1. Introduction

Empirical generalisations (EGs) are a central feature of scientific investigation. An oft-cited example is Boyle’s Law, stating that for any given body of gas, the greater the volume, the less the pressure. This relationship between the pressure $P$ and the volume $V$ of a body of gas is expressed formally as $P = C/V$, where $C$ is a constant. There are hundreds of other generalisations in the annals of the natural, biological and biomedical sciences, and even in the social sciences. The purpose of this article is to examine some of the characteristics of EGs as they relate to marketing by addressing six questions:

Q1: What is an empirical generalisation?
Q2: Why are empirical generalisations so important?
Q3: What is a “good” empirical generalisation?
Q4: How are empirical generalisations established?
Q5: What inhibits the establishment of empirical generalisations?
Q6: How is the establishment of empirical generalisations to be facilitated?

2. Definition of an Empirical Generalisation

An EG is a relationship between two or more variables that has been observed across a range of conditions. Because the relationship is observed to occur and recur it is regarded as a pattern, regularity or law, and it can be represented formally using a mathematical, graphic or symbolic language. By knowing that an observed relationship holds under a range of conditions (and that it does not hold under other conditions) it is possible to use knowledge of the relationship for practical purposes, such as making routine predictions and stating principles. It is also possible to start to theorise why the relationship occurs, and why it holds under some conditions and not others, thereby moving from empirical description to theory-building.

Some aspects of this definition are worth emphasizing:

(a) It is a relationship, such as the fact that pressure varies inversely with the volume of a body of gas (Boyle’s Law). Or – outside of the world of the natural sciences – the observation made by Auerbach almost a hundred years ago that there is a regular relationship between the size of cities and their rank (defined in terms of population size) (the “rank-size rule”).

(b) The relationship is regarded as a pattern, regularity or law that has been observed to occur and recur. This results from repeated empirical observation. Boyle’s Law was investigated hundreds of times before it achieved its
taken-for-granted status: “[It] has been found to hold for different gases and mixtures of gases, different amounts of gas, different kinds of apparatus, different experiments, different times and places, when pressure is increasing and when pressure is decreasing” (Ehrenberg, 1975, pp 72-73). The rank-size rule has been studied across dozens of countries and across time for certain countries (summarised in Haggett, Cliff and Frey, 1977, chapter 4).

(c) Typically, the relationship is known not to hold under certain conditions. The empirical relationship between pressure and the volume of a body of gas does not hold “when the temperature changes, when there is a chemical reaction, when there is a leak in the apparatus, when there is physical absorption or condensation of the gas, or when we tried to prove the law at school” (Ehrenberg, 1975, pp 73). These conditions may themselves form the basis for further EGs. Given the limitations of Boyle’s Law, for example, a natural extension would be to determine the relationship between pressure and volume when the temperature changes. The answer to this query forms the basis of the Gas Equation: PV=RT, where R is constant for any given amount of gas and T is the temperature (Ehrenberg, 1995, pp 145-146). The rank-size rule has been applied at the state level and it fits quite well for U.S. cities, but in Australia five of the six states are dominated by exceptionally large urban centres and the rank-size ratio is much lower than the rule would predict (quoted Haggett et al., pp 112-113). But, with urban growth in places like the Gold Coast and the Sunshine Coast, Australian states may not now be so exceptional – this is a question for further empirical investigation.

(d) The relationship can be represented using a mathematical, graphic or symbolic language, thus Boyle’s Law is simply denoted as P = C/N, or we can say the product PV remains constant C. The rank-size rule is shown as p_i = p_1/i, where p_i is the population of the i_th town in the series 1,2,3, ... n in which all towns in a region are arranged in descending order by population, and p_1 is the population of the largest city. When population size of city is plotted against rank of city the relationship is visualised as a reverse J-shaped concave curve (or a downward sloping straight-line when plotted on a log-log scale).

(e) Knowing there is a relationship means it can be used to make routine predictions and formulate principles. Thus, if we know C and V it is a simple matter to calculate P – but note that this is only simple because the underlying EG has been established. In the absence of an EG it would be very risky, if not foolhardy, to make predictions of P; while the relationship might be known to hold once, this tells us nothing about whether it would hold again in another place or at another time. In marketing there is a practical need for predictions and principles, but too often all we know is that the relationship holds once (Ehrenberg and Bound, 1993).

(f) Theorisation is possible too. Once it is known that pressure varies inversely with the volume of a body of gas (under certain known conditions), it is worthwhile to theorise why this is so. Similarly, once it is known that the largest city will dominate a country (in a predictable way) there is interest in knowing why. Is it because of the agglomeration of business activities, the centralisation of government functions, the focus of transportation systems, or all of these factors working in combination? And, if the effect is so systematically observed, is it virtually inevitable – “a law” - or is it something that occurs very often – “a lawlike relationship” – but it could be altered (perhaps, in the case of the rank-size rule, it could be altered through regional economic policies)? In marketing, approximate lawlike relationships are to be expected, not laws in a strict or universal sense.

The definition of EGs provided here also helps us appreciate what is not an empirical generalisation.

