

ON THE LURE OF METRICATION: ATTITUDES AND SOCIAL REPRESENTATIONS

GEORGE GASKELL

London School of Economics, UK

It is in the triangle of mediation of theory, hypothetical constructs and empirical observations that much social psychological research progresses. Without agreed indicators a construct is restricted to the realm of speculation. Measurable indicators of concepts have a particular attraction as they afford the opportunity to make quantitative comparisons. There are risks, however, in that the objectification of a construct in specified operations may come to be the construct itself. The developments in metric scaling of the attitude construct were little short of revolutionary, introducing a paradigm that has guided the discipline for some 70 years.

As the theoretical scope of social representations has been elaborated the issue of measurement has come to the fore. What is needed, some argue, is an agreed methodology with which to conduct empirical research. Doise and colleagues (1993) have promoted some of the recent developments in multi-variate analysis, others, including Roiser (this volume) look back to the original work on the measurement of attitudes as a possible approach.

The substance of the correspondence between Moscovici and Guttman, quoted by Martin Roiser in his thought provoking article, is also seen in Jaspars and Fraser's chapter on attitudes and social representations (Farr and Moscovici, 1984). Here it is argued that "a cumulative attitude scaling procedure implicitly presupposes a common or shared cognitive representation in order to differentiate between the evaluative response patterns of the subjects. The same can be said of other attitude scaling procedures" Without this assumption it would not be possible to "develop an attitude scale of the Guttman type"; p 111. Thus Fraser and Jaspars argue that "it would be more useful to consider attitudes as individual response dispositions based on collective representations"; p. 123.

Martin Roiser takes this argument further to suggest that the Guttman scale is a possible methodology for elucidating within group consensus and between group disagreements regarding a social object. He cites the example of nuclear armaments in Britain and, with an 'inter-ocular' Guttmanesque procedure, shows a convincing

similarity between the positions taken by significant political groupings and measures of public opinion. Martin Roiser suggests that more research is required. As a contribution to this effort let us look at some examples of the Guttman and the Thurstone scales, which help to elucidate their strengths and weaknesses for research on social representations. At issue is whether these scales, as originally conceived, should be adopted as empirical tools for the study of social representations or whether they are merely suggestive of possible approaches to the study of social representations.

An interesting and recent application of the Guttman scalogram is found in 'Political action in Europe and the USA' by Alan Marsh (1990). This multi-country study looked at patterns of conventional and unconventional political participation in a context of the shift from the conformity and passivity of the 50s and early 60s to the strikes, occupations, boycotts and collective violence of the late 60s to the present day. While the findings make fascinating reading, for example the repertoire of political action has expanded since the 70s with young people in the vanguard, it is with the lessons of the methodology for social representations research that we are currently concerned.

Following extensive pilot research Guttman scales were constructed comprising seven items covering the spectrum of conventional participation ranging from reading about politics in the papers (low), to attending political meetings (high). Unconventional participation was similarly assessed with seven items including petitions, demonstrations, blockades and unofficial strikes. The pilot research showed that people had elaborated views about the justification and effectiveness of different forms of protest and that these ideas were general principles rather than context specific. While there were some differences between the countries in the dominance ordering of the forms of participation, for each country scales with high coefficients of reproducibility (around 0.95) were produced.

With Guttman scales it is in principle possible to make comparisons over time, to explore the characteristics of groups of people having the same scale values and differences between groups having different scale values. That is to say shifts in opinion, consensus and disagreements. In this context one might go beyond the mere scale value to develop two perspectives. Firstly the segmentation issue; who in terms of sociodemographics characteristics agrees and disagrees with whom? And secondly, towards an understanding of the scale value, what does any one group agree about and why, and how do they differ from other groups? The answers to such questions about the meaning of the scale values would require, additional information from interviews, media coverage and exposure etc. In combination these approaches would give a typology of representations of interesting groupings.

In Marsh's study some of these paths are investigated. Cross national comparisons showed that the Americans are far more likely to actively participate in politics than Europeans, but less likely to vote in elections; the Dutch are the most protest prone country and the Austrians the least. Within country analyses focus on gender and age differences and a five fold typology of political action. But with only 10 and 60 people in the various countries reaching the highest scale category, notwithstanding a sample of 1000, the scale had to be collapsed into three categories. Seven items is hardly a long scale, but without an even distribution of respondents across items, or a very large sample, it may not be possible to say much about potentially interesting minority positions.

