PROGRESS IN MARKETING KNOWLEDGE

Malcolm Wright, Ehrenberg-Bass Institute, University of South Australia

Zane Kearns, Massey University

Abstract

By drawing on an empiricist tradition, and on the literature discussing the philosophical problems of marketing knowledge, we seek to develop practical guidelines for developing empirical generalisations. In particular we suggest three criteria for the development of marketing knowledge: ensuring falsifiability and theoretical competition; overcoming uncertainty through replication; and using extension to develop generalisations and identify boundary conditions. By way of demonstration, these criteria are applied to assessments of the Dirichlet model, the Servqual instrument, and market share modelling. We conclude with renewed plea that more of the academic research effort in marketing be devoted to replicating and extending existing results, and determining the conditions under which existing theories do, *and do not*, hold.

Correspondence should be addressed to Malcolm Wright at:

Ehrenberg-Bass Institute, University of South Australia

GPO Box 2471, Adelaide, South Australia 5001

AUSTRALIA

Email Malcolm.Wright@marketingscience.info

The authors would like to thank two anonymous JEMS referees for helpful comments on an earlier version of this paper.

INTRODUCTION

In recent years, considerable attention has been paid to the development of empirical generalisations in marketing. Leone and Schultz (1980) were amongst the first to attempt to catalogue marketing generalisations, although they found very few. Since then there has been renewed interest in this area with a number of criticisms of the current state of marketing knowledge, and suggestions for improved methodological practice (eg. Armstrong 1991; Armstrong and Schultz 1993; Ehrenberg and Bound 1993; Hubbard and Armstrong 1994; Ehrenberg 1995; Barwise 1995; Bass 1995; Armstrong, Brodie, and Parsons 1997). This work recently culminated in a workshop on empirical generalisations at Wharton, and the associated special issue of Marketing Science (1995).

A separate stream of work has focused on the philosophical problems of marketing knowledge (eg. Anderson 1983, 1986, 1988a, 1988b; Hunt 1984, 1990, 1992, 1993a, 1993b; Peter 1992; Zinkhan and Hirscheim 1992). Although this work and the work on empirical generalisations both address the problem of knowledge in marketing, the two bodies of literature have not been strongly linked.

We believe that there is a need to integrate these various approaches. In particular, we hope to use the philosophical debate on marketing knowledge as a source of practical guidelines for (i) undertaking research to develop empirical generalisations, and (ii) assessing the progress made by current research programmes towards such generalisations.

Zaltman (1991) has already discussed the issues associated with assessing progress towards meeting the Marketing Science Institute's research priorities, but his position gave priority to the sociological processes involved in the development and dissemination of knowledge. This provides few practical guidelines for ensuring progress in developing empirical generalisations.

This paper presents a contrasting position; one which focuses on the logical rather than the social structure of marketing knowledge, and one which is based on the use of intersubjectively verifiable observations (ie. observations which can be verified by other people) to develop and improve empirical generalisations. Based on this approach, we

propose a set of criteria for making progress in developing marketing knowledge, and outline philosophical justifications for these criteria. By way of demonstration, we then discuss several high profile areas in the marketing literature.

EPISTEMOLOGICAL BACKGROUND

In attempting to develop objective knowledge about the world empiricists face two major problems. First, any such knowledge is based on fallible observations. Second, it is logically impossible to prove any universal statement or theory to be true.

We accept that empirical knowledge is fallible (see also Hunt, 1990, 1993a). Zaltman (1991) also discusses the uncertainty inherent in the quest for knowledge, as do other authors, such as Peters (1992), Anderson (1983), and Hirschman (1986). While these authors have pointed out that the problems of observation make the naive realists' idea of objective "Truth" untenable, Hunt has demonstrated that this does not require the adoption of a relativist epistemology (Hunt 1984, 1990, 1992, 1993a, 1993b). If we act with prudence we can still obtain reasonably objective knowledge about the world through observation. The fact that this knowledge is fallible puts marketing in a position no different from that enjoyed by physics or biology.