First, there are empirical studies that do not generalise; for instance, isolated facts (“today the market share of Colgate is 60%”), one-off cases (the story behind the launch of Macleans Whitening Gel Stripe toothpaste in the Sydney marketplace), ad hoc anecdotes (a manager’s interpretation of the Macleans story), and isolated experiments (a one-off test among 30 students of a hypothesis about toothpaste brand choice). These empirical studies have their place; at the very least they can stimulate ideas and thoughts that deserve to be studied systematically at a later date. Unfortunately, the temptation is to generalise from these kinds of studies – to generalise from a Harvard-style case, to treat the anecdote as a fact, and to assert general conclusions from isolated experiments. The dangers of doing this without offering caveats, or with a disregard for cautionary statements to warn readers that results should not be generalised, have been emphasized repeatedly (Leone and Schultz, 1980; Monroe, 1992; Hubbard and Lindsay, 2002).

Second, there are generalisations that are not empirical. Abstract mathematical theorems based on sets of well-defined axioms fall into this category. So too do the
normative theories of economics; these serve the useful purpose of helping us to think how the world could be or ought to be, given certain assumptions or axioms. Such theories exhibit many characteristics that are associated with EGs (there is a focus on relationships, a formal language is employed, etc.), but they are not, and not intended to be, EGs.

Also in this category are abstract frameworks – e.g., Maslow’s hierarchy of needs, Howard and Sheth’s model of consumer behaviour, and Dick and Basu’s model of customer loyalty. Arguably, these are elegant conceptualisations of the combined effects of attitude and behaviour on various aspects of consumer activity, but in their initial formulations they were not operationalised nor tested. Such abstract models are empirically and theoretically problematic, in that there is no systematic empirical evidence to support the framework and no formal language to express the relationships. They are simply boxes and arrows to depict a logic that seemed plausible to their authors.

Looking at the discipline of marketing we see many empirical studies that might generalise and many generalisations that could have an empirical grounding, but as they stand they are articles of faith or statements of internal logic whose “truth status” is unknown (Uncles, 2002; November, 2004). They could be generally right, or generally wrong.

3. The Importance of Empirical Generalisations

EGs are of crucial importance for four major reasons.

First, they are a basic form of marketing knowledge (see Rossiter 2002 on the five forms of marketing knowledge). They describe relationships that are observed to exist over a range of different circumstances which have been studied systematically. Good description, in turn, provides a basis for making routine predictions and highlights associations that are worth explaining. This knowledge then can be embodied in principles. This whole process is clearly and unambiguously articulated, allowing independent verification of any claims made by the original researchers. The Double Jeopardy pattern is an example that has been widely observed over time, across places and in many product categories. It states that competing brands vary little in terms of average frequencies of purchase compared to substantial differences in the size of their customer base. Many other candidate EGs have been proposed, ranging from the observation that the price elasticity for closely substitutable brands is -2.6, to the finding that higher market share brands are less deal elastic (see a summary table in Bass and Wind, 1995).

Second, EGs are the building blocks for more complex knowledge. Boyle’s Law was a stepping-stone on the way to deriving the more general Gas Equation (the Law is a special case of the Equation when temperature T is constant). Or, from the area of buyer behaviour analysis, consider the NBD-Dirichlet model (Ehrenberg, Uncles and Goodhardt, 2004). This is an amalgam of observations, patterns and models that have accumulated over four decades. Early on it was shown that for frequently bought goods the distribution of the number of purchases of a brand or product category is predictable. An empirically-grounded theory was put forward to account for these patterns (encapsulated in Repeat-Buying theory and the NBD model). More recently, attention has turned to brand choice, as well as purchase incidence (including the Double Jeopardy pattern that links the two). Once again, it proved possible to put forward an empirically-grounded theory (this time in the form of the NBD-Dirichlet model). This more complete model integrates findings that were previously considered separately, but this would not have been possible without the building blocks being thoroughly established as basic forms of marketing knowledge. The Generalised Bass model (GBM) in the diffusion literature is another example of a more complete model built upon earlier partial models and observations. In its original form the Bass model describes the diffusion of innovation, whereas only in the GBM are decision variables such as price and advertising included (Bass, 1995).

Third, EGs have utility. They can be used. For instance, knowledge of Double Jeopardy shows a marketing manager what to expect when attempting to grow sales; sales growth is more likely to come from securing more customers than it is from expecting a dramatic increase in the frequency of purchase from existing customers. This implies programs designed to enhance loyalty (e.g., loyalty, affinity and rewards programs) will also need to expand the customer base if they are to succeed. The alternative path of securing more frequency than expected given the size of the customer base could be attempted, but that would mean going against the grain. That would be hard to achieve, especially if directly competing brands simultaneously launch programs to enhance loyalty. Another sense in which EGs are useful is that they highlight deviations (where there are norms it is easier to see deviations). Some of these deviations will prove to be random, and of little general interest.
Others, however, might prove to be systematic and deserving of further investigation. For instance, a number of systematic discrepancies are associated with the NBD and NBD-Dirichlet models (for some market leaders annual purchase frequencies are a unit or so higher than predicted by the models, Fader and Schmittlein, 1993).

