



**Wulf Schlund**, graduated psychologist, has been working in market research for 33 years. He worked for Research International, IfD Allensbach, Schwartauer Werke and Braun, before he started his own business: 1978 till 1992 ISM, 1993 till 1998 ISM The Research Alliance, since 1999 (after selling the name) ISM Global Dynamics. Today Wulf Schlund is major shareholder of ISM Global Dynamics and President of Global Dynamics.

**A**t a time when all the talk is of changing consumer behaviour and the possibility of new approaches in market research, it is only normal to put practicalities on the back burner. It's simply not fitting to start rummaging around in measurement problems again, problems which are as old as the measurements and scales themselves.

But if we *do* do it, the reasons are fairly simple: We have yet to come up with a solution to the problems. We have ignored them – practically buried them. All scales used until now in market research have had their advantages, but they have also had their disadvantages. And there has been enough research on the pros and cons of each scale. Rather than sieving through the issue again we can summarise by saying that whenever the benefits of a scale were recognised, somebody eventually found fault in it and people moved on to the next one until somebody picked holes in that one too.

The main issue here is not the specific

**Wulf Schlund**

## Swept under the carpet again: International scaling

problems presented by individual scales. It is much more important to get to the root of the scales' problems and name the reasons why every scale we have used up until now is insufficient.

### Problems setting the points

What is the ideal number of points on a scale: 2, 4, 6, 8, 10 or 3, 5, 7, 9, 11? These numbers more or less cover off the spectrum of possibilities used for scales in market research. Each range has its followers, as well as its advantages and of course its disadvantages. Even if it is obvious whether specific disciplines and measurements need a smaller number of points than others which need a larger number of points, arguments and disagreement also within disciplines are going on. So what we have now is a colourful mix of different scales doing the rounds.

### The name game

By the very latest, with a four point scale you can expect to run into problems with names. We have all benefited from numerous attempts to standardise and even validate scales, and it was an important process. But it only really served to provide us with fixed names and whenever different names were used we ended up with different results. And each set of names caused differences from country to country and from culture to culture.

### One pole or two?

Do one-pole scales provide the same results as bi-polar scales? Are bi-polar scales actually even allowed if opposite poles obtain other names than pure negations? The cynic would say that bi-polar scales are just better for presentations and should not even be used to

measure attitudes, beliefs and images. In any event they only exacerbate the problems of international comparability.

### Coping with differences

Within disciplines, or when trying to measure the same thing, the co-existence of so many different scales creates more bother than it's worth and is more grief than actual use. It basically goes to prove that despite decades of long-winded research there is still a lot of uncertainty on how best to measure things. It is far from ideal. If anything it creates more problems: It is preventing us from dovetailing existing data more.

### The problem with all graded scales

All graded scales have a problem in common. They look like interval scales and are sold on that basis, but aren't. This is nothing new or controversial. But in the absence of a better alternative sweeping the problem under the carpet is making things worse. Market researchers the world over have come to terms with the situation and have simply shrugged their shoulders and carried on using some or other scale to measure attitudes, images and propensity to purchase.

We had enough reason to tackle this problem again, with market research gaining in importance internationally. The difficulties already encountered by using false interval scales worsens when you enter the international arena. We only need to examine what has happened as a result of leaving problems unsolved:

- The pattern of responses depends on the verbal anchors placed along the point scale.



- Insufficient differentiation on the positive end of the scale.
- Different response patterns in different countries.

After becoming accustomed to wrong scales the whole time we were turning our back on the problem, and having become quite fond of them, it is quite an arduous process to remind ourselves of the unfortunate consequences for measurement theory. Below, four examples have been plucked from the plethora of unfortunate consequences, none of which are new to us:

- Parameter procedures for measuring the significance of differing results are basically a no if you use wrong interval scales – i.e. the scales used in market research. They should not be used.
- The same applies to all parametric regressions and correlations, upon which the whole of multivariate statistics are based.
- Creating total scores across a number of items is inappropriate.
- Results from international studies are basically not comparable from country to country, or culture to culture.

It would be difficult to identify the worst factor, all of which are related to each other. But as now more than ever we have to ask whether a European consumer already exists, or whether there will not be one for another ten years, the problem which causes us the most inconvenience is the lack of internationally comparable results.

Unfortunately we have not even researched whether consumers in the North German State of Schleswig-Holstein or Hamburg deal differently with a five point scale with named points than consumers in the Southern German States of Baden-Württemberg or Bavaria. We have indeed found significant differences in this respect between the Germans and the Spanish – also incidentally between the Swedes and Italians. Diagram 1 shows a typical example of parallel studies in Germany and Spain:

What the graphs show only too well is how difficult it is to make a comparison between German and Spanish research results even though they are on the same subject and use the same measurement tools. It is difficult enough trying to interpret the results. How can you interpret such different results:

- Do the Germans feel so much pressure to conform or are individual opinions just more homogeneous?
- Are the Germans more disciplined in making judgements or are they just not as interested in the survey question?
- Are they just less enthusiastic or scared of telling the truth?