The second problem, the impossibility of conclusive proofs of universal statements or theories, was recognised and discussed by Popper (1935). As an alternative to trying to prove theories true, Popper proposed instead that our theories must be falsifiable (ie. there must be some conceivable observation which could contradict them), and that we should subject them, and *competing theories*, to the strongest tests possible to determine which provides the best predictions and has the least serious falsifying instances. His ideas are reflected in other modern solutions to this problem. For example, Chalmers suggests that we address the fallibility of our knowledge by "pitching [it] against the world in the most demanding way possible given existing practical techniques" (Chalmers 1990, p. 7). Armstrong, Brodie, and Parsons (1997) suggest a practical way of ensuring stronger tests of theories by recommending that we always test multiple competing hypotheses, and prefer the hypothesis that, over time, performs the best for given conditions.

Some researchers would object to the requirement for such direct tests of theories, claiming that some types of knowledge do not directly yield testable predictions. For example, managers may wish to commission interpretive studies of consumer behaviour, which although they do not make specific predictions, are said to provide insights which, in turn, improve the quality of managerial decision making.

Modern consumer behaviour texts often take this approach - that *understanding* leads to better decisions - although they typically do not describe this as interpretive research. For example, one popular consumer behaviour text states that '"[this book] is based on the belief that a knowledge of the factors that influence consumer behaviour can, with practice be used to develop sound marketing strategy" (Hawkins, Best, and Coney 1992, iii). Another says that "Managers who want to satisfy consumers need an indepth understanding of those consumers. *Understanding Consumer Behaviour* identifies the essential elements of consumer behaviour and provides the knowledge and skills to analyze the reasons for consumers' behaviour. The text also shows how understanding consumers can be used to develop effective marketing strategies." (Peter and Olson 1994, v).

However, evidence is still required to support the claim that such interpretive knowledge improves decision making, or alternatively to determine which type of *understanding* produces the best results. If there is no such evidence then, as Hunt (1993b) points out, it is hard to see how such knowledge could be distinguished from astrology, palmistry, or crystal therapy.

Another possible objection is that a focus on falsifiability and the rejection of relativism emphasises the logical process of justification at the expense of the creative process of theory development (Hunt, 1993b, outlines the distinction between discovery and justification in marketing). For example, revolutionary new theories can be generated by inspiration, or by subjective processes of interpretation, neither of which rely on empirical research; Ehrenberg (1995) described this as the "Theoretical-Empirical" approach (although he did not support its use). We contend that while justification is amenable to a logical description, the intuitive and creative elements of discovery defy formalisation, and different methods varying unpredictably in their effectiveness between different individuals and research programs.

Nevertheless, we acknowledge that discovery and creativity are important in ensuring scientific progress, particularly in developing falsifiable theories, and in thinking of ways to extend and modify theories. However, creativity still needs to be combined with the logical aspects of justification; in particular, the value of creative outputs can only be demonstrated by their falsifiability, by successfully testing their predictions, replicating these tests, and extending them to a variety of different conditions. As Edison said "Genius is one percent inspiration and ninety-nine percent perspiration".

Furthermore, creative theory development often depends on empirical results for its genesis (the "Empirical-Theoretical" approach of Ehrenberg, 1995). Consequently we see empirical results and creative development as inextricably intertwined. Therefore, while creativity has an unquestionably important role in progress, we do not attempt to prescribe a method for being creative. In any event, we do not believe that there is a shortage of creativity in marketing theory.

Given this background, we claim that progress is achieved by developing falsifiable theories, overcoming uncertainty through replication, attempting to extend our theories to new situations, and by identifying areas in which the theory or technique systematically fails. These failures are the primary means of identifying areas where *existing* theories eventually need to be modified, improved, or replaced.

In the next section, we outline these procedures in some detail.

CRITERIA FOR PROGRESS IN MARKETING KNOWLEDGE

Falsifiability and Theoretical Competition

The first criterion for any empirical theory is that it be falsifiable; that is, it must be possible to conceive of an observation which would contradict the theory. Falsifiability was claimed by Popper (1935) to be the key point of demarcation between science and non science.

In establishing this criterion we are not restricting ourselves to the original falsificationist programme, rather we are establishing *falsifiability* as the first hurdle which must be passed by any theory purporting to have empirical content.

For a theory to be falsifiable it must be capable of making testable predictions. Under the original falsificationist programme, these predictions are to be compared with observations, and with the predictions of other theories (or "multiple competing hypotheses") to determine how well the theory performs. The best performing theories: (i) have less serious falsifying instances than competing theories; (ii) make a wider variety of predictions; and (iii) successfully predict some outcomes which are at odds with the predictions of competing theories.

