

## **Validity in market research practice: 'new' is not always 'improved'**

Associate Professor Linda Brennan  
Australian Graduate School of Entrepreneurship (AGSE)  
Faculty of Business and Enterprise  
Swinburne University of Technology, Melbourne  
Cnr Wakefield and William Streets, Hawthorn  
PO Box 218, Hawthorn, Victoria, 3122, Australia  
Telephone: 61 3 9214 4861  
Facsimile: 61 3 9819 2117  
*e-mail:* [lbrennan@swin.edu.au](mailto:lbrennan@swin.edu.au)

Jim Camm  
Associate Director, Automotive Practice Group  
The Nielsen Group  
Level 5; 312 St Kilda Road  
VIC 3004, Melbourne, Australia  
Telephone: 61 3 9251 3205  
Email: [jim.camm@acnielsen.com.au](mailto:jim.camm@acnielsen.com.au)

Janusz K Tanas  
PhD Candidate, President Student Research Council  
Australian Graduate School of Entrepreneurship (AGSE)  
Faculty of Business and Enterprise  
Swinburne University of Technology, Melbourne  
Cnr Wakefield and William Streets, Hawthorn  
PO Box 218, Hawthorn, Victoria, 3122, Australia  
Telephone: 61 3 9214 5874  
Facsimile: 61 3 9214 8381  
*e-mail:* [jtanas@groupwise.swin.edu.au](mailto:jtanas@groupwise.swin.edu.au)

**Keywords:** Validity, methodology, market research

## **Validity in market research practice: ‘new’ is not always ‘improved’**

### **ABSTRACT**

Market research suppliers want to provide a differentiated product or service offering to their clients. However, the frequent need to appear ‘new’ results in adaptation of research to different contexts. This has consequences for validity. In this paper we make the case for tempering the existing enthusiasm for constant ‘innovation’ in market research methodologies. In particular, we argue that unfettered adaptation can lead to the generation of invalid findings. We demonstrate some methods of how invalid results can develop. Finally we explain current concepts of ‘validity’ and provide an innovative way of showing the relationship between the dozen or more varieties of validity that are commonly used in the literature(s). This paper aims to remind practitioners and academics alike, that concepts of validity are important and that there is no point in having a perfectly reliable, but completely invalid, measurement tool.

### **INTRODUCTION**

Innovation is a necessary precondition for progress and subsequent advancement in any field. We accept this as self-evident. At the same time it is legitimate to ask the question whether or not all innovation is ‘good’ for the field of market research, or can innovation lead us away from good practice in research?

### **THE PRESSURE TO APPEAR NEW AND DIFFERENT**

There is, no doubt, a pressure from clients of market research to innovate. There is also pressure on marketing academics to make ‘new’ contributions to the body of knowledge (BOK). While the focus of this paper is on market research suppliers, the lessons presented here are apposite for academics as well.

The problems of marketing are complex and often appear to be intractable, relating as they do, to the behaviour of consumers. Therefore, there will always be some unfinished business, some unturned stone or a yet-to-be-explored angle awaiting every brand and marketing manager. Furthermore, there is constant pressure on these same marketers to get an edge on their competitors. Competitive advantage does not come about by doing more of what the competition is already doing. It comes about by being out there and taking the lead in the market, being creative, taking risks, and dealing with the consequences in a complex and rapidly changing business environment (Hult, Ketchen, and Slater 2005; Vanderkaay 2005). Indeed, there are more success stories in marketing that relate to risk taking and imagination than there are relating to in-depth analysis and understanding (Temporal 2006).

On the market research supplier side, there is the desire to differentiate and position their offerings in order to improve margins. It is a truism in marketing that undifferentiated,

commodity-like offerings have the lowest margins. Related to the complexity of the marketplace is the observation that it is often the case that 'out-of-the-box', validated methodologies do not match the specific needs presented by a new marketing problem.

One may state that on both sides, there is also a desire to reduce costs. However, the risk of extreme cost reduction strategies is the standardisation of the service, which in turn can also encourage the application of a methodology to an inappropriate situation. It is cheaper to provide a pre-packaged solution to the client's research problem than it is to develop a new methodology. Questionnaires, recruitment strategies and overarching methodologies are often 'leveraged' from one project to another and the cost of development amortised over several projects.

All these factors combine to create demand for something that appears new and different (at least to the client). Via the mechanism of competition, the research community supplies it by offering innovative methodologies. Barker (2002 p 152) put it succinctly:

'...the world of commercial qualitative research is awash with apparently new techniques, new models, new approaches. The self image of the sector is of creative developers, inspired gurus, bearers of always new wisdom and unique insight to the tables of brand managers and advertisers. The commercial pressures to innovate (occasionally for the sake of it) are indeed very great – client research briefs often literally beg for new ways of approaching problems. We know our competitors will be proposing living with a sample of families for two weeks in order to gain those hidden nuggets of insight, which will provide the leverage that brand of cat food really needs; so we need to go one step further, one step wackier.'