In some respects the Guttman scales developed by Marsh and his colleagues are similar to the social distance measure of Bogardus which gives an index of social exclusion and ethnicity. In both cases there is an almost explicit ordering of the scale items. Someone who says that they would exclude a person from ethnic group X from their country is hardly likely to consider the same person as a suitable candidate for a family marriage. Equally it would be surprising to find someone who avoids politics in the newspapers attending meetings and campaigning. Roiser calls these pseudo scalograms as the “pattern of responses seems to be largely predetermined, rather than being socially constructed”. As such, along with Guttman, Roiser rule this type of scale out of court.

There are both methodological and conceptual considerations which lead me to worry about this distinction between pseudo and genuine scales. First, there is the problem of *ex post* interpretation. The test for the presence of a scale is the coefficient of reproducibility (CR). This is 1 minus the number of responses that must be changed in order to achieve a perfect scale. For Guttman 0.85 was the recommended criterion for an acceptable approximation to a perfect scale. Yet apparently it is possible to achieve a CR of 0.89 with randomly generated data for 100 responses to 6 stimuli. As the sample size gets smaller, so is it more likely that the CR criterion will be statistically achieved but not necessarily with any social significance. So if the ordering of the stimuli is determined by the analysis, rather than *a priori*, capitalising on chance is a risk.

Second, given that in the development of an attitude scale a researcher is likely to have some guiding concepts in mind, and/or to conduct pilot interviews, and/or to collect materials from other sources such as media reportage, it would be surprising if they had no idea about the underlying ordering of the scale items. Even a person from Mars who read the papers would soon realise that in Britain people seem to be more exercised by murderers than by those pinching a packet of sweets from a shop. Having an open and enquiring mind in research is one thing, making the *tabula rasa* a necessary precondition to accept the subsequent findings seems rather curious.

Thirdly so-called pseudo scales, as has been illustrated with the Marsh study, provide a basis for a variety of different analytic approaches, compatible with the issues of concern to social representational theorists. The theory is concerned to explore representations of different groupings in a society. In Guttman scale terms this would be evidenced by comparisons among those taking different scale positions on a common metric and among those whose ordering of the metric is different. In this latter context a CR of 0.85, indicating a common ordering of the items need not be the ultimate criterion. One may observe in the patterns of deviations from the perfect scale so-called non-scale types. These might well be of interest in a study of minority positions, or in a columnised society in which a social object is represented differently by different groups. Guttman acknowledges that for two populations a universe may be scalable but with different orderings of objects and categories but does not develop the implications of such a finding in relation to group differences. Quoting a study in which a scale worked for some Army personnel but not for the Air Force, he concludes more generally that if a domain is scalable for one population but not for another population, then the populations differ in kind rather than degree, ie they differ in more than one dimension and it is therefore not possible to compare the two.

The Thurstone procedure for scaling attitudes was developed out of the principles of psychophysics. While the individualisation of the attitude has been documented elsewhere

it is notable that Thurstone recognised the potential of a metric scaling of attitudes for descriptions of and comparisons between social categories. He outlined four uses of mean values on a scale:

1. The average or mean attitude of a particular individual on the issue at stake,
2. The range of opinion that he is willing to accept or tolerate,
3. The relative popularity of each attitude of the scale for a designated group as shown by the frequency distribution for that group, and
4. The degree of homogeneity or heterogeneity on the issues as shown by the spread or dispersion of its frequency distribution.

The first two uses focus on the individual, the second two on the social representations. Commenting on a comparative judgmental task Thurstone says that the results describe as much a group of respondents as that group's view of the stimulus (in this case different nationalities) and thus could be used to look at cultural similarities and dissimilarities. Scale values, he argued, could also be used to compare several different groups in their attitudes on a disputed issue. Such comments seem compatible with current trends in European social psychological thinking, but will we now flock to the Thurstone methodology. There is still the problem of establishing scale values for items. The achilles heel is the procedure of sorting the items into eleven categories, the metric scale representing the attitude variable. This approach was adopted presumably because the number of pair-wise comparative judgments were found to be far too onerous; for example 30 statements would involve 435 separate judgments. (Subsequent work in multi-dimensional scaling has shown that the complete matrix of judgments provides an vastly over-determined solution, thus the required number of pair-wise comparisons can be considerably reduced.)