The idea of multiple competing hypotheses (Armstrong et al 1997) is particularly important as it helps to overcome the "confirmation bias" inherent when testing a single theory - the tendency to keep working and reworking the analysis until the preferred result is obtained. As Greenwald, Pratkanis, Leippe, and Baumgardner (1986) have pointed out, even falsificationist research may be subject to this bias. Another method of overcoming confirmation bias, suggested by Greenwald et al (1986) is to also give some emphasis to identifying the *conditions* under which a theory performs well. We consider this in more detail in a subsequent section.

For a theory to be falsifiable *in principle,* we must be able to conceive of some evidence that would refute it. If we cannot do this, Popper would say that the theory is unscientific, and provides no basis for progress.

For example, Newtonian physics was falsifiable, and indeed was falsified by Eddington's observation of light displacement around the sun during a solar eclipse (Popper 1935). This falsification established a boundary to the application of Newtonian physics, and highlighted the requirement for new or improved theories beyond this boundary. By contrast astrology is not falsifiable, because we cannot conceive of any evidence which would be accepted by astrologers as a refutation of their "predictions". In general, an unfalsifiable theory can never be contradicted, so there is no possibility for the type of empirical failure which forms the basis of improvements to empirical theories.

Theories that appear to be falsifiable can be rendered unfalsifiable *in practice* by the actions of their supporters. For example, Popper (1935) noted that Marxist historical theory was originally falsifiable, but was rendered unfalsifiable by the introduction of ad hoc defences against disconfirming evidence. Modifications of a theory can be legitimate if they introduce new falsifiable claims, but the ad hoc modifications of the Marxists did not meet this requirement.

The criterion of falsifiability requires that a theory be falsifiable *in principle*. The question of whether a theory is falsifiable *in practice* can only be addressed by considering whether its proponents subject the theory to rigorous empirical tests, and how they respond to the results of those tests.

Some theories in marketing have been criticised as unfalsifiable. In particular Tuck (1976) and Ehrenberg (1988) criticised comprehensive models of consumer behaviour, such as those of Howard and Sheth (1969) and Engel, Blackwell and Kollat (1978), for being consistent with any possible observation, and therefore being unfalsifiable (although these textbooks are now somewhat dated, similar models are still implicit in many of the current major consumer behaviour texts). It may be that comprehensive models of consumer behaviour are not intended to yield falsifiable predictions directly. Rather, they may be interpretive models that are intended to be applied to improve managerial decision making. However, the claim that the application of these models improves decision making is a falsifiable claim, which should still be empirically assessed.

Failures of interpretivist theories are sometimes attributed to the implementation of the theory rather than to the theory itself. While this will occasionally be true, the use of this argument places a burden of proof on the advocates of the theory to demonstrate successful applications. This burden of proof is required to avoid unfalsifiability *in practice*.

Overcoming Uncertainty Through Replication

A major criticism of falsification and other empirical approaches is that the observations on which they rely (including falsifying observations) are subjective and uncertain, and

so cannot be reliably used to either falsify, justify, or even objectively compare theories (Chalmers 1982, Anderson 1983).

This subjectivity and uncertainty arises because of factors such as:

sampling error;

the dependence on fallible "instrument" theories underpinning the data collection procedures;

the possibility of some confounding influence such as an unrecorded marketing activity or fluctuations in the business environment;

the possibilities of psychological subjectivity, bias, error, or plain dishonesty from the observing researcher.

Hunt addresses several of these points in his arguments for truth and objectivity in marketing (1990, 1993a), but fails to recognise the elegance of replication as a solution to these problems. The importance of replication as a check on observations has been stressed recently by Lindsay and Ehrenberg (1993) and Hubbard and Armstrong (1994), but these authors failed to clearly identify that replication also addresses the philosophical problems of observation. Replication helps to overcome the philosophical problems of observation through the completion of studies with:

further samples to reduce the problem of sampling error (a similar function is performed by meta-analysis of existing studies);

different data collection methods to make our observations somewhat independent of the bias arising from any one particular data collection method (a similar point is made by Chalmers, 1990);

different times or places to reduce the effects of confounding influences;

additional observers to reduce the influence of the subjectivity, bias, error, or dishonesty of a single observer on the results (what Hunt calls intersubjective verifiability; 1991).