Research can be boring (McAnena 2004). However, the quest for imagination in research has been with us for a long time. In 1961 Blankenship (p 34) wrote: '*Creativity in research provides no defense (sic) for use of inadequate research standards.*'

### **MARKET RESEARCH CROSSES DISCIPLINARY BOUNDARIES**

There are many sources of innovation in market research (innovation in this circumstance being the presentation of something new to the client or within the context). First, there is the transfer of methodology developed in other fields into marketing research. Second, there is transfer of knowledge. For example, concepts in psychology are adapted into market research. Finally, there are extensions of existing techniques within the domain of market research itself. The question arises – what mechanisms are in place to ensure that such adaptations of theory and method across disciplinary boundaries are applied appropriately? Is the expertise developed in the (for example) psychometric domain directly applicable to the matters embedded within the advertising research domain? Of course, the answer to that is; "it all depends." Academics and practitioners alike must be in a position to evaluate when it is appropriate to adapt methods and theories from one domain to another. This would, of necessity, require some research training in a multitude of disciplines.

On the whole, market research is a theory borrowing discipline, much like the better documented case for marketing (Murray, Evers, and Janda 1995; Deshpandé 1983). Market research practitioners bring their theories from a wide variety of disciplinary

backgrounds; ranging from applied statistics to anthropology. While there is value in the variety of approaches to solving problems offered by this diversity, there are dangers that researchers will adopt ideas from others without a full understanding of the theoretical underpinnings.

Within market research, therefore, there is a large number of methodologies and approaches that have been borrowed from other domains and used with new populations, situations and contexts. In addition, both these borrowed methodologies and those developed within market research are ‘tinkered with’. By this we mean that practitioners (and possibly academics) change the methodology without considering the effect it will have on the validity of the methodology. For example, ethnography is a trendy ‘method’ that is applied as a term without much thought as to the ontological and epistemological foundations, nor the methodologies that were originally validated in the discipline. Participant-observation, for example, does not mean simply observing; it means actively participating, reflecting and observing, writing about the observation and being assimilated into the community under consideration. Most clients won’t pay for this kind of research so it rarely takes place; although there is much non-participant ‘observation’ that is passed off as ‘ethnography.’

These borrowings and tinkering helped shift the market research discipline into a credible and scientific status within the domain of business decision-making. However, we argue that the market research discipline got lost on the way to true scientific status.

This is because:

1. There is pressure to constantly produce something that appears new – even if it is just a name change, and
2. Opportunities exist to borrow and tinker with valid methodologies – there are no limits to the number of adaptations that can be made to a method

which leads to the proliferation of methodologies that have not been validated or revalidated for the new populations, situations and contexts they are used for. This in turn leads to the generation of research findings that might or might not be correct.

### **WAYS THAT ESTABLISHED METHODS CAN BE INVALIDATED**

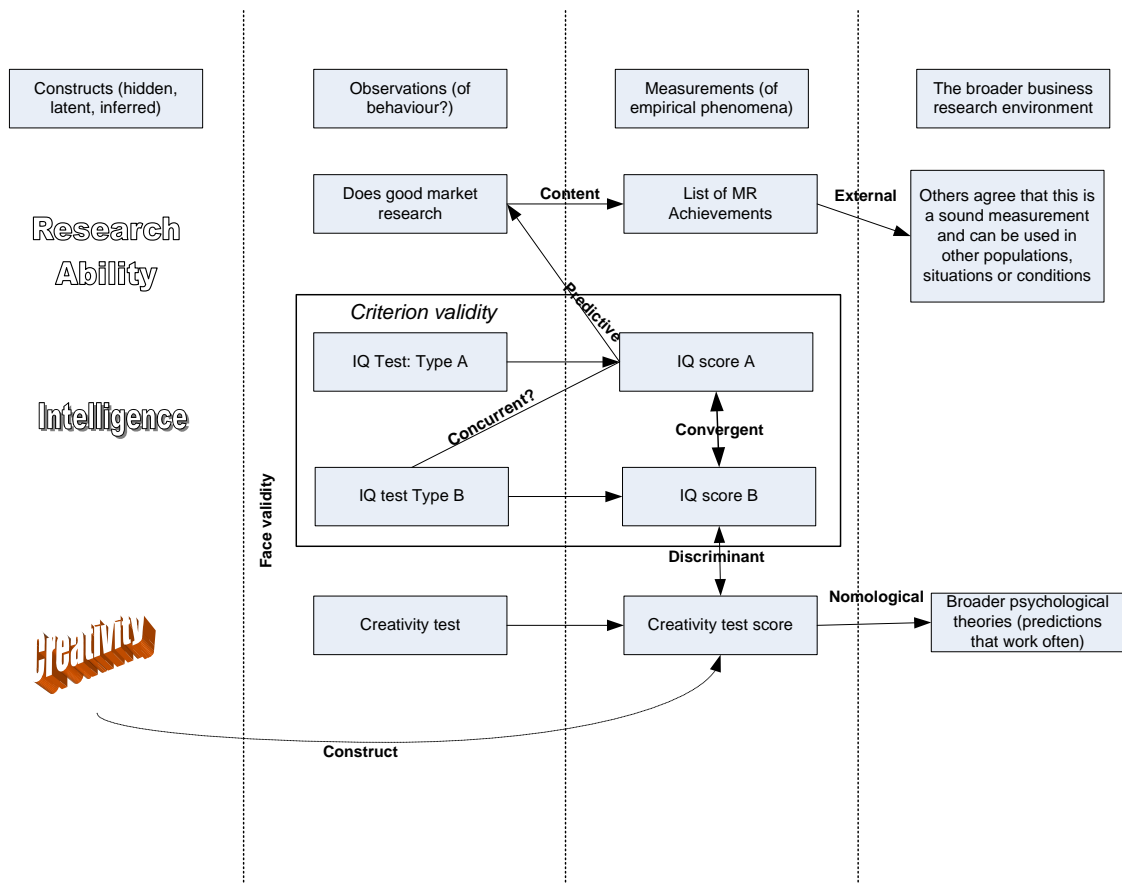
This argument rests squarely on what we mean by ‘valid’. Therefore, we now review the basic concepts of validity. Figure 1 illustrates in a novel way the many types of validity as summarised from the typology provided in Appendix A. The concepts of validity and reliability hold a central place in measurement science. In this paper, we put reliability to one side and discuss only validity. We note in passing, however, that there is confusion in the literature with many authors apparently using the terms interchangeably. For now, a working definition of reliability is ‘that a measure will yield the same result if repeated under the same conditions’ and a working definition of validity is ‘that a measure measures what it purports to measure’. An early example of a validation issue in market research is provided by Blankenship (1961):