This single example highlights the problem well enough. It is not difficult to imagine that it would be no easier to compare European results with those from Asia, where even between Japan and China there are noticeable differences in how they see scales. Inciden-

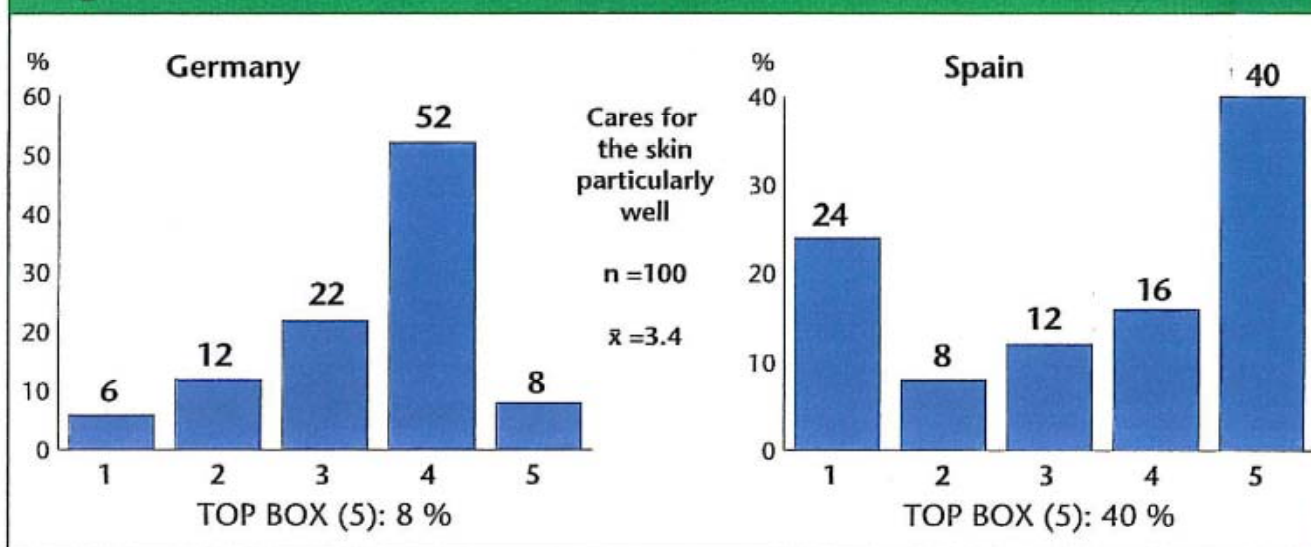
tally there are still such differences to this day between former East German and West German consumers even if they are not as marked as they still were up to five years ago.

But is this problem really that bad, or should we just treat it less seriously?

You might be justified in answering this question with a *yes*. Never before have the United States of Europe been so close. We are on the verge of introducing the Euro. Market research has been working with the wrong scales for such a long time, not bothering with the international incomparability of results, that with the light at the end of the tunnel there seems no real point in changing things now.

Of course some people see it differently: first, we haven't got a clue how long it will take before *the* European consumer emerges, if at all. Second, it is not market research's job to invest in long-term studies to find out whether and at what rate the Italians, Spanish, French, Germans, British and Swedish are starting to think the same way. Market research is too busy taking quick measures of product acceptance, concept, understanding and advertising effects, or maybe even the growing similarities between lifestyles. But even for this it would be good to have the right measurement tools without the flaws of producing differences between consumers in different countries which have nothing to do with the question at hand.

**Diagram 1:**





We at Global Dynamics also share in the latter view and it is the reason why we have never given up our search for a better way to score internationally than is possible with current scales.

### Can we conceive of a solution to the problem?

People in the USA and England have tried to find a solution to the problem with psychometrics. In the process they came to a solution which is remarkable if not only for the fact that it is quite radical. Their solution was to do away completely with a scale. The psychometric experts consider it permissible – or, according to measurement theory, justified – only to make paired comparisons, or at most triad rankings (freely placed, i.e. not on a scale). From their point of view this data collection method is the only justified alternative to point scales for measuring attitudes and images as well as purchase propensity. Only data generated in this manner is valid and can be analysed and worked up into the actual distance between subjects.

This somewhat extreme solution can not be criticised from a measurement theory point of view. But in practice it does cause significant problems; or in other words the solution can rarely be

put in practice. You just need to imagine an image study to be carried out with this method into five brands with 20 items. An image scale which took five minutes up until now would have to be re-arranged into a one hour work session. A proper 30 minute interview would be a half day workshop. This sort of research would start to become so expensive and take up so much time that clients would accept neither the timescales nor the costs. It is an extreme solution but it has the advantage of being able to summarise the problem to a T. But unfortunately the disadvantage is that it is just not practical.

We now find ourselves facing quite a problem. We know what the problem is, but don't really know what to do about it. We will have to keep looking. But in so doing the impractical, extreme psychometric solution was at least of benefit to us in providing a more specific way than before to define the *parameters* for the solution. Now

*This article brings the principle problem of using scales up for discussion. The main issue here is not the specific problems presented by individual scales, it is more important to get to the root of the scales' problems, the problem all of them have got in common. We can't simply continue ignoring this problem in the era of global research, knowing that no existing scale in use allows international comparison. Will there be a way out?*

that we know these parameters the next step will be to focus more closely on a solution which fulfils these parameters in full whilst still only needing a reasonable amount of time and money. By considering the problem and the somewhat extreme psychometric solution outlined above we have come up with five parameters for a new and implementable solution. They are as follows:

- No verbal scale.
- No point scales whatsoever.
- A measurement tool without predefined distances.
- Spatial freedom to place subjects.
- Matching current scales on time and cost factors.

Even reading these parameters, the task still almost seems as hard as turning a circle into a square. But there is a solution. And we believe we may have struck upon it and are currently testing it out in key areas, before validating it.

ISM  
GLOBAL DYNAMICS



Too bad that currently used scales do not allow international comparison.

SCS, our new Stimulus Comparison Scale, does.

SCS®

The better tool.

+49 6173 92480

+49 6173 1088

info@globdyn.com