Fourth, **EGs are a guard against falsehoods and unsubstantiated claims to knowledge** (Hubbard and Armstrong, 1994; Hubbard and Lindsay, 2002; Wright and Kearns, 1998). Falsehoods take several forms. In a research context, the falsehood may be the perpetuation of erroneous and questionable results – something that can easily occur when a relationship is asserted and popularised on the basis of a single “innovative” or “original” study. A further study might quickly show the relationship to be false. This would not be surprising because there are plenty of things that can go wrong in a research study, particularly one that is highly innovative or original (not least: sampling error and imperfect data collection procedures; measurement error and encoding error; analytical error, confounding influences and omitted variables; interpretation error, subjectivity and bias; and plain dishonesty from the observing researcher). Added to which is the fear that editors and reviewers are biased against null results, possibly leading to the proliferation of Type 1 errors (i.e., erroneous rejection of the null hypothesis) (Hubbard and Armstrong, 1992). Without additional studies these errors go unchallenged.

In a managerial context, many things that marketers believe to be true will be true because managers have opportunities to test their ideas everyday (these “truths” are rules of thumb and folk wisdom, or what have been dubbed “managerial factoids” and “CUGs” – currently useful generalisations). But, equally, there will be things that managers wish to be true that are not – because they are predicated on myths, wishful thinking and an “anything goes” mentality that ignores constraints and market norms. The problem is how to distinguish between the truths and falsehoods in the absence of systematic empirical studies to identify generalisations, norms and constraints.

4. **“Good” Empirical Generalisation**

Even when we have EGs it is inappropriate to treat them all alike. Some rely on shoddy evidence. Undoubtedly, some are trivial and trite. Others are very well established and describe important relationships. What, then, are the characteristics of a “good” EG? A very clear answer to this question was provided by Barwise (1995), who suggested there are five criteria to consider when assessing an EG.

(a) **Scope-Boundary Conditions.** An EG that is known to be true over a certain range of conditions is said to have scope. Thus, Boyle’s Law holds for different gases and mixtures of gases (which can be specified), different amounts of gas (which can be quantified), different kinds of apparatus (which can be demonstrated), and so forth. The rank-size rule holds for different countries. Because these relationships are known to be true it is possible to make routine predictions: we can predict pressure P from knowing C/V or we can predict the population of the jth city knowing p_j/i. However, EGs do not have to be true everywhere. Boyle’s Law does not hold when the temperature changes, and even the more general Gas Equation only holds in a strict sense for perfect gases. The rank-size rule might not hold within a megalopolis where several large cities have coalesced (e.g., the north-eastern seaboard of the U.S. from New Hampshire to Maryland) or in regions where urban growth is strictly governed by planners (e.g., Randstad-Holland in the Netherlands).

A good EG is one where these “boundary conditions” are known as well as the scope. The ability to articulate boundary conditions is a sign of the maturity of an EG, as well as offering protection against vagueness and overstatement. The range of possible boundary conditions is vast – goods versus services, high versus low market shares, new products versus existing ones, and so forth. If boundary conditions have not been examined the generalisation must be regarded as preliminary.

(b) **Precision.** If routine predictions are to be made it is desirable to have generalisations that are reasonably exact (e.g., a relationship that can be expressed algebraically allows an analyst to insert known values and calculate unknowns). By contrast, loose or equivocal generalisations may be too vague for making routine predictions or, worse, misleading. On the other hand, too narrow a focus on precision tends to mean that other criteria are overlooked. For instance, a very precise result can be obtained by fitting a regression line to a single-set of data; this provides a “specifically best fit”, but it tells us nothing about whether the result will generalise to another dataset or another context. Of far more value is something that provides a “generally good fit”, even though it will be only approximately true in any specific instance. Therefore, EGs are best seen as approximations, which ignore minor perturbations for the time being.
(c) **Parsimony.** The world is complex and therefore it is easy to add layers of complexity to a research study; extra variables, more contingencies, and further nuances and subtleties. A quagmire of complexity now surrounds notions such as “trust” and “value”. Yet, in scientific investigation, less is more. Something that is simply expressed (though not necessarily simple in conception) is likely to be more memorable, more tractable, and more readily used. What could be more simple than \( P = C / V \), or \( p_i = p_j / i \), or \( E = MC^2 \)? Some of the most impressive leaps in scientific investigation have occurred when vast amounts of nuanced data have been reduced to quite simple EGs. The work of Kepler in astronomy and Mendeleev in biology are supreme examples of empirically-grounded generalisations that offer parsimony and precision. This is no less important in contemporary marketing, where we are faced with a morass of company, customer and competitor data, but are short on simple EGs that summarise patterns, regularities and laws.

(d) **Relevance.** A precise and parsimonious result, that is known to hold over a range of conditions, is useful in at least three ways: (i) to make routine predictions (e.g., gravitational force at sea level); (ii) to offer a benchmark or norm against which new information can be compared (e.g., gravitational force on the surface of the Moon is one-sixth that at sea level on Earth); and (iii) to provide a context for thinking about practical issues (e.g., in order to fly we will have to overcome the gravitational force of the Earth). Significantly, these are scientific uses of EGs, not engineering ones. Knowledge of gravity does not directly tell you anything about the various ways to overcome it (which are as ingenious as rockets, jet-engine aeroplanes, helicopters, propeller planes, balloons and blimps). Nevertheless, it helps to know that there is a gravitational force. Marketers are sometimes frustrated by this – they want answers to specific practical issues, not generalisations. They appear to want to build a better rocket (preferably one that is seen as distinctive from all other rockets), without necessarily knowing what a rocket has to do if it is to fly.

