

Dirk Frank and Wulf Schlund

Packaging research revisited

A method for international research

Packaging is one of the most important marketing instruments. As an essential part of the marketing mix it deserves the same level of attention from a brand manager as advertising, pricing and the product itself. The myth, charisma and success of “super brands” like Coca Cola or Marlboro are well-known testimonials for the effects of a unique packaging.

The marketing functions of packaging

A good packaging has to serve the brand in more than one way:

- Based on a clear appearance it makes sure that the brand is recognised quickly.
- In connection with a high *brand recognition* it gives the brand a strong impact in the shelf against the competition.
- By signalling the USP it communicates the brand’s uniqueness.
- In doing so it supports and reinforces the brand core.
- It nourishes the brand affinity in the form of brand preference and purchase interest.

- It strengthens the emotional loyalty to the brand.
- It maintains consumers’ trust in the brand.
- It provides space for a premium price strategy by its high value impression.
- It supports the sympathy for the brand by an esthetical design.

This list clearly focuses on the communicative functions of the packaging. Functional aspects of a pack normally are obvious and easier to test: often together with the product itself. Sometimes the functional aspects can also be included in a pack test which deals mainly with the communicative functions of the pack.

Out of the list presented above marketing managers can evaluate their existing pack testing approach regarding its methodological power: How many of the functions listed above does the approach answer in a valid and unbiased way?

Pack test methods so far

There is a lot of pack testing methods: from qualitative explorations via communication tests in a competitive context, impact tests, diagnostic shelf tests (studio and in situ purchase simulation) and there are tailored combinations of these methods.

Among the more psychometric approaches a differentiation criteria is the way

the alternative design routes are presented: Does the consumer see only one pack design (*monadic approach*) or several designs (*comparative test*)? Within the sector of comparative testing you find the *direct comparison* of several alternatives by simultaneous assessment as well as the *successive assessment* of the alternatives.

The group of monadic approaches has a common idea: Similar to the real market situation the respondent sees only one pack design. After the assessment of this one pack design sometimes further design variants are shown for evaluation. This kind of compromise is called *semi-monadic* approach.

The big question – in particular for the semi-monadic approach – is whether to *include the competition* or not. In practice there is a wide range of procedures from the complete lack of competition via the pure presentation as a frame of reference up to its full evaluation. Full or partial evaluation of the competition brands provides more information and makes sure that the respondent is not aware which of the brands shown is the test brand.

The strengths and the weaknesses of these different approaches ranging from comparative to monadic testing have been listed and discussed often enough. The fundamental difference between mo-

monic and comparative testing can be summarised in the following question: *Do we want to treat the consumer as a marketing expert or as a consumer?* If you prefer to treat the consumer as a consumer, your pack test approach should be a monadic one.

A monadic pack test means as many samples as pack designs to be tested. This is expensive and it becomes even more expensive when you include the current pack design with an additional test sample. But the higher costs compared to the comparative approaches are more than justified, because what you obtain is valid results. Also to include the current pack design is more than worthwhile: You need to make sure that one of the new design alternatives is better than the current one and you also need to know in which way. How else could you find out than by also testing the current pack design?

Lacks and limitations

It does not matter whether monadic or comparative testing, the pack test approaches so far have been limited to investigating only segments of the criteria mentioned above. No one of them includes all of the important criteria.

Among the limited proportion of criteria pack tests have been measuring so far is the pack's attractiveness, the pack profile, perhaps the purchase interest and

the brand image. But the pack's effect on the brand image has never been separated into the current brand image and the influence of the pack design.

Criteria such as *brand recognition* have never been investigated in the same test. Brand recognition is sometimes investigated in a separate test, however, not very often as this involves considerable additional cost. Another important measure of pack performance is the *shelf impact* which was rarely tested at all for many years. Shelf impact means the ability of a pack to stand out from competition in a typical shelf situation.

In particular for international pack testing there is a number of further limitations:

1. Test designs and measurement procedures are not yet standardised enough for a valid international comparison.
2. Traditional rating scales are not really suited for international comparison.
3. Testing prototypes of pack designs in several markets has been too expensive and troublesome: production costs and the risk of damage and loss.

