be approaching this completely cold: basic principles will still apply. In other words, there should be some sort of conceptual framework available already. In which case, why attempt to reinvent the wheel? If you are testing a new interface, just get your team to read the usability test transcripts and watch the video tapes, fit the findings into the collective conceptual framework and identify from this what needs to be done next.

Why, then, might expert UX designers still advocate affinity diagram methodology? Possibly because of the social psychology of using these, such as the need to involve stakeholders so as to give an impression of democracy and ensure buy-in. While politically expedient, this sort of activity runs the risk of diluting the expertise of the people in the room and inducing errors into the newly-created framework: people who have never attempted to design an interface will categorise data by superficial surface features rather than underlying properties. I see this for myself every time I see "Pain Point" as a big bold category on a completed affinity diagram used to analyse qualitative usability test data. [*In this context, a pain point would be an interface defect that causes a catastrophic failure in goal attainment and, of course, considerable frustration as a result.*] Yes, pain points need to be fixed but why exactly did the interface fail?

Was the user given too little information? Too much information? Ambiguous information? Information that conflicted with the user Mental Model? Was the user expected to rely on memory from a previous screen?

Fundamentally, the issue for an affinity diagram outcome that includes pain point as a category is: should the navigational failures have been categorised by how disastrous they are? In other words, categorised by the ***effect*** of bad interface design. One consideration is that, unless an interface has been constructed incompetently, there is always a degree of bad luck for a sequence of steps to result in a pain point. One person's pain might be another person's fleeting bemusement and minor irritation. The information design failure is the same but, for each person, the outcomes are different. Hence, categorisation by effects would seem to be both atheoretical and rather hit and miss. Not a surprise: *effect* is a *surface feature*, a symptom of bad design. I would argue that it would be far better for ***underlying causes*** to form the basis of the categorisation: why did the interface fail? This has the added advantage that the sorts of remedies necessary are built into the findings.



As part of my UX design course, I was asked to analyse qualitative data obtained from usability tests. However, with my experience as a cognitive psychologist researching into how people interpret and make inferences from information, and why they make errors, I already had a useful conceptual framework in place. I therefore 'cheated' with this exercise, generating a set of cause-of-error categories from a first pass through the data, then organising the highlights of the usability issues into these categories. In other words, my pre-existing conceptual framework enabled me to take a top-down approach to analysing the data. Pain points are highlighted within categories. Only a strong devotee of affinity diagram methodology could assert that my analysis was inappropriate and had not generated useful conclusions. My conceptual framework can also be seen in action where I discuss information design failures on the London Underground map.

I also wonder whether there are additional social psychology factors might deform affinity diagrams, and that can be exacerbated if too many novices are taking part. The people who shout the loudest, metaphorically, will have the most influence on the outcome, for example, during the element identification stage, perhaps creating such a large number of cards on certain aspects of the data that these take over the categorisation, or by their insistence on organising the cards into their own personal categories. According to Wikipedia (8th July 2024) "In many cases, the best results tend to be achieved when the activity is completed by a cross-functional team, including key stakeholders." I would love to know what evidence the writer had to make this claim and, given the qualifier "in many cases" whether the sub-optimal results obtained in other, alluded to situations, cancelled-out the supposed advantages of inviting everyone into the room.

Summary and synthesis

What is the paradox of affinity diagrams? If you are studying a genuinely new topic, and you don't have a useful conceptual framework to apply, then using affinity diagrams might appear to be the best solution but,

if you don't have a useful conceptual framework, then this is exactly the situation where the outcome will be superficial categories and affinity diagrams will give poor results. Put simply, people who need affinity diagrams the most are least likely to get sound, or even useful results from them. Inevitably, this issue feeds back to qualitative methodologies in general. Not only do these yield data that are difficult to analyse (see Part 1) but, also, qualitative data might actually be booby-trapped: the available ways to analyse them are likely to result in unsound answers for the unwary.

What is the way forward? I am not advocating abandonment of qualitative methods, or even affinity diagrams. Instead, what is needed is a thorough understanding of the various methods available, their strengths and weaknesses, and how to get the best out of them. There is no space in this balanced approach for empty-headed quasi-religious insistence that certain ways are the best.

One-off affinity diagrams are of questionable value in determining the objective way of the world. On first pass, the tendency of people to interpret unfamiliar concepts by their surface structure, in conjunction with the fallible human lens of expectations, beliefs and prejudices, risks distorting any objective truth out there beyond recognition. On the other hand, an affinity diagram is, by implication, a potentially excellent way of discovering people's expectations, beliefs and prejudices themselves. Hence, this methodology can be used as a tool to understand how stakeholders see the problem and the design brief within their own organisational context: what they think of their current products, what they are hoping for from their new products, how their new products should look, and perhaps any political undertones and warring factions that will cause difficulties along the way. If I was entering an organisation in a consultancy role, the first thing I might try would be to conduct some brainstorming/affinity diagram exercises just to find out what I might be up against in terms of organisational pathology.

UX designers, in their role as competent experienced professionals, should not be using affinity diagram methodology for routine work. Knowledge and experience, in the form of appropriate sophisticated conceptual frameworks, should be in place already, enabling them to critique existing products, identifying their strengths and weaknesses. Designers should also be able to create basically competent products themselves, for example web pages that are straightforward to use, that present tasks as logically coherent sequential steps, that present necessary information clearly and exactly when it is needed, and that don't trap users who make mistakes in deep rabbit holes. This is similar to what I would expect from a competent architect designing a house and a builder implementing the plans.

Assuming this is the case, then a few user tests should be all that is needed to confirm basic usability and that no issues have been unaddressed. Any specific design questions for which there is not a clear consensus can be resolved with A/B testing, using quantitative or qualitative methods depending on the particular objectives for that instance. However, in some situations where UX designers are venturing into new territory, the team might feel that their collective conceptual framework for effective design might be rather stretched, so that they are unsure as to how to design a usable product to perform the novel task. Here, extensive user testing, which incorporates qualitative methodology, is justified. Complex qualitative data will need some form of appropriate analysis, such as affinity diagrams but, in this situation, it should be only the designers who will be take part, not the stakeholders. The purpose of this exercise is to understand the problem and see how it fits in with the existing conceptual framework of the UX designers, with a view to enhancing and expanding this as necessary. Non-experts at UX design will not be able to function at the necessary conceptual level to make a useful contribution. However, it would be a fascinating exercise to take the same usability data and conduct two affinity diagram analyses, one conducted by the experts UX designers and one by the non-expert stakeholders. The different sophistication of outcomes should be revealing.

In the spirit of balance, in the third and final part of this series I will round off by taking a look at some issues when applying quantitative methodology to usability testing, most notably, what I call the Usability Gap.