'Many years ago a researcher didn't believe the answers he was getting on direct questions as to the publications people read. Everyone reported 'prestige' publications; few admitted to readership of the low prestige type. He got a very different picture when he had solicitors collect magazines from the same homes "for charitable purposes." What people said they read, on direct question, and what they actually read, on the basis of magazines in the home, were quite different things' (p 35).

Quite simply, the methodology of direct questioning was found to be invalid. It did not measure what it purported to measure. It does have what is known as 'face' validity but fails the test of content validity.

Face validity exists when the measure 'looks as if' (on the face of it) it *should* measure a particular attribute (McGartland Rubio and Kimberly 2005). This is usually easy for others to agree with, so it is also called consensus validity (Heeler & Ray, 1972) or logical validity (McGartland Rubio and Kimberly 2005). Both terms are often used synonymously with content validity (Sireci 1998). However, face validity does not imply expert acceptance of the measure's validity as does content validity. Interestingly; there is a large volume of research based on the assumption that face validity has been established before all measurement takes place. However, because it is such a base assumption, there is no way to test face validity outside of the context. We argue that the expression face validity has two meanings. First, a face belongs to an individual idea. For example in Figure 1 – the construct of 'intelligence' is a single idea, which may or may not be uni-dimensional. Second, 'on the face of it' means simply that it '*looks as if*' it would be OK to measure a phenomenon. However, we never ask - who does it look that way to? In the establishment of face validity, an important but often missing question is - whose face? Whose perspective are we looking from? For example, few people in Melbourne (Australia) would argue that a 24 degree day with sunshine is a 'nice day'. We could 'on the face of it' establish 24 degrees and sunshine as two separate variables of the construct 'nice day.' However, if we do not know which temperature measure we are using (Fahrenheit or Celsius) then the 'nice day' could be very different depending on whether you are located in the US which uses Fahrenheit or somewhere which uses Celsius. Furthermore, artists might prefer clouds and brooding skylines, pilots might be concerned about invisible turbulences and drought stricken farmers might prefer lots of rain. Thus, perspective is important in the establishment of validity. Nunnally clearly outlined the implications of face validity being established within a context rather than within a set of measures in 1967. However, this excellent advice seems to have been lost somewhere along the way.

The usual method of establishing face validity in the social sciences is often a relatively cursory process (Rossiter 2002). Most validation is done with university students and academics (Basil, Brown, and Bocarnea 2002). These people are not representative of the general population. Market researchers are usually adamant about random sampling but if face validity is established with people who are not representative of the population then the first step is flawed.



**Figure 1: Relationships between ‘Validities’**

Figure 1 illustrates that only face and content validity can be established in the absence of other measures because all other forms are defined in terms of their relationship to other variables or measures – thus content and face validity are ‘formative’ in nature. That is, validity of one measure is based on the supposed validity of another measure. We argue that if the first ‘validity’ is not well established, then the rest of the process becomes meaningless.

In order to describe these relationships, imagine a hypothetical group of market researchers who have both Intelligence and Creativity (well don’t we all?). In the diagram these are shown under the heading ‘Constructs (hidden, latent, inferred)’. In the next column we have ‘Observable Behaviour’, which includes performing two different IQ tests and a creativity test. In addition, we note that these market researchers do ‘good’ market research. In the next column we show measurements that could arise relating to each of the behaviours. For instance, IQ test score arise from undertaking IQ tests and we could assemble list of market research achievements for each researcher who does good market research. The rightmost column represents aspects of the broader environment.