(e) **Basis for theory-building.** Having established an EG in terms of its scope and boundary conditions, there is then something to explain. There is then a need for theory to account for the scope and boundary conditions. Why does pressure vary inversely with the volume of a body of gas? Why does the relationship between pressure and volume change when the temperature changes? Why do largest cities dominate in a predictable way?

From a scientific standpoint, there would be little point in trying to explain these relationships if we did not already know of their existence. It is for this reason that Ehrenberg (1994) argues strongly for an “empirical-then-theoretical” approach (EtT) to knowledge generation. The researcher starts by systematically exploring data, looking for patterns, from which EGs are derived and, only then, is there a basis for theorising. This sharply contrasts with the “theory-in-isolation” approach (TIL), whereby the researcher puts forward an original theory or hypothesis, tests it with a small, isolated and arbitrary dataset, and assesses the test using inferential statistics.

Unfortunately the EtT versus TIL debate quickly degenerates into a chicken-before-the-egg problem (see Bass 1995 and commentaries by Bemmaor, Rossitter and Schmittlein in response to Ehrenberg 1994). Of course, hardly any research in an applied area such as marketing is likely to be purely theoretical. Likewise, no research is purely empirical. There is no tabular rasa. No pure induction. The process of undertaking research is theory-laden (in the sense of having to decide what to study, how to take measurements, which variables to consider, etc.). Ehrenberg is the first to emphasize the importance of (empirically-grounded) prior knowledge when examining new data. For instance, in looking at new data in product categories as diverse as chemical additives, paper and packaging, gasoline and aviation fuel contracts, the NBD and NBD-Dirichlet models have been used as reference points or benchmarks – this has seemed preferable to mindless data mining. See also Nelder (1986) who advises statisticians to take more account of prior knowledge and think more about the context from which data have been obtained.

Therefore, the issue is not necessarily whether the empirical (E) or the theoretical (T) should come first, but the status accorded E and T in the research process. Those seeking EGs emphasize the importance of data, systematically analysed, with a view to finding patterns (EtTEtT...). By contrast, those who fail to see the importance of EGs tend to be content with one-off studies that test a theory without any attempt to uncover patterns; were a series of tests to be undertaken in a systematic fashion patterns might become more apparent (TETE...).

4. **Establishing Empirical Generalisations**

To establish EGs empirical observations must be accumulated across studies. This much is agreed. However, marketers approach the task in two very different ways.
4.1 Reviews, Meta-Analyses and Best-Evidence Synthesis

By far the most common approach is to generalise from a set of empirical studies. For instance, a number of studies might have been undertaken in which the relationship between “advertising expenditure” and “brand sales” is investigated. At some point an attempt will be made to review and summarise the findings of these studies. Typically, this is done with a loose narrative discussion – the reviewer mentions result (b) from paper A and result (c) from paper B, but we never hear about result (b) in paper B, or result (c) in paper A, or result (a) at all. The problem is compounded by the fact that the studies are disparate: constructs are defined and measured in different ways, sample designs and sizes vary, significance thresholds differ, and so forth. It is as if we are being invited to compare apples and bananas, in a very casual way.

Meta-analysis offers the prospect of providing greater rigour. Formally, meta-analysis is the “analysis of analyses”. Models are compared using formal statistical procedures, enabling quantitative generalisation in that both the sign and values of a relationship can be reported (parameter estimates, elasticities, effect sizes, etc.). For instance, a meta-analysis of 128 studies shows advertising elasticity to be approximately 0.3 (Farley, Lehmann and Sawyer, 1995). Or, more specifically, those studies focused on short-run effects reveal an advertising elasticity of 0.27 and among long-run studies the advertising elasticity is 0.39. In this manner we end up with a situation where red apples are compared to green ones.

The statistical methods involved in meta-analysis provide useful techniques for combining datasets and for testing moderating variables – which may also be boundary conditions. However, meta-analysis does not promote either parsimony or an understanding of more subtle relationships. Like any technique, it is also restricted by the quality and type of data available. Usually it has to work with dependent and independent variables that have been measured in quite different ways and under different circumstances. Previously used analysis techniques might be quite different too; for instance, where the effects of advertising have been investigated with regression models and structural equation models. Moreover, the original studies can be of variable quality – it is then as if rotten apples are mixed with those that would win a prize, whereas it would be better just to consider the prize-winners.

An alternative to meta-analysis is to conduct a very formal and systematic review, taking care to record all significant results and to document key aspects of the research procedure (methods, sample sizes, etc.). The list then needs to be reviewed carefully, to determine which studies merit further consideration. Poor studies are discarded, leaving only the best evidence with which to construct the final review. Such “best-evidence synthesis” has been advocated in marketing (Rossiter, 2004; see also Slavin, 1986). The determination of judges and judging criteria are very important elements of this approach. For this reason perhaps there are few explicit examples in marketing where best-evidence syntheses has been undertaken, although in spirit the “advertising principles” and “forecasting principles” projects come close (Armstrong and Pagell, 2003 and associated commentaries).

The underlying problem with all these approaches is that they are trying to make sense of very disparate studies, where the establishment of EGs was not on the mind of the original researchers. This differs from approaches found in many other branches of science.

4.2 Replication

“That which isn’t worth replicating isn’t worth knowing” - this quote from Mittelstaedt and Zorn (1984) highlights the importance of replication as one of the pillars of normal scientific investigation. The point has been emphasized time and again (Monroe, 1992; Hubbard and Armstrong, 1994; Hubbard, Vetter and Little, 1998) and it was a central theme in the Journal of Business Research special issue on replications and extensions in marketing and management research (Easley and Madden, 2000; Easley, Madden and Dunn, 2000).