Hubbard and Armstrong (1994) point out that marketing has a preponderance of unreplicated studies. The findings of such studies cannot be viewed with confidence until they have been reproduced by different researchers using different data collection methods under different conditions. Replication provides an effective method for ensuring that our research results are robust, but it also plays a vital role in the development of empirical generalisations by testing the whether the initial findings still occur under different conditions. This is discussed further in the next section.

Using Extensions to Develop Generalisations and Identify Boundary Conditions

The principles outlined so far ensure that theories are structurally sound (falsifiable), the best available (of the multiple competing hypotheses), and have been subject to empirical tests (with results confirmed to be reliable by replication). These conditions are necessary for meaningful progress, but they are not entirely sufficient. After all, the aim of science according to Chalmers is "to establish highly general laws and theories applicable to the real world ... Further, it is understood that the generality and degree of applicability of laws and theories is subject to continual improvement" (1990, p. 7). How then can we achieve high generalisability and continual improvement?

Lindsay and Ehrenberg (1993) suggest that we follow close replications (designed to ensure the reliability of our observations) with differentiated replications. Differentiated replications are designed to discover whether the theory generalises to different conditions. This can include changes both to variables that are part of the theory, and to variables that have been excluded from the theory for the sake of parsimony, but might plausibly have some effect.

Differentiated replications also help researchers to improve theories by identifying conditions under which the theory systematically fails. This provides direction for further developments and improvements and may even occasionally lead to a revolutionary change or "paradigm shift" in the area under investigation. Testing the limits of our theories is usually a prerequisite for progress: unless we can identify areas where a theory systematically fails, we have no reason to suggest the theory is inadequate, and no reason to attempt to make further progress.

Another view of this point is proposed by Greenwald et al (1986). As noted earlier, they point out that theory testing often involves a confirmation bias - a tendency to keep working and reworking the analysis until the preferred result is obtained. One way to avoid this is to actively seek to identify boundary conditions, or conditions under which the phenomena of interest no longer appears, and so they present the identification of boundary conditions as a key component of progress in research. We agree.

An immediate differentiated replication can also perform the functions of a close replication. The difficulty is that if the differentiated replication does not confirm the tentative result, or if the evidence is mixed (some confirmatory, some not) it is hard to determine whether this is due to the unreliability of the original observations, or to the failure of the theory to generalise to the new conditions. This reservation stated, a program that makes immediate use of differentiated replications, while inherently more risky, can progress more quickly.

The procedures we have outlined in the last three sections are not a complete description of what is needed to achieve progress in marketing knowledge. As we noted earlier, creativity also has a key role to play, for example, in choosing what areas to examine and designing the actual research. We have not, however, emphasised this creative process for two reasons: (i) there do not appear to be any reliable guidelines in this area; and (ii) there is no shortage of creativity apparent in marketing research.

APPLICATIONS

To demonstrate the application of our criteria for progress in marketing knowledge, we now briefly discuss three illustrative areas of the marketing literature.

Negative Binomial Distribution / Dirichlet Model

One of the most successful empirical generalisations in marketing science is the Negative Binomial Distribution (NBD) / Dirichlet model. This has been recently reviewed in the Marketing Science Special Issue on Empirical Generalisations in Marketing (Uncles, Ehrenberg, and Hammond 1995).

The predictions of the model are clearly falsifiable *in principle*. We can well imagine (for example) regularly observing the presence of small brands with repeat purchase rates much higher than their larger competitors, a phenomenon excluded by the Dirichlet model. As it happens, we very seldom observe this phenomenon. Indeed, the excellent fit of the model has been replicated and extended to a number of different product fields and countries by a variety of different researchers.

The model has also been modified and improved over time. The most obvious progress in this area has come from the shift from the NBD to the NBD-Dirichlet formulations of the theory, although the original development of the NBD also drew together a variety of earlier unconnected observations. While there has already been a great deal of progress in this area, research continues, both to investigate the application of the model to new situations (recent examples include: store choice, Uncles and Ehrenberg 1990; and pharmaceuticals, Stern and Ehrenberg 1995), and to investigate whether the mathematical formulation of the model can be further improved to address existing empirical discrepancies or boundary conditions (some of these issues are discussed in Barnard, Ehrenberg, Hammond, and Uncles 1994). Progress in this area has been considerable, and is ongoing (see also Ehrenberg and Uncles 1997).