A new test system for international use

There is no doubt, the lacks and limitations listed above mark a clear need for improvement. In particular

international pack testing requires some better solutions than so far. At GLOBAL DYNAMICS we investigated all these problems step by step to find a solution. Finally we found a solution for each problem and combined them to a new pack test system. These are the main characteristics:

- An interview flow which includes definitely more test criteria than ever before.
- An experimental design for separating brand and pack effects on the brand image.
- A new scaling instrument much better suited for international use.
- The option of multimedia software instead of traditional mock-ups in a quality which allows its application in nearly every market.

More criteria within one test

Maybe surprising at first glance but a matter of fact is the number and variety of criteria the new test system can include. The challenging task here was to avoid any bias effects from one criterion to another. And this is about the typical interview flow of the system:

1. Brand recognition

- Which brands seen?
- Any products in a new packaging? Which ones?

- New elements of this packaging
- Remembered details of single brands

2. Brand attraction effects

- Brand affinity/purchase interest towards all brands shown
- Brand choice

3. Communication effects

- Brand communication/image effects
 - Brand personality
 - Product expectation

- Optional: price expectation or brand price trade off

4. Appeal

- Appeal/attractiveness
- Characteristics profile
- Optional: comprehension of information

5. Shelf impact

- Speed of recognition
- Accuracy of recognition
- Brand confusion

As it has been unusual to include a *brand recognition* test and to measure the *shelf impact* of a new pack design together with all the other packaging effects within the same interview, we briefly outline the character of these two modules

and afterwards explain the methodology.

The brand recognition test answers the question whether the new pack design lets the consumer still recognize the brand. At the same time it discovers which of the pack's characteristics communicate the brand foremost: the shape, the colours, the logo or the typeface. The brand recognition test simulates the consumer's first contact with the product in a realistic way by a short time exposure among the major competition. The packs are shown in a one-to-one size in the form of a slide or by a LCD-projector. For a typical

experimental arrangement: a series of slides or computer pictures is shown to the respondent, each of them showing a shelf with the packs of different brands. Here the relevant competition is included. Some of the slides include the test brand, others do not. The respondent's task now is to decide for each slide whether the test pack is present or not, as quickly as possible. As soon as the respondent ascertains this information, he pushes a button and the finding time is automatically recorded.

Only then the respondent tells the interviewer whether the test pack was in or not.

This way for



display see figure 1.

Figure 1

The brand recognition test can also be used as a *product differentiation* test within a range of products of the same brand.

The *shelf impact* measurement within our system is done by a standardised

each shelf arrangement you obtain two measures: the speed and the accuracy of recognition. It is on purpose that the *standout* test is conducted at the end of the interview: This way the respondent becomes familiar with the new pack design, which compensates for the advantage of the current

pack in terms of prior awareness. And these are typical results of a shelf impact test conducted for Hardenberg (see figures 2 and 3):

tion was much slower for the new design: the Doppelkorn drinkers needed 1.97 sec on average compared to 1.36 sec for the current design. As beautiful as the new label design

For the methodology it is critical that the brand recognition test is done at the beginning of the interview and the shelf impact measurement at the end of the interview. The whole test

has to be done as studio interviews using simulated shelves. As a monadic approach this pack test needs as many cells as pack designs to be tested with 120 respondents each (see figure 6).



Figure 2: Doppelkorn “old”

In a packaging test for Hardenberg Doppelkorn a new packaging design was tested and compared to the current label. Among other criteria also the shelf impact was investigated. The new design achieved better results than *current* in a number of criteria, but had the weaker shelf impact (see figure 4). The current design was identified with much more accurate certainty when present (97 per cent versus 87 per cent) and less often mixed up with another brand when not present (90 per cent versus 98 per cent). In particular the speed of recogni-

Figure 3: Doppelkorn “new”

was, its branding was too weak (see figure 5).