Let us imagine that these market researchers first do a test that purports to measure their creativity and obtain a creativity test score:

**Construct validity** is the extent to which the test score (outcome of measurement) is a measure of what we understand by 'creativity' (Cronbach and Meehl 1955; Smith 2005). The creativity test is made up of a large number of individual test items. In this sense, creativity is not something which can be directly observed, it must be inferred from some other observable behaviour. The more abstract the concept, the more difficult to assess validity (Nunnally, 1967):

**Discriminant validity** is the extent to which the IQ and Creativity tests differentiate between the two constructs 'creativity' and 'IQ' (Campbell 1960). That is to say if we believe that a person could have high IQ and low creativity, or vice versa, then the tests will be able to show this:

**Convergent validity** is the extent to which the two IQ tests provide the same results (Reichardt and Coleman 1995; Campbell and Fiske 1959). That is to say if some of our market researchers scored highly on one IQ test but obtained a low score on the other we have cause to doubt that the tests are measuring the same thing. Let us assume that there is a belief that a high IQ score indicates that the researcher will perform good market research:

**Predictive validity** is the extent to which the test score for IQ predicts how well researchers conduct market research (Cronbach and Meehl 1955; Nunnally 1967). It might well be the case that IQ does not correlate at all with good quality market research and it is found that creativity is a far better predictor. The question then arises of how we know whether a researcher is doing good work? The measure we have suggested is to assemble a list of market research achievements. Researchers with longer lists of achievements would then be said to have done better research than those with shorter lists of achievements:

**Concurrent validity** is the extent to which the test score could, for instance, differentiate quantitative researchers from qualitative researchers (Heeler and Ray 1972). It is similar to Predictive validity, but does not imply future activity. Predictive validity and Concurrent validity are usually considered to be variants of Criterion validity (Diamantopoulos 2005). In our researchers' situation, an IQ test would potentially differentiate between levels of verbal and/or numerical skills.

**Criterion validity** is the measurement of how well one variable (or set of variables) predicts or correlates with an outcome based on information from other variables (includes concurrent, predictive and diagnostic validity) (Diamantopoulos 2005). In this case our IQ tests are related to each other but may be used to determine different elements of IQ depending on the needs of the person testing. One variable (eg numerical skills) might be sufficient in some circumstances and can be used as a 'valid' alternative to a battery of tests.

**Content validity** is the extent to which a test is made up of a representative sample from the domain of interest (McGartland Rubio and Kimberly 2005). We note that the issue of circularity, which is inherent in many aspects of validity, now arises. How do we know that the items we have chosen correlate with ‘good market research’ unless we already have an idea in mind for what constitutes ‘good’ market research? This is a vexing question. It is partly answered by the concept of ‘face validity’, which is not shown as a set of relationships on the diagram because it should be applied in all cases:

**Face validity** is the extent that the measurement measures the quality that it purports to measure on the basis of a common sense assessment (McGartland Rubio and Kimberly 2005). Our hypothetical group of market researchers are only a small subset of all market researchers. Having established that the ‘list of achievements’ is a valid measure of quality of market research, we can now ask the question whether or not the idea of using this type of measure is applicable for the broader market research community, or even for other professions. For instance, is ‘list of achievements’ a valid measure for the quality of social work, police work or sales?

**External validity** is the extent to which the results can be generalised to other populations, situations or conditions (Peter 1981; Redmond and Griffith 2003). Finally, we return to our IQ scores. It is reasonable to ask whether the ideas we have developed – that creativity is different to IQ and that IQ, not creativity, is a good predictor of quality of market research as measured by lists of achievements – fits in with broader psychological and social theories.

**Nomological validity** is the extent to which the pattern of results for a theoretical network is consistent with broader theoretical networks (Peter 1981). For instance, it could be the case that there is strong evidence in another domain, say, sales, where it has been shown that IQ does not predict performance at all and that lists of self-reported achievements are poor indicators of sales ability.

In Appendix A we provide a more detailed review of the validity literature, and introduce the classifications ‘Formative Validity’ and ‘Prognostic Validity’ as organising concepts for the various measures of validity.

### **SO IF WE KNOW ALL THAT, WHERE DO THINGS GO WRONG?**

Now that we have established some of the concepts relating to validity, we turn now to three specific examples of how methodologies become invalid through ‘tinkering’ with methodologies.

#### Example 1: Scales

The workhorse of market research is the questionnaire, and one of the most important aspects of a questionnaire is the use of scales. There is established methodology for devising and validating scales; (cf. Celsi et al. 1992; Salzberger 2000; Rossiter 2002; Churchill 1984; O'Connor and Eskey 2005; Gerbing and Anderson 1988). We ask the question, how often is this established methodology used to devise scales in market research? The design and construction of questionnaires is prefaced on face validity being

This paper is an early draft of a paper which eventually appears as:  
Brennan, L., Camm, J., & Tanas, J. (2007). Validity in market research practice: ‘new’ is not always ‘improved’. *Der Markt*, 46(1-2), 3-12.



established ‘up front.’ Face validity can be established by asking the *population of interest* to confirm the researchers’ ideas of ‘what looks as if’ it might be correct and to make adjustments to the scales as necessary. Let us now assume that a ‘validated’ scale is pressed into service for a new population, situation or context. Very few clients are willing to pay for the required experimentation to re-validate the scale and very few suppliers are willing to admit that they are not sure if their proprietary instruments are valid in the new context. Thus, we inevitably have a drift in the certainty of the outcomes. Unless you test it within the population, situation and context you cannot be sure that it ‘really’ works. Nomological validity implies repeated measures. For example, the SERVQUAL scale (Parasuraman, Zeithaml, and Berry 1985) is still being used consistently to evaluate ‘service quality’ in different contexts (cf. Donnelly et al. 2006; Badri, Abdulla, and Al-Madani 2005; Gounaris 2005) although issues with its discriminant validity were identified as early as 1990 (Carman).