Servqual

On the other hand the debate about "Servqual", an instrument for measuring service quality, has shown only limited progress (see Cronin and Taylor, 1992, 1994, and

Parasuraman, Zeithaml and Berry, 1994 for a discussion of Servqual versus Servperf, and Buttle 1996 for a summary of other aspects of the Servqual debate). The original authors have been modifying their views and their instrument, and so might appear to be making progress. Before we can evaluate progress, however, we need to consider what falsifiable theory is being proposed. Three possible candidates can be identified.

1. Higher Servqual ratings lead to higher profits (or share, or some other measure of "success").

2. The use of the Servqual instrument gives results which enable managers to improve profits (or share, or some other measure of "success").

3. Service quality is a multidimensional construct (with five important dimensions) which can be measured by a specific procedure.

Formulations (1) and (2) have a clear dependent variable. However the Servqual debate has not specifically investigated the effects of service quality on profitability (or any other clear dependent variable). Until such tests are attempted we will not have any idea of whether propositions (1) and (2) hold at all, let alone the circumstances under which they hold, or how this version of Servqual theory can be improved. Consequently, progress on propositions (1) and (2) appears to be almost non-existent.

Formulation (3) has been extensively examined, so some progress might be thought to have occurred in this area. However it has also been subject to a large number of theoretical and empirical criticisms, (again, see Buttle 1996 for a summary). A particularly serious criticism of the Servqual scale is that the dimensions and components vary markedly from situation to situation. This may be because they are highly situation specific, or it may be because the procedure is not reliable. In any event, no generalisable patterns seem to emerge from this research. The result is a proliferation of suggestions for modifications, but little assessment of these proposed modifications through replication in different circumstances. Progress thus far in this area has been limited to falsifying the original formulation of formulation (3), and suggesting a number of largely untested replacements. Meaningful progress requires that these replacements be subjected to dispassionate enquiry to determine which performs the best, and whether this performance generalises to different circumstances.

More promising lines of research have recently been presented by Anderson, Fornell, and Lehmann (1994) who report a test of the links between other measures of customer satisfaction, and market share and profitability, and Rust, Zahorik, and Keiningham, (1995) who propose an approach for identifying the financial return on service quality. These authors do not use the Servqual scale, but they do establish a clear falsifiable proposition - that changing service quality has a benefit (or cost) which can be measured and predicted. Further progress will occur as this claim is tested in a variety of different industries and countries, as the research is replicated by different researchers, and the conditions under which this claim does (and does not) hold are identified.

Market Share Modelling

Market share modelling is a pervasive activity in both academia and industry, especially since the advent of data collection from supermarket checkout scanners. We do not propose to review the voluminous literature here, but will offer a few general observations.

This area involves the development of a large number of different types of models, applied to a wide variety of data sets, and so has key ingredients available for considerable progress. There are also two key problems.

The first problem is the criteria by which a particular model formulation can be evaluated. The typical approach is to compare a model with a naive or competing model to see which performs the best; (although a naive model should be a relatively weak competitor, it often performs, comparatively, surprisingly well). The models may be assessed either on their fit to the data they were estimated from, or on their fit to a holdout sample, using measures such as R^2, Absolute Average Error, Thiel's U, or the Bayesian Inequality Criterion. This is similar to Armstrong et al's (1997) recommendation of testing multiple competing hypotheses.

However, measuring statistical fit to a single set of historical data overlooks other issues such as parsimony, explanation, and prediction. In particular, the application of market share modelling is to *predict* the outcomes of marketing mix decisions, so tests of predictive validity are essential.

A test of the model against a holdout sample is sometimes described as a test of predictive validity, but this is not really true for most mature markets. Such markets tend to be stable with very similar data in adjacent time periods. Even promotional activity tends to follow a similar pattern and to be relatively constant from one year to the next. In such markets, a test against an adjacent holdout sample is really just a test of the robustness of the model to minor changes in conditions. Failure implies the model is not robust, but success does not necessarily imply the ability to predict the effects of major marketing changes, which is the purpose of market share modelling.

Furthermore, some models, such as those using Guadagni and Little's (1983) loyalty parameter, only forecast the *next* time period (or purchase). Such models are often reported as having been tested against a substantial holdout sample, but they should properly be seen as having been tested against a large number of *one period* holdout samples. There is a limited amount of change possible in *one period*.