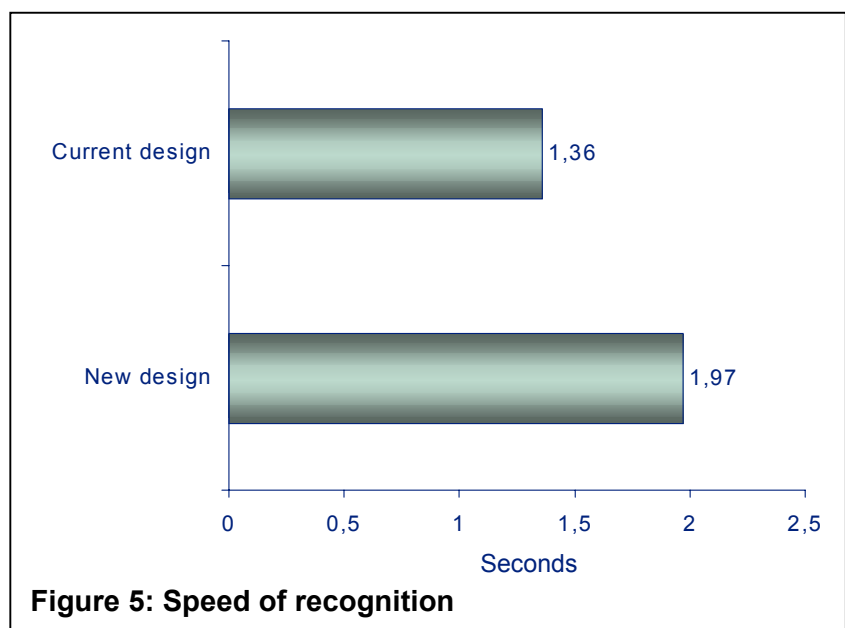


Figure 5: Speed of recognition



Figure 6

The approach is strictly psychometric avoiding both unnecessary qualitative and utopian quantitative elements (for example in situ purchase tests for which you would need samples of 400 respondents per cell).

Separating brand and pack effects

So far pack tests have never investigated a pure image effect achieved by the pack on the brand's image.

We do this in the following way: We are using an experimental design which lets one part of the test sample at the beginning rate the existing brand image as based on just the brand's name. The other part of the sample rates the image communicated by just the pack design which does not contain the brand's name. Later on we investigate the brand image based on the branded pack. By this procedure we find out to which degree the new pack design is suitable for the brand (see figure 7).



Figure 7

An international scaling instrument

Psychometricians and other measurement experts have always argued against the existing scaling instruments and these are their main objections:

- Poor discrimination
- No real interval scales
- Artificial correlations
- Cultural differences in scale perception: grades and naming.

Out of these points of fundamental criticism we concluded a particularly strong need for improving the situation for international research. So in co-operation with American and English measurement and scaling experts, we have developed a new scaling instrument which is able to bring down cultural biases to a theoretical minimum: the SCS Stimulus Comparison Scale (see our article in *planung & analyse* 6/2000). This scaling instrument avoids the major points of criticism as on the respondent's side it is no real scale, but just a distance without any grading, but with a positive and a negative pole (see figures 8 and 9).



Figure 8



Figure 9

Subjects are rated by moving a marker on the scale. For particular applications a number of markers are used. The trick of the scale is that at the back of it the interviewer can read scores immediately and as quickly as he would note down a number on a traditional scale.

- The markers allow a similarly quick administration as traditional scaling procedures.
- Marker signs ease the notation for the interviewer.
- A 20° inclination makes the back invisible for the respondent and improves the legibility for the interviewer.

The advantages of the scale over the classical rating scales are obvious:

- No grades, no names which could cause cultural perception differences.
- More valid classifications: no halo effects, less response patterns
- Better differentiation: full use of the whole scale range, no clumping on the positive side.

In the meantime a number of validation studies have confirmed the expectations, which were based on measurement theory. It looks, as if the SCS is in fact the most suitable scale for international use.

Also for pack testing technological progress has brought considerable new options. So we do not have to spend any longer too much money and too much time for producing expensive mock-ups which can be damaged easily. We – or the client's design agency – can now produce multimedia pictures in a quality which allows their application in practically any market.

In the area of pack design, the multimedia technique allows traditional research techniques to be greatly enhanced by showing or even creating new three-dimensional packs on the screen. With either plain physical packs or fully labelled branded products, rotation, allowing reverse

labels, usage instructions and other details to be studied. With the computerised approach, all of the cost, timing and other problems of making and transporting pack mock-ups are avoided.

Abdruck mit Genehmigung des Deutschen Fachverlages, planung & analyse, Mainzer Landstrasse 251, 60326 Frankfurt am Main, Tel.: 069-7595-2019, Fax: 069-7595-2017, redaktion@planung-analyse.de, www.planung-analyse.de