However, this was not known at the time and as he developed the more economical method using judges to determine the scale values, he stipulated that these scale values should not be affected by the opinions of the judges who sort the statements. He quotes a study by Hinckley, later published in 1932, in which three groups, pro-black white students, anti-black white students and educated blacks acted as judges. Since the three resulting scales were "practically identical" in terms of high inter-correlations he concluded that there was no effect of the judges attitudes on the scale metric. Note that the use of correlation as an indicator of similarity is inappropriate in this context as it takes no account of absolute scores. However a replication by Hovland and Sherif showed Hinckley's result to be an artifact of Thurstone's correction for 'carelessness' which led to the exclusion of a large number of pro-black judges. So a scale is possible as long as some of the judges who take a minority position are excluded; hardly what a social representations approach would find attractive. At the end of the day Thurstone's bold claim that 'attitudes can be measured' may be tenable in a relative sense of measurement but certainly not in any absolute sense.

For both Guttman and Thurstone, the pioneers of attitude scaling, the sine qua non was the uni-dimensionality of a scale. Without establishing this criterion there could be no comparisons between individuals, groups or countries. A single dimension may be an ecologically valid description of some social objects, for example a right to left scale is probably appropriate for a two party political system. Other social objects may be represented by people in more complex terms and, as such a such, a single dimension would not capture the universe representations of the object. Subsequent developments in

varieties of multi-dimensional scaling have shown the heuristic value of the assumption of multi-dimensionality in the mapping of representations of social phenomena. There are many techniques available, factor analysis and variants of MDS and correspondence analysis. These statistical procedures have been misused (for example casually labelling principal components as social representations) as often as they are employed to good purpose. But at least, in principle, these approaches allow for the constructivist presumption that the stimulus is not a given; a social object may be differently represented by different people and groups. In the early days this may have been recognised, but it was not a matter for further research.

All in all these considerations lead me to the conclusion that we are unlikely to profit from a return to these attitude scaling procedures as originally specified. Can anything be salvaged? It is worth acknowledging similarities in the problematics of early attitude scaling and aspects of social representations theory. The requirement to sample the universe of a social object, the use in some studies of claims drawn from ordinary people and the mass media, the opportunity to compare and contrast groupings with different scale positions, the interest in changes in scale positions over time or following the recent media coverage of an issue. All these aspects of attitude scaling resonate with issues in social representations. But there are also differences. Thurstone was at pains to develop a method to scale the affect for and against a psychological object. Like the measurement of a table, he argued, it is not possible to capture all the complexity of the attitude in a single number. While he recognised that it would be interesting to elucidate the cognitive basis underlying affective responses, the task of attitude measurement stopped short of this. With Guttman it seems as if establishing a scale to his criteria and charting the percentages of the population at different scale points was an end in itself.

To this extent attitude scaling methods can only be a small part of the social representational project. We need to go beyond opinion research and analyses concerning 'how many' or 'how much', and ask questions about the 'whys'. What are the meanings and constructions of social objects for different grouping?, where do these opinions come from and how are they transformed and sustained in social life? It is highly unlikely that these issues will be captured in metric quantification, even of a multi-dimensional nature. Yet a combination of the qualitative and quantitative approaches may prove to be productive synthesis - time will tell.

REFERENCES

- Doise, W., Clemence, A. and Lorenzi-Cioldi, F. (1993) *The Quantitative analysis of Social Representations*. Hemel Hempstead: Harvester Wheatsheaf.
- Guttman, L. (1944) A basis for scaling qualitative data. *American Sociological Review*, 9, 139-150.
- Hinckley, E.D. (1932). The influence of individual opinion on construction of an attitude scale. *Journal of Social Psychology*, 3, 283-296.
- Jaspars, J. And Fraser, C. (1984) Attitudes and social representations. In R.M. Farr and S. Moscovici (eds.) *Social Representations: Cambridge CUP*.
- Marsh, A. (1990) *Political action in Europe and the USA*. London: Macmillan.
- Thurstone, L.L. (1928) Attitudes can be measured. *American Journal of Sociology*, 33, 529-544.

Thurstone, L.L (1931) The measurement of social attitudes. *Journal of Abnormal and Social psychology*,26, 249-269.

George Gaskell
Dept of Social Psychology
London School of Economics
Houghton St
London
UK