Ultimately, the falsifiability of market share models depends on their predictive validity, and better ways need to be found to assess this. Prediction to an adjacent holdout sample is a start, and may validate a forecast under conditions of *no change*, but prediction to a holdout sample which is not adjacent to the estimation sample would be a stronger test, as would the use of a holdout sample which included major marketing changes which were absent from the estimation sample. The best test would be prediction of future market shares after some planned major marketing action - in effect a field experiment.

Some excellent experimental work on market mix impacts has emerged in recent years as a result of the BehaviorScan experiments (eg. Lodish, Abraham, Livelsberger, Lubethkin, Richardson, and Stevens 1995, Lodish, Abraham, Kalmenson, Livelsberger, Lubethkin, Richardson, and Stevens, 1995). Of course these are not true field experiments, but are rather conducted under highly controlled conditions using a previously recruited panel of respondents. These experiments are also usually used to assess alternative marketing mixes, rather than to test market share models. Nevertheless, this work points out a path that market share modelling also needs to follow.

Research on market share models could also be improved by having stronger multiple competing hypotheses. Comparing performance against a naive model is again a start, but comparison of seriously competing models and data collection methods would be better. For example, the predictive ability of a multiplicative model based on scanner data could be compared with that of a multinomial logit model based on experimental data from the same population. Similarly, the predictive validity of a market share model could be compared with that of managerial judgements or of survey based demand estimates.

The second problem is a lack of replication among researchers. The practice seems to be that each researcher builds their own model for their own data sets. The resulting analysis provides tentative results, but these do not often seem to be followed up by replications and extensions. Instead, they merely seem to form a basis for further theory development and speculation. Consequently, much of this work does not progress to replication and extension, but seems trapped in perpetually developing new theories and getting tentative results, never going on to the "normal science" component of consolidating and extending the empirical results (c/f. Kuhn 1970).

What we really need to know is whether particular model forms, specifications, and estimation methods predict equally well in different conditions (different product categories, different geographical areas, different researchers, etc), and which of the alternative forms, specifications, and estimation methods performs the best over these various conditions. Finally, we need to know in which areas the model systematically fails. These questions can only be addressed through replication.

As noted earlier, this section has not attempted to review the voluminous literature on market share modelling. Instead, it is intended, along with the previous illustrations, to demonstrate how the application of our criteria can guide the development and assessment of a research programme.

SUMMARY

Progress in developing marketing knowledge is impossible unless the theory involved is falsifiable. Empirically testing a theory involves an assumption of falsifiability, but to have confidence in such tests we need to replicate them in different circumstances to

ensure that the original results are not spurious. Developing generalisations also relies on replication as a means of extending the theory to new situations, and of investigating the influence of different variables. Falsifications (if they can be replicated, and are not just problems of observation) indicate the empirical limits or boundaries of a theory, and suggest directions for future research and theoretical improvements - we should actively seek to find these boundary conditions. By comparing competing theories, we also acknowledge that although none of our theories may be absolutely true, this does not mean that we know nothing at all, or should abandon them all, or are incapable of making progress.

Consequently, we contend that progress in marketing science requires the following steps.

1. Develop a falsifiable theory or theories. This requirement addresses the question of whether our theory is scientific, or merely "pseudo"-scientific.

2. Undertake a number of tests of the theory, including attempts to extend the theory to new situations. Undertaking close replications guards against the problems of observation, while differentiated replications establish the generalisability of the theory and identify any boundary conditions.

3. Improve the theory by modifying in response to the boundary conditions identified in Step 2, and then return to Step 2 to test this modification. Alternatively, generate a revolutionary new theory (effectively returning to Step 1).

Steps 2, and the first part of Step 3 are the type of activity characterised as "normal science" by philosophers of science such as Kuhn (1970). Most work within a scientific discipline is directed towards these activities. This does not relegate creative processes to unimportance, as creativity is essential to generate and modify theories, and to think up severe tests of theories. However, we believe that the problem facing marketing is not a lack of creativity (as represented by theory development) but rather a lack of rigour in following the processes of "normal science".