There are three forms of replication: repeat studies, close replication and differentiated replication. The latter two are particularly important from the standpoint of establishing EGs (Lindsay and Ehrenberg, 1993).

(a) Repeat studies are uncommon in marketing and many researchers would regard them as unnecessary: why re-do what has been undertaken already? Indeed, where we have confidence in the way the original study was designed, executed and interpreted there is little reason for repeating the work. However, as was pointed out earlier, many things can go wrong in a research study, particularly in a highly innovative one. A repeat study, conducted by an independent research team, enables sampling schemes, data collection procedures and measurements to be checked, and it is an opportunity to
assess the interpretations and biases of the original researchers. The rejection of cold fusion because of the failure to replicate results is a poignant example of the importance of these factors. Because repeat studies are so rare outside the natural sciences it is hard to say how many cold fusion stories are lurking in marketing – probably there are quite a few.

In the absence of repeat studies some researchers have resorted to using hold-out samples. These are of no value if the hold-out sample comprises randomly selected cases from the full dataset – by definition, the random sample must conform to the data from which it is sampled, other than for random sampling error. By contrast, a non-randomly selected hold-out sample might be of some use (e.g., separating out the most recent year from a 10-year record, or separating out the very young and very old from a sample of people of various ages). These are not tough tests, but they help establish the truth of the first study and begin the process of seeing whether an empirical relationship holds more than once.

(b) Close replication is where most aspects of a study are kept invariant, but not quite everything. There might be interest in observing the dependency on cars of shoppers at Edmonton Mall, with surveys undertaken in August and December – across these two surveys there may be no major change in the population of shoppers, the sampling procedure, the measurement techniques, the methods of analysis, nor the geography, but the fact that temperatures drop dramatically in the Canadian winter means that at least one condition varies (this could have an impact on car dependency or it might be that most shoppers drive irrespective of the season). Close replications like this serve as a quick check to see if an empirical relationship holds again, and whether there is at least some basis for generalisation.

(c) Differentiated replication is designed to discover whether a relationship generalises to different conditions. The conditions themselves are deliberately chosen, with a view to establishing the scope and boundary conditions of an EG. The fact that conditions are deliberately chosen means this form of replication can be undertaken in a reasonably formal and systematic way, which contrast with the more opportunistic approach of close replication. With differentiation the emphasis is not on doing more replications just for the sake of it, but of carefully considering new conditions

Figure 1: The Marketing Data Matrix

![Marketing Data Matrix Diagram]
where the EG might hold, observing whether it does hold, and then theorising why it holds. Or, in the case of a boundary condition, why the EG does not hold.

Effective replication (especially of the differentiated form) requires researchers to shift from a reliance on single sets of data (SSoDs) to many sets of data (MSoDs) (Ehrenberg, 1995). This can be visualised as drawing a sample from a data matrix comprising different places, times, product categories, etc. (see Figure 1, adapted from Haggett et al., 1977, pp 15). The researcher decides whether to move through the matrix along one dimension (say, time) or to sample from a block of the matrix (say, two time periods, one place, and ten product categories).

For example, in a study of consumer purchasing Uncles, Hammond, Ehrenberg and Davis (1995) looked at average purchase frequencies and share of category requirements in 34 separate product categories – not just Omo in the Laundry Detergent category for instance. This study was confined to the U.S., however in related work place was varied, with geographical extensions to Australia and New Zealand (Wright, Sharp and Sharp, 1998). Similarly, in looking at the exceptionally high average purchase frequencies of some leading brands, Fader and Schmittlein (1993) examined not just one product category but 39 in Japan and 28 in the U.S. All these examples concern the purchasing of packaged goods, but further studies have examined categories as disparate as aviation fuel contracts and drug prescribing behaviour (summarised in Ehrenberg et al., 2004). Through a process of replication the data matrix can be examined, seeing where EGs hold or fail.

The ideal is to have the validity of an EG tested by different researchers (inter-subjectivity), under varied conditions (deliberately and systematically selected from the data matrix), using multiple methods (with a view to achieving convergent validity). A summary of the whole process is shown in Figure 2. In practice, few EGs in marketing will be derived in exactly the manner described here because of constraints (available time, resources, access to databases, etc.) and of the need to rely on less formal processes. Nevertheless, researchers should not be deterred from moving in the direction of

Figure 2: The Process of Establishing Empirical Generalisations

```
Data & Priors

Original study
Is the observed result real?

More data
Repeat study
Was the original study conducted properly?

More data
Close replication
Does the original result hold again?

More data
Differentiated replication
What is the scope of the result?
Is it an EG?

What explains the EG?
What principles are implied by the EG?
What boundary conditions are there?

Varied conditions
Different researchers
Multiple methods
```
the process described in Figure 2. A good starting place is with the original study. Consider the researcher embarking on a project in the area of customer satisfaction, he can choose to focus on only high-involvement situations or examine low-involvement ones too, he can select only high-contact services (hairdressing) or low-contact services (on-line banking) as well. If a relationship was found across all four conditions the researcher might be well on the way to finding an EG. Starting the process of replication in this way with the original study serves three purposes: (i) it is a check on whether the relationship holds again (at least once), (ii) it demonstrates the relationship is replicable (if it can be replicated once it possibly can be replicated many times), and (iii) it signals the importance of replication (the original researcher believes this to be important, so perhaps others will be encouraged to conduct follow-up studies).