#### Example 2: Projective Techniques

Projective techniques are used extensively in qualitative research. We have no issue with the use of stimuli to promote discussion and to help respondents express their feelings and emotions. However, projective techniques are often taken further; whereby images chosen by respondents are interpreted by the market researcher. This interpretation by the market researcher introduces the full range of validation issues. It may be stated that a detailed discussion is beyond the scope of this paper, but let us just provide two simple cases. First, it is a commonly held belief that colour can be validly interpreted. For instance, red is often thought to mean aggression whereas green means tranquillity and furthermore that these attributes can be inferred when respondents associate these colours with products or services. The issue here is whether the *interpretation* is valid. But psychologists and psychiatrists who attribute meaning to psychological responses to colour cannot agree on which colour means which and in a cross cultural context colour has entirely different meanings depending on your cultural background (Jacobs et al. 1991; Aslam 2006). The temptation for an innovative market researcher would be to take a validated methodology based on interpretation of colour and apply it in a different culture.

Similarly, after more than 50 years of usage, there is still debate about the interpretation of Thematic Apperception Test (TAT) pictures (Hibbard 2003; Ackerman et al. 2001; Lilienfeld, Wood, and Garb 2000; Cramer 1999). If the experts can’t agree on what it all means, how could the ‘face’ who is being asked to review the pictures? We note that many research companies have proprietary projection ‘kits’ that are applied across populations, situations and contexts. If the experts can’t agree on TAT after 50 years, how valid might some instances of proprietary projections be?

#### Example 3: ‘Construct’ development methodologies and face validity

Some methods have developed entire mythologies that have grown around them without recent peer review and methodological challenge. A case in point is Geert Hofstede’s now famous cross cultural dimensions (cf. Hofstede and Bond 1988; Hofstede 1994; Hofstede 2001). Hofstede analysed a large (secondary) data base of employee values scores collected by IBM between 1967 and 1973. The data was from more than 70

This paper is an early draft of a paper which eventually appears as:  
Brennan, L., Camm, J., & Tanas, J. (2007). Validity in market research practice: ‘new’ is not always ‘improved’. *Der Markt*, 46(1-2), 3-12.

countries, from which Hofstede first used the 40 largest only, and afterwards extended the analysis to 50 countries and 3 regions. Since 2001, scores are listed for 74 countries and regions, partly based on replications and extensions of the IBM study on different international populations. As of mid-June 2006, the search term ‘Hofstede’ produced over 45,000 citations in *scholar.google.com* – while not the most authoritative source, a clear indication of the volume of researchers who are using Hofstede’s concepts. One of the dimensions identified by Hofstede and Bond (1988) from an original study of Chinese Cultural Values is something that is now called ‘long term orientation’ (Merkin 2004). The ‘constructs’ in this dimension are ‘measured’ by the items shown in Table 1:

**Table 1: Measures of long term orientation**

DIMENSION LABEL (CONSTRUCT)		
	<b>Long term orientation</b>	<b>Short term orientation</b>
<b>Scale Items</b>	Persistence (perseverance)	Personal steadiness and stability
	Ordering of relationships by status	Protecting your face
	Thrift	Respect for tradition
	Having a sense of shame	Reciprocation of greetings, favours and gifts

You can see that most of the words used in these ‘measures’ have very little to do with 1) time or 2) long term outlook or 3) orientation (taking a relative position). They may have more construct validity if compared with the original Chinese Values Scale as put forward by Bond (but we assert that they do not). However, the use of these items to ‘measure’ whether or not an individual approaches their life from a long term, future orientated, perspective would appear to be inconsistent with dictionary definitions of the words. Notwithstanding the ambiguity of meaning, the measure continues to be used apace even after Fang (2003) made it very clear that these concepts were being fundamentally misused. Any basic face validity test would have identified the problem: that after borrowing the methodology from the original Chinese it now makes little sense. Put another way, the original items may have been valid, but they have been applied uncritically throughout the world in new populations, situations and contexts. No doubt, the use of passive students or disinterested employees in their corridors resulted in uncritical ‘evaluations’ of these items.

A researcher who simply uses an existing scale without challenging the appropriateness of the use within the context and within the population of interest (the face) might very easily use a test score to infer that there are substantive differences between groups. These differences might not actually exist were the test items revalidated. Churchill (Churchill 1979) suggested that validation of test items could be done via relatively informal sources (colleagues, convenience samples, etc). However, we argue that by doing this, researchers take the risk that the construct development is undertaken in a population that is not able to discern inherent semantic issues such as that described above.

## **WHY IS THE PROBLEM OF VALIDITY SUCH AN ISSUE IN MARKET RESEARCH?**

We have provided three examples of how invalid methodology can infiltrate market research. This leads to the question of why validity needs to be an important issue in market research, more than when compared to other sciences or engineering. Validity arises as a problem in market research because there is often a lack of linkage back to the 'real world.' In order to explain this, consider a comparison with the science of bridge building. In bridge building, methodology and design principles are constantly validated by virtue of bridges having to pass the test of not falling down. Think of an innovative young engineer devising an imaginative new approach for calculating resonant frequencies in a bridge. A wrong calculation can result in vibrations being amplified in the bridge structure, which is undesirable and may lead to the bridge falling down. In the real world, the brutal check on the validity of any new approach is whether or not the bridge vibrates (and therefore whether it remains standing under pressure). But in market research, we can propose a solution to a problem, have it implemented, observe some outcomes and then move on without having to live with the consequences. It is very difficult to determine whether that hoped-for increase in satisfaction and subsequent sales did not occur because the methodology findings were invalid, or whether uncontrolled factors in the marketplace intervened.