The examples in this paper have illustrated the application of these steps. “Normal science” appears to have been successfully applied in the case of NBD/Dirichlet theory. In the case of the Servqual instrument, “normal science” uncovered the problems with

the original instrument, but the various modifications and replacements do not appear to have been subject to the same standard of investigation. Perhaps they will in any event be replaced by competing theories about the link between service quality and *profitability*. Market share modelling does not appear to have progressed to the "normal science" stage, but instead is mainly concerned with process of theoretical development found in Step 1.

Too often, we seek to make "progress" merely by publishing more and more theoretical refinements, confirmed through one-off empirical studies. If the criteria presented in this paper are followed, this type of activity will be recognised as a minor first step. Progress in "normal science" primarily occurs through replications and extensions, and identifying the conditions under which our theories do, *and do not*, hold.

REFERENCES

Anderson, Eugene, W.; Fornell, Claes and Lehmann, Donald R. (1994), Customer Satisfaction, Market Share, and Profitability: Findings From Sweden, *Journal of Marketing*, 58, July, 53-66.

Anderson, Paul F. (1983), Marketing, Scientific Progress, and Scientific Method, *Journal of Marketing*, 47, Fall, 18-31.

---- (1986), On Method on Consumer Research: A Critical Relativist Perspective, *Journal of Consumer Research*, 13, September, 155-173.

---- (1988a), Relative to What - That is the Question: A Reply to Siegel, *Journal of Consumer Research*, 15, June, 133-137.

---- (1988b), Relativism Revidivus: In Defense of Critical Relativism, *Journal of Consumer Research*, 15, December, 403-406.

Armstrong, J. Scott (1991), Prediction of Consumer Behaviour by Experts and Novices, *Journal of Consumer Research*, 18, September, 251-256.

Armstrong, J. Scott; Brodie, Rod. and Parsons, Andrew. (1997), Hypotheses in Marketing: empirical evidence v. expert opinion v. current practice, *Unpublished working paper*, Auckland University, Auckland.

Armstrong, J. Scott and Schultz, Randall (1993), Principles involving marketing policies: an empirical assessment, *Marketing Letters*, 4, 3, 253-265.

Barnard, N.R.; Ehrenberg, A.S.C.; Hammond, K. and Uncles, M. (1994), "Two Debatable Dirichlet Discrepancies", *Working Paper*, South Bank Business School, London.

Barwise, T. Patrick (1995), Good Empirical Generalisations, *Marketing Science*, 14, 3, Part 2, G29-G35.

Bass, Frank M. (1995), Empirical Generalisations and Marketing Science: A Personal View, *Marketing Science*, 14, 3, Part 2, G6-G19.

Buttle, Francis (1996), SERVQUAL: review, critique, research agenda, *European Journal of Marketing*, 30, 1, 8-32.

Chalmers, Alan F.. (1982), *What is this thing called Science? second edition*. University of Queensland Press: St Lucia, Queensland.

Chalmers, Alan F. (1990), *Science and its Fabrication*. University of Queensland Press: St Lucia, Queensland.

Cronin J. Joseph and Steven A. Taylor (1992), Measuring Service Quality: A Reexamination and Extension, *Journal of Marketing*, 56, July , 55-68.

Cronin J. Joseph and Steven A. Taylor (1994), SERVPERF Versus SERVQUAL: Reconciling Performance-Based and Perceptions-Minus-Expectations Measurement of Service Quality, *Journal of Marketing*, 58, January, 125-131.

Ehrenberg, A.S.C. (1994), Theory or Well Based Results: Which Comes First? in Laurient, G.; Lilien, G; and Pras, B (eds) *Research Traditions in Marketing,* 79-108. Kluwer.

Ehrenberg, A.S.C. (1988), *Repeat Buying: Facts, Theories and Applications.* Oxford University Press.

Ehrenberg, A.S.C. and Uncles, M. (1988), Dirichlet Types*, Working Paper*, South Bank University.

Engel, James F.; Blackwell, Roger T. and Kollat, David T. (1978), *Consumer Behaviour: 3rd Edition*. Dryden Press.

Greenwald, Anthony G.; Pratkanis, Anthony R.; Leippe, Michael, R. And Baumgardner, Michael, H. (1986), Under What Conditions Does Theory Obstruct Progress? *Psychological Review*, 93, 2, 216-229.

Hawkins, Roger J.; Best, Kenneth and Coney, A. (1992), *Consumer Behaviour for Marketing Strategy*, 5th edition, Irwin.