5. Problems Establishing Empirical Generalisations

A number of marketing EGs have been mentioned above. Noteworthy as these examples are, should there be many more of them? Opinions are divided. One view is that there are as many EGs as would be expected for a youthful discipline; compared to physics, say, marketing has had much less time to search for and find EGs. An alternative view is that given the predilection of marketers to make predictions, state principles and theorise (e.g., in textbooks and in the classroom) there should be far more EGs than we in fact see. There is concern that perhaps too many predictions, principles and theories rest on insufficient systematic empirical evidence.

A key problem is the absence of formal replications. Despite the importance of replication as an essential component of normal science, it is not routine in marketing. Zinkhan, et al. (1990) found that only 4.9% of empirical papers published in marketing journals were replications with extensions. In a similar study, Hubbard & Armstrong (1994) found that replications with extensions made up just 2.4% of published empirical papers, and that the frequency of publication of replications had been decreasing since the 1970s. A study across the disciplines of marketing, management, accounting, economics and finance for the years 1970 to 1991 reported similarly low rates of replication, averaging 6.2% of articles (Hubbard and Vetter, 1996). More recently, Danaher and Brodie reported ongoing problems in finding replication and extension studies in the area of price elasticity research (Danaher and Brodie, 2000).

Significantly, of the replications that have been published in marketing, Zinkhan, et al. (1990) found that 46% gave results that conflicted with the original study, while Hubbard and Armstrong (1994) found 60% gave conflicting results. Furthermore, sustained investigations of particular topics by a succession of researchers show that some widely-known original studies need to be couched around with caveats. In marketing, for instance, there have been at least eight replications of Aaker and Keller’s (1990) work on brand associations (summarised by Bottomley and Holden, 2001).

There are several reasons for the lack of replications, and therefore EGs, in marketing.

(a) The marketing mindset is not in tune with the goals of empirical generalisation. There are those who argue that because of the uncertain and unpredictable nature of human agency we are unlikely to find patterns or regularities in marketing (the argument is that management disciplines are more art than science and, as such, more concerned with specific contexts than generalisations) (for a critical discussion see Dowling, 2004). This viewpoint is hard to sustain when faced with evidence showing the existence of regularities, but a contemporary twist on the argument is to suggest that the combination of human agency and major contextual discontinuities (brought about by the Internet, the knowledge economy, coopetition, and so forth) makes most of our existing EGs look out-of-date and challenges our ability to find new ones (Sheth and Sisodia, 1999). In a world where the only constant is change, how can there be enduring EGs?

Others argue that marketers have an engineering mentality which requires the development of useful tools and techniques, in contrast to a truly scientific approach. They seek to build the marketing equivalent of a rocket, and are not too concerned to find the marketing equivalent of gravity. Yet others in marketing are prepared to accept there are EGs, but claim most of the patterns and regularities that have been found are trivial or merely corroborate what managers already know. As November (2004) writes: “… the proposition that, in general, advertising has a positive impact on sales is hardly news to a marketing practitioner” (pp 46).

The mentality carries over into marketing practice itself (or is a reflection of marketing practice), in that marketers are rewarded for the innovations they make. They seek to launch new brands, change existing ones, re-think communications strategies and be creative in
Empirical Generalisation in Marketing

getting goods to market. If any of these initiatives do not work the damage is unlikely to be life-threatening, unlike in scientific and biomedical areas – perhaps the manager loses her job (if she has not already had the wit to change job). It is as if marketers are doctors in a gold rush – they are very busy tackling business problems and making some difference, but they do not always make the most informed decisions, are not completely clear why they make a difference, and do not have the time to reflect on decisions and outcomes.

When taken together these mindsets are not compatible with the patient, steady and long-term outlook that is required to undertake a program of research leading to the establishment of EGs. And yet, at the same time, most marketers are quite prepared to make predictions and espouse principles in textbooks, in the classroom, in the popular press, and in practice as if there are empirically observable patterns, regularities and laws.

(b) The process of establishing EGs is perceived by marketers as time-consuming and laborious. This is a correct perception – for good reason. Like most things in life, effort has to be put into achieving quality outcomes; an Olympian swimmer must swim an untold number of laps before competing at an elite level, and the fledgling pianist must practice daily if he is to master the piano. Scientific inquiry is no different: “science is 10% inspiration and 90% perspiration”. Perhaps the problem in marketing is not so much that the process is seen as time-consuming and laborious, but that it is also seen as tedious, boring, trivial, unglamorous and ultimately not worth doing. Why examine ten datasets when it is possible to publish from just one? Why yet another study? What is the incremental contribution of an additional study? Why not leave the task of replication to others? In fact, why not leave the perspiration to others and focus on the inspiration? The problem is compounded by a Catch-22 effect that a researcher who finds no support for the original study will be accused of not repeating the study carefully enough, whereas if support is found then there is the view that nothing new has been learned (Monroe, 1992). So, why bother?