Another example, from everyday life, is the 'theorising' that occurs for the latest 'whodunit' in the news. When the details are sketchy, unfettered imagination amongst armchair criminologists is free to roam far and wide. However, as time progresses and more 'facts' are brought to bear there is a process of pruning the wild thicket of theories that flourished when the news first breaks. This sort of theorising about social events is not so different to the exploratory or discovery stage in science. Evidence accumulates and ideas have to fit the ever increasing body of evidence. Furthermore, in many sciences we have the luxury of not just accumulating new evidence, but of being able to go back to the laboratory and design specific experiments that can differentiate between competing theories (or hypotheses).

The point we wish to make is that whether it be bridge building, theorising on whodunits or conscious laboratory experiments, there is a process of continuously confronting the methodology with the evidence of the real world. But for market research, it is often the case of delivering results and moving on. This fact places a huge onus on the researcher to ensure that methodology is validated beforehand. It explains why there can be an accumulation of methodologies of doubtful usefulness.

These days we are fortunately not too often subjected to unfettered innovation in bridge building. And while armchair theorising about titillating social events probably does not do much harm, the same is not true in market research. We have an obligation to ensure that new methodologies and approaches are valid and reliable. Further, while reliability can be easily established through the application of statistical techniques, we argue that there is simply no point in having a very reliable but invalid instrument. Notwithstanding this, if a measure is not even reliable at a 'reasonable' level (Cronbach 1951), there is no point in using it even if all aspects of validity have been accounted for. We suggest that

This paper is an early draft of a paper which eventually appears as:  
Brennan, L., Camm, J., & Tanas, J. (2007). Validity in market research practice: 'new' is not always 'improved'. *Der Markt*, 46(1-2), 3-12.

the foundation for 'good' research is an assessment of validity at the formative stages of any project.

This paper is an early draft of a paper which eventually appears as:  
Brennan, L., Camm, J., & Tanas, J. (2007). Validity in market research practice: 'new' is not always 'improved'. *Der Markt*,  
46(1-2), 3-12.

## APPENDIX A: TYPOLOGY OF VALIDITY IN MARKETING RESEARCH

*Validity comes before reliability – something should be at least valid before it can be considered reliable*

**Note: This list is meant to be illustrative not exhaustive**

Term	Description	Authors
Validity	The measure measures what it purports to measure	(Carmines and Woods 2005)
<b>Formative validity</b>		
	<b>This is the term we propose which encompasses the ‘measures’ of validity which come before the collection of data – therefore is an integral part of project design</b>	<b>Brennan and Camm (this paper)</b>
Content validity	The extent to which an instrument measures those qualities that it purports to measure – often used synonymously with face validity. However content validity implies expert acceptance of the measure(s).	(Sireci 1998) Thorndike and Hagen 1971 (Cronbach and Meehl 1955) (Nunnally, 1967) (Diamantopoulos 2005)
Face validity	Exists when the measure ‘looks as if’ is should indicate a particular variable. This is usually when others agree so it is also called consensus validity. This is often used synonymously with content validity. However, face validity does not imply expert acceptance of the measure’s validity.	(Turner 1979) (Heeler and Ray 1972) (Diamantopoulos 2005)

<b>Prognostic validity</b>		
	<b>This is the term we propose which encompasses assessment of the predictive, descriptive and reflective measures of validity AFTER the data has been collected. It will incorporate an understanding of design, methods, data collection and sources of error. It may also include some tests of reliability where these tests contribute to an understanding of the prognostic value of the data.</b>	<b>Brennan and Camm (this paper)</b>
Construct validity	The extent to which a test may be said to measure the theoretical idea (construct). (Includes criterion, convergent, concurrent and discriminant validity), can only be inferred, as a construct is an abstraction	(Carmines and Woods 2005) (Churchill 1979) (Redmond and Griffith 2003) (Heeler and Ray 1972) (Peter 1981)

This paper is an early draft of a paper which eventually appears as:  
 Brennan, L., Camm, J., & Tanas, J. (2007). Validity in market research practice: ‘new’ is not always ‘improved’. *Der Markt*, 46(1-2), 3-12.