Hirschman, Elizabeth C. (1986), Humanistic Inquiry in Marketing Research: Philosophy, Method, and Criteria, *Journal of Marketing Research*, 23, 237-249.

Howard, John A. and Sheth, Jagdesh, N. (1969), *The Theory of Buyer Behavior.* Wiley.

Hubbard, Raymond, and Armstrong, J. Scott (1994), Replications and extensions in marketing: Rarely published but quite contrary, *International Journal of Research in Marketing*, 11, 233-248.

Hunt, Shelby D. (1984), Should Marketing Adopt Relativism? in Anderson, Paul F. and Ryan, Michael, J (eds) (1984), *Scientific Method in Marketing.* American Marketing Association: Chicago.

--- (1990), Truth in Marketing Theory and Research, *Journal of Marketing*, 54, July, 1-15.

--- (1992), For Reason and Realism in Marketing, *Journal of Marketing*, 56, April, 89-102.

--- (1993a), Objectivity in Marketing Theory and Research, *Journal of Marketing*, 57. April, 76-91.

--- (1993b), *Modern Marketing Theory: Critical Issues in the Philosophy of Marketing Science.* South-Western.

Kuhn, Thomas. (1970), *The Structure of Scientific Revolutions 2nd Edition*. University of Chicago Press, Chicago.

Leone, Robert P. and Schultz, Randall, L. (1980), A Study of Marketing Generalisations, *Journal of Marketing*, 44, Winter, 10-18.

Lindsay, R. Murray and Ehrenberg A.S.C. (1993), The Design of Replicated Studies, *The American Statistician*, 47, 3, 217-228.

Lodish, Leonard M.; Abraham, Magid M.; Livelsberger, Jeanne; Lubethkin, Beth; Richardson, Bruce; and Stevens, Mary Ellen (1995), A Summary of Fifty-Five In-Market Experimental Estimates of the Long-Term Effect of TV Advertising, *Marketing Science*, 14, 3, Part 2, G133-G140.

Lodish, Leonard M.; Abraham, Magid M.; Kalmenson, Stuart; Livelsberger, Jeanne; Lubethkin, Beth; Richardson, Bruce; and Stevens, Mary Ellen (1995), How TV Advertising Works: A Meta-Analysis of 389 Real World TV Advertising Experiments, *Journal of Marketing Research*, 32, May, 125-139.

Marketing Science (1995), *Special Issue on Empirical Generalisations in Marketing*, 14, 3, Part 2.

Parasuraman, A.; Zeithaml V.A.; and Berry, L.L. (1994), Reassessment of Expectations as a Comparison Standard in Measuring Service Quality: Implications for Further Research, *Journal of Marketing*, 58, January, 111-124.

Peter, J. Paul (1992), Realism or Relativism for Marketing Theory and Research: A comment on Hunt's "Scientific Realism", *Journal of Marketing*, 56, April, 72-79.

Peter, J. Paul and Olson, Jerry C. (1994), *Understanding Consumer Behaviour*, Irwin.

Popper, Karl R. (1935), *Logik der forschung.* Julius Springer Verlag.

Rust, Roland R.; Zahorik, Anthony J. and Keiningham, Timothy L. (1995), Return on Quality (ROQ): Making Service Quality Financially Accountable, *Journal of Marketing*, 59, April, 58-70.

Stern, Philip, and Ehrenberg A.S.C. (1995) Differentiation and Segmentation in the Pharmaceutical Market, *Marketing Downunder*, 1995 Marketing Science Conference, Sydney.

Tuck, Mary (1976), *How Do We Choose?* Methuen: London.

Uncles, M.; Ehrenberg, A.S.C. and Hammond, K. (1995), Patterns of Buyer Behaviour: Regularities, Models, and Extensions, *Marketing Science*, 14, 3, Part 2, G71-G78.

Uncles, M. and Ehrenberg A.S.C. (1990), The Buying of Packaged Goods at U.S. Retail Chains, *Journal of Retailing*, 66, 3, 278-296.

Zaltman, Gerald (1991), Assessing Progress on Meeting MSI Priorities, *Marketing Science Institute Report*, Number 91-107.

Zinkhan, George, M and Hirscheim, Rudy (1992), Truth in Marketing Theory and Research An Alternative Perspective*, Journal of Marketing*, 56, April, 80-88.