(c) Publication processes are blamed for having a bias against replications, which in turn reduces the chances of finding EGs. This perceived bias takes various forms: editors, it is argued, do not see replication as worth the journal space; referees primarily look for innovative work (especially new techniques); and the criteria used by editors and referees focus on technical proficiency not the establishment of empirically-grounded knowledge, nor the repeated application and use of techniques, nor even the likelihood of replication and use (replicability is rarely listed as an evaluative criterion). The overall effect of this is to discourage researchers from undertaking replications – with the effect of turning a perceived bias into the reality of few published replications (Armstrong, 1982; Madden, Easley and Dunn, 1995). Why engage in this activity if it is not rewarded? Why raise this with early-career researchers if it will not help their career prospects? This, of course, is to miss the point of normal scientific inquiry: “It largely is (or should be) a professional activity concerned with developing “objective” (meaning “widely agreed upon”) knowledge and understanding and also practical applications of this, rather than with the researcher’s self-promotion and possible tenured appointment” (Lindsay and Ehrenberg, 1993, pp 218).

Of course, some replications are published. But the perception of bias carries over into the arena of published replications. Specifically, there is a view that replications which do not support the original study are more likely to be published (this might help to counter the problem of proliferating Type I errors, but is then in danger of committing Type II errors). Also, invariably an extended discussion of the deviations is seen as more interesting than an account of “mere regularities” (in fact the reverse ought to be the case: exceptional deviations are of limited interest – because they are unique, idiosyncratic and exceptional perturbations – whereas it is the pattern or regularity that deserves our attention and any systematic deviations from the main effect).

(d) There are problems of a methodological nature. First, there is the issue of how to design replication studies. This is not easy. Figures 1 and 2 highlight some considerations, but there are no neat answers to questions such as which cells in the data matrix to examine initially, or how many axes of the matrix ought to be considered, or how many replications are enough. In practice, data become available to researchers in different ways. Perhaps there is ready access to temporal information, but only for one country – the researcher has to accept these constraints for the time being, and exercise common-sense in the way such data are used. Real datasets are messy. It is not possible to control for every extraneous factor. In such circumstances it is not possible to be highly prescriptive. There are guides (e.g., Lindsay and Ehrenberg, 1993), but these stop-short of offering a cook-book approach, in contrast to the more controlled world of experimental design.
Second, the statistical toolkit with which most marketing researchers are equipped is deficient for looking at MSoDs. The toolkit they have is ideal for drawing inferences from an isolated SSoD where \( n \) is greater than 30 and less than 1,000, to which a “best-fit” model can be applied. This was exactly the situation faced by W.S. Gosset, the man who published “Student’s t” as the solution to the problem of describing the sampling distribution of the mean for single, fairly small, datasets (Ehrenberg, 1975, p 302). But that was in 1908. Nowadays we have multiple datasets with which to undertake replications, and yet we remain under the sway of “the cult of the isolated study” (Nelder, 1986). There is little understanding of such basic ideas as sampling cases from the data matrix, or of the concept of an “approximately good fit” across a number of datasets or of the desirability of identifying “significant similarities” (which, arguably, is of more use than the habit of focusing most effort on looking for “significant differences”). Procedures are available, ranging from “eye-ball” through data reduction and exploratory data analysis, to formalised pattern-recognition techniques (such as neural networking, data-mining and visualisation), and finally to expert systems. On the whole these procedures have not entered mainstream research in marketing, either in terms of usage or training. The problem was articulated twenty years ago by Nelder (1985) and there has been surprisingly little improvement over the intervening years: “The search for the reproducible result is the search for sameness, in contrast to the emphasis on the significant difference from a single experiment. The problem here is that the scientist or technologist may be quite unaware that statistics has a useful role in this kind of activity. Certainly most textbooks give little sign that it has”.

6. Facilitating the Search for Empirical Generalisation

“That there is more order in the world than appears at first sight is not discovered till the order is looked for” (quoted by Haggett et al, 1977, pp 1). This statement upholds us to look for patterns, regularities and laws, with a view to establishing EGs. This can be facilitated in several ways; here we mention five key steps.

First, the establishment of EGs needs to be given higher priority in the discipline. At an institutional level this has been recognised for some time. For example, twenty years ago the American Marketing Association had a taskforce concerned with the development, dissemination and use of marketing knowledge; periodically empirical generalisation has been on the Market Science Institute’s list of priority areas; and recently it appeared as a special interest topic for the U.K.’s Academy of Marketing. However, institutional “calls to action” do not by themselves change the attitudes and behaviours of individual academics. These changes are more likely to occur if the search for EGs is seen as a routine component of normal scientific investigation. This means: including the topic on the syllabus of research training programmes and doctoral colloquia, having students undertake replications as part of their research training, and getting the topic featured in standard market research and market analysis textbooks. Some progress has been made in having research students undertake replications, but there are few signs that this training has flowed down to non-research students and almost no standard market research textbook includes a discussion of EGs, replication, meta-analysis or the analysis of many sets of data (e.g., neither McDaniel and Gates, 2002 nor Smith and Albaum, 2005 address these topics although both claim to be introducing marketing students to scientific methods).

Second, research needs to be set up in such a way that it can be replicated. “Replicability … is the Supreme Court of the scientific systems” (Collins, 1985), but it is not uncommon in marketing to be faced with a lengthy paper that does not contain the basic information which would allow an independent researcher to replicate the study (e.g., sample sizes, sample characteristics, copies of the questionnaire and other survey instruments, software and algorithms used, assumptions regarding the treatment of missing values, descriptive statistics, etc.). This information may not be an exciting read, but it can be placed succinctly in appendices or on the journal website. Online repositories of raw data might be useful too, although without knowing how the data have been cleaned and edited it would remain difficult to undertake replications.