Concurrent validity	The extent to which particular measures predict other criterion measures concurrently – that is there is concurrent variation in the outcomes (predictive validity)	(Heeler and Ray 1972) (Cronbach and Meehl 1955) (Nunnally, 1967)
Convergent validity	Synonymous with concurrent and predictive validity - the extent to which the measure correlates or converges with similar measures of the same variable. That is, the variables are measuring the same construct.	(Heeler and Ray 1972) (Diamantopoulos 2005) (Rossiter 2002) (Reichardt and Coleman 1995)
Discriminant validity	Antonymous with convergent validity it is the extent to which a measure is different to other measures. A valid construct is different from other constructs - also called divergent validity.	(Heeler and Ray 1972) (Rossiter 2002) (Diamantopoulos 2005) (Reichardt and Coleman 1995)
Criterion validity	The measure of how well one variable (or set of variables) predicts an outcome based on information from other variables (Includes concurrent, predictive and diagnostic validity)	(Nunnally, 1967) (Redmond and Griffith 2003) (Diamantopoulos 2005)
Nomological validity (lawlike)	The degree to which predictions from a formal theoretical network containing the concept under consideration are confirmed. Thus, there are formal hypotheses derived from theory –this probably means peer reviewed	(Venkatraman 1989) (Diamantopoulos 2005) (Peter 1981)
Predictive validity	The extent to which particular measures predict other criterion measures - often measured at the same time as concurrent validity. That is you can use one measure to predict the outcomes of another. Usually implies pre-test and post-test methodologies.	(Heeler and Ray 1972) (Nunnally, 1967) (Cronbach and Meehl 1955)
Diagnostic validity	Used mainly in the health sciences. Does the test outcome consistently relate to the diagnosis of the client/patient condition?	(Redmond and Griffith 2003)
Substantive validity	A subset of construct validity, it is the extent to which the scale items have a theoretical linkage to the construct.	(Schwab 1980) (Garver and Mentzer 1999)
Internal validity	When any differences between outcomes are attributed solely to the effect under investigation. The role of theory in this type of validity is to differentiate as something different from other constructs. Includes reliability, convergent validity and	(Redmond and Griffith 2003) (Peter 1981)

	discriminant validity.	
External validity	The extent to which results can be generalised to populations, situations or conditions. Includes nomological, content and face validity	(Redmond and Griffith 2003) (Peter 1981)

## REFERENCES

- Ackerman, Steven J., Mark J. Hilsenroth, Amanda J. Clemence, Robin Weatherill, and J. Christopher Fowler. 2001. Convergent Validity of Rorschach and TAT Scales of Object Relations. *Journal of Personality Assessment* 77 (2):295-306.
- Aslam, Mubeen. 2006. Are you selling the right colour? A cross cultural review of colour as a marketing cue. *Journal of Marketing Communications* 12 (1):15-30.
- Badri, Masood A., Mohamed Abdulla, and Abdelwahab Al-Madani. 2005. Information technology center service quality: Assessment and application of SERVQUAL. *International Journal of Quality & Reliability Management* 22 (8):819-848.
- Barker, Andy. 2002. Nature, nurture and inertia: Where do qualitative researchers get their ideas from? *Qualitative Market Research* 5 (2):152-153.
- Basil, Michael D., William J. Brown, and Mihai C. Bocarnea. 2002. Differences in Univariate Values Versus Multivariate Relationships. Findings from a Study of Diana, Princess of Wales. *Human Communication Research* 28 (4):501-514.
- Blankenship, A. B. 1961. Creativity in Consumer Research. *Journal of Marketing (pre-1986)* 25 (6):34-38.
- Campbell, D T, and Donald. W Fiske. 1959. Convergent and discriminant validation by the multitrait-multimethod matrix. *Psychological Bulletin* 56:81-105.
- Campbell, DT. 1960. Recommendations for APA test standards regarding construct, trait, and discriminant validity. *American Psychologist* 15 (August):546-553.
- Carman, James M. 1990. Consumer perceptions of service quality: an assessment of the SERVQUAL dimensions. *Journal of Retailing* 66 (1 Spring):33-35.
- Carmine, Edward. G, and James A Woods. 2005. Validity Assessment. *Encyclopedia of Social Measurement* 3:933-937.
- Celsi, Richard, Simeon Chow, Jerry Olson, and Beth Walker. 1992. The Construct Validity of Intrinsic Sources of Personal Relevance: An intra-Individual Source of Felt Involvement. *Journal of Business Research* 25 (2 September):165-185.
- Churchill, G A Jr, Peter, J P. 1984. Research design effects on the reliability of rating scales: a meta analysis. *Journal of Marketing Research* XXI (November):360-375.
- Churchill, Gilbert A Jr. 1979. A paradigm for developing better measures of marketing constructs. *Journal of Marketing Research* 16 (February):64-73.
- Cramer, Phebe. 1999. Future Directions for the Thematic Apperception Test. *Journal of Personality Assessment* 72 (1):74-92.
- Cronbach, L J. 1951. Coefficient Alpha and the internal structure of tests. *Psychometrika* 16 (3):297-333.
- Cronbach, Lee J, and Paul E Meehl. 1955. Construct validity in psychological tests. *Psychological Bulletin* (52):281-302.
- Deshpandé, Rohit. 1983. "Paradigms Lost": On theory and method in research in marketing. *Journal of Marketing* 47 (Fall):101-110.
- Diamantopoulos, Adamantios. 2005. The C-OAR-SE Procedure for scale development in marketing: A comment. *International Journal of Research in Marketing* 22:1-9.

This paper is an early draft of a paper which eventually appears as:

Brennan, L., Camm, J., & Tanas, J. (2007). Validity in market research practice: 'new' is not always 'improved'. *Der Markt*, 46(1-2), 3-12.



- Donnelly, Mike, Neil J. Kerr, Russell Rimmer, and Edward M. Shiu. 2006. Assessing the quality of police services using SERVQUAL. *Policing: An International Journal of Police Strategies & Management* 29 (1):92-105.
- Fang, Tony. 2003. A Critique of Hofstede's Fifth National Culture Dimension. *International Journal of Cross Cultural Management: CCM* 3 (3):347-368.
- Garver, Michael S., and John T. Mentzer. 1999. Logistics research methods: employing structural equation modeling to test for construct validity. *Journal of Business Logistics* 20 (1):33-57.
- Gerbing, David W, and James C Anderson. 1988. An updated paradigm for scale development incorporating unidimensionality and its assessment. *Journal of Marketing Research* XXV:361-370.
- Gounaris, Spiros. 2005. Measuring service quality in b2b services: an evaluation of the SERVQUAL scale vis-à-vis the INDSERV. *Journal of Services Marketing* 19 (6):421-435.
- Heeler, Roger M, and Michael L Ray. 1972. Measure Validation in Marketing. *Journal of Marketing Research* IX (November):361-370.
- Hibbard, Stephen. 2003. A Critique of Lilienfeld et al.'s (2000) "The Scientific Status of Projective Techniques". *Journal of Personality Assessment* 80 (3):260-271.
- Hofstede, G H. 2001. *Culture's consequences*. 2nd ed. Thousand Oaks, CA: Sage.
- Hofstede, Geert. 1994. Business cultures. *The Unesco Courier* Apr (4):12.
- Hofstede, Geert, and Michael Bond. 1988. The Confucius Connection: From Cultural Roots to Economic Growth. *Organizational Dynamics* 16 (4):4-21.
- Hult, G. Tomas M., Jr David J. Ketchen, and Stanley F. Slater. 2005. Market orientation and performance: an integration of disparate approaches. *Strategic Management Journal* 26 (12):1173-1181.
- Jacobs, Laurence, Charles Keown, Reginald Worthley, and Kyung-II Ghymn. 1991. Cross-Cultural Colour Comparisons: Global Marketers Beware! *International Marketing Review* 8 (3):21-30.
- Lilienfeld, Scott O., James M. Wood, and Howard N. Garb. 2000. The Scientific Status of Projective Techniques. *Psychological Science in the Public Interest* 1 (2):27-66.
- McAnena, Fiona. 2004. Taking insight from dull data. *Marketing (UK)*:33-.
- McGartland Rubio, Doris, and Kempf-Leonard Kimberly. 2005. Content Validity. In *Encyclopedia of Social Measurement*. New York: Elsevier.
- Merkin, Rebecca S. 2004. Cultural Long-Term Orientation and Facework Strategies. *Atlantic Journal of Communication* 12 (3):163-176.
- Murray, Jeff B., Deborah J. Evers, and Swinder Janda. 1995. Marketing, theory borrowing, and critical reflection. *Journal of Macromarketing* 15 (2):92-.
- Nunnally, Jum C. 1967. *Psychometric Theory*. New York: McGraw Hill.
- O'Connor, Thomas. R. 2005. Types of Scales and Indexes. *Encyclopedia of Social Measurement* 3.
- Parasuraman, A., Valarie A. Zeithaml, and Leonard L. Berry. 1985. A Conceptual Model of Service Quality and Its Implications for Future Research. *Journal of Marketing* 49 (4 Fall):41-45.
- Peter, J Paul. 1981. Construct validity. *Journal of Marketing Research* XVIII (May):133-145.

This paper is an early draft of a paper which eventually appears as:

Brennan, L., Camm, J., & Tanas, J. (2007). Validity in market research practice: 'new' is not always 'improved'. *Der Markt*, 46(1-2), 3-12.

- Redmond, Elizabeth C, and Christopher J Griffith. 2003. A comparison and evaluation of research methods used in consumer food safety studies. *International Journal of Consumer Studies* 27 (1 January):17-33.
- Reichardt, Charles S, and S C Coleman. 1995. The criteria for convergent and discriminant validity in a multitrait-multimethod matrix. *Multivariate Behavioral Research* 30 (4):513-538.
- Rossiter, John R. 2002. The C-OAR-SE procedure for scale development in marketing. *International Journal of Research in Marketing* 19 (4):305-336.
- Salzberger, Thomas. 2000. An alternative way of establishing measurement in marketing research- its implications for scale development and validity. Paper read at ANZMAC 2000, at Griffith University, Australia.
- Schwab, Donald P. 1980. Construct validity in organizational behavior. In *Research in Organizational Behaviour*, edited by B. Staw and L. L. Cummings. Greenwich, Connecticut: JAI Press: 3-43.
- Sireci, Stephen G. 1998. The Construct of Content Validity. *Social Indicators Research* 45 (1 - 3):83-117.
- Smith, Gregory T. 2005. On Construct Validity: Issues of Method and Measurement. *Psychological Assessment* 17 (4):396-408.
- Temporal, Paul. 2006. *Asia's Star Brands*. Singapore: John Wiley and Sons.
- Turner, Stephen P. 1979. The concept of face validity. *Quality and Quantity* 13 (1):85-90.
- Vanderkaay, Sharon. 2005. The Business of Selling Knowledge: Part I. *Canadian Architect* 50 (2):27.
- Venkatraman, N. 1989. Strategic orientation of business enterprises: The construct, dimensionality, and measurement. *Management Science* 35 (8):942-962.