Third, systematic approaches to research should be encouraged. Ideally, this is achieved by building replications and extensions into programmatic research, along the lines described in section 4.2. But it also can be achieved through systematic empirical tests of prior findings in areas where replications have been rare, as discussed in section 4.1 in relation to systematic literature reviews and formal meta-analyses. Literature reviews should be comprehensive and insightful, and should provide useful, evidence-based, generalisations for researchers. Meta-analyses that statistically combine...
prior results to develop robust findings should pay particular attention to potential moderator variables and possible boundary conditions. Importantly, these activities cannot be done effectively without some innovative thinking and inspiration. As we saw earlier, charts like Figures 1 and 2 can offer guidance, but there is no cookbook to say how replications should be undertaken.

Fourth, the methodological challenges of searching for EGs need to be better understood and studied. A statistically significant result obtained from one study does not say anything about whether the result will hold again if the study was replicated under different conditions. Getting a statistically significant result once is merely a first step (all it says is that the observed result is probably real and not due to sampling error), there then needs to be a process of replication. When statisticians like Gossett were writing it may have been difficult to obtain additional datasets, but today we can readily access datasets from across the data matrix. We are drowning in data and have virtually unlimited computing power. Therefore, methods, procedures and techniques that are appropriate for the analysis of MSOsDs need to come to the fore.

Fifth, there may be a need to change aspects of our publication processes. Opinions appear to be divided on how to do this. One view is to see empirical generalisation as a specialist topic – not unlike substantive topics in marketing (“brand management”, “marketing communications”, etc.). It then follows that there should be specialist journals (e.g., Journal of Empirical Generalisations in Marketing Science, www.empgens.com) and special issues (e.g., Marketing Science in 1995, the Journal of Business Research in 2000, and this issue of the Australasian Marketing Journal). In tangible and visible ways these “specials” keep the topic of empirical generalisation on the agenda. The danger, however, is authors in mainstream journals continue to ignore the topic, seeing it as something others do. This puts EGs in a ghetto.

An alternative view argues that the mainstream must change. This is most likely to be brought about if refereeing criteria alter. Referees need to ask: has there been any attempt by the authors to replicate the work (a “replicated” criterion)? Can the work be replicated from the information provided in the article or from a publicly-accessible website (a “replicable” criterion)? It might also be useful to question the likelihood of the work being replicated. Such criteria are now to be found in leading psychology, political science and social research journals where there has been a shift away from reliance on isolated experimental studies.

An intermediate position is to include a replications section in mainstream journals or, at the very least, to have a section devoted to “rejoiners, research notes and commentaries” where EGs can be published and discussed. Such sections are highly valued in other disciplines (e.g., in the decision-sciences journal Interfaces and in the Journal of the Royal Statistical Society in statistics, both of which publish commentaries and sometimes re-analyses alongside the original article). The record in marketing is patchy: the Journal of Marketing Research has had notes, abandoned them and reinstated them; the Journal of Consumer Research currently has a “re-inquiries” section; Marketing Science has created a special submission category for empirical papers. Unfortunately, as the experience of Journal of Marketing Research shows, these sections do not necessarily survive changes of editorship.

Steps such as these ought to lead to more EGs appearing in print – generating the marketing equivalent of Boyle’s Law and Auerbach’s rank-size rule. Collectively, the actions described here should help advance the discipline and provide the rewards and recognition that individual researchers seek. In doing this, replication and the establishment of EGs should become second-nature, not a rare and exotic activity undertaken by a minority of researchers.

References


**Acknowledgements**
We thank all those contributors and reviewers who were involved in the Special Issue of the Australasian Marketing Journal on Empirical Generalisation in Marketing. Our gratitude is also extended to attendees at the special EG session during ANZMAC 2003, hosted by the University of South Australia. Discussions at this event helped to develop and refine our thoughts.

**Biographies**

**Mark D. Uncles** is Professor of Marketing, School of Marketing, University of New South Wales. He started his academic career as a quantitative geographer, in which area empirical generalisation is readily accepted as a primary goal of research. In marketing, the search for patterns, regularities and laws has been integral to his work in the areas of buyer behaviour, brand and store choice, and loyalty.

**Malcolm Wright** is NZ Managing Director, Robt. Jones Holdings Ltd. Previously he was Associate Professor in the Department of Marketing at Massey University. He has long standing interests in empirical generalisation, research methodology and the Dirichlet model. His work has appeared in a variety of international journals including the International Journal of Research in Marketing, the European Journal of Marketing, the Journal of Product Innovation Management and the Journal of the Market Research Society.

**Correspondence Addresses**
Mark D. Uncles, Professor of Marketing, School of Marketing, University of New South Wales, Sydney NSW 2052. Telephone: +61 (2) 9385 3385. Facsimile: +61 (2) 9663 1985. Email: m.uncles@unsw.edu.au
Malcolm Wright, NZ Managing Director, Robt. Jones Holdings Ltd., Level 6, 101 Lambton Quay, Wellington, New Zealand. Telephone: +64 (4) 473 5683. Facsimile: +64 (4) 473 5685. Email: mjwright@paradise.net.nz