Is consumer memory (really) Dirichlet-like?

Lara Stocchi

This research note discusses the suitability of the Dirichlet model in approximating the non-linear processes that characterise brand retrieval in consumer memory. This is achieved through the use of the Central Limit Theorem, which highlights a *significant sameness* between the combination of the statistical distributions describing information encoding, activation and retrieval, and the distributions that form the Dirichlet model. This demonstration reinforces the suitability of the Dirichlet model for predicting a brand's retrieval propensity through the analysis of the network of concepts consumers associate with the brand in their memory, collected by the use of surveys.

Keywords: Dirichlet model, brand associations, encoding, activation, retrieval, Central Limit Theorem.

Introduction

Romaniuk (2013) has provided empirical evidence that the Dirichlet theory, a widely accepted empirical generalisation typically used to predict patterns in buying behaviour, can be used to model market statistics based on consumer memory associations. More specifically, her work illustrates that the same sets of statistical distributions used for describing purchase timing and brand choice provide a robust prediction of the rates and shares of the brand associations held in memory by consumers tracked through surveys. Importantly, this novel empirical application of the Dirichlet model is grounded, in the first place, on a striking similarity between brand statistics derived from brand associations (e.g. a brand's mental market share, associative penetration and associative rate) and buying behaviour metrics (e.g. a brand's market share, purchase penetration and purchase frequency). It is also based on the assumption that consumer memory is stochastic in nature and that consumers store information pertaining brands as per the Associative Network Theory of Memory (ANT) of Anderson and Bower (1973), which describes memory as a system of interlinked concepts and information retrieval as a stimuli-based probabilistic process. Accordingly, the aim of applying the Dirichlet model to brand associations is to estimate a brand's retrieval propensities relative to competitors, also defined as *brand salience* or *mental availability* (Sharp 2010).

Whilst the accuracy of the predictions that the Dirichlet theory generates for brand memory associations documented by Romaniuk (2013) would suggest that the model suitably describes retrieval propensities for brands competing in a given category, the empirical results alone do not provide enough insights into the robustness of this theory in simulating the underlying cognitive processes that characterise information retrieval in ANT-like memory structures. In particular, the model is fitted to the discrete counts of brand associations to infer a brand's retrieval propensity by analysing (i) how many consumers could provide at least one association (i.e. how many consumers could effectively 'think of' the brand and retrieve it from memory); and (ii) how many associations each brand obtained relative to the other brands in the market. Consequently, this approach does not directly assess the impact on brand retrieval of: (i) chances of encoding, i.e. chances of establishing associations between brands and other concepts in memory; and (ii) the chances of reaching a sufficient level of neural activity ('activation') that can effectively trigger information retrieval.

In the psychological literature there are well-documented non-linear (stochastic) models that fully describe these processes and how the chances of information retrieval arise within ANT-like memory systems. In particular, psychological literature documents the existence of at least three distinct 'layers' of stochastic processes concurrently underpinning the chances of information retrieval: a first layer is represented by information *encoding* (i.e. establishing associations between concepts in memory); a second layer is represented by information *activation*; and a third and final layer is represented by *latency of retrieval*. Although the empirical results obtained when fitting the Dirichlet model to brand associations provide *a priori* support to the assumption that the Dirichlet distributions can approximate the combination of these three layers, there is the risk that the model could somewhat oversimplify the complex cognitive mechanisms that characterise information retrieval, hence leading to misleading predictions. This is of concern for those who wish to use the Dirichlet theory to model brand associations. Should this theory not provide a robust (albeit comprehensive) description of brand retrieval propensities, then further supplementary analytical approaches would be necessary to track and discern the distinct cognitive processes that underpin brand retrieval propensities.

In an attempt to address this issue, the aim of this research note is to determine the degree of *significant sameness* between the distributions forming the Dirichlet theory and those describing the stochastic processes that underpin information retrieval documented in the psychological literature. This is achieved through two steps. First, by reviewing in detail all three layers of stochastic processes that underpin retrieval in ANT-like memory structures. Second, by using the Central Limit Theorem to compare such distributions against the Dirichlet distributions.

First layer of stochastic processes in memory: information encoding

Anderson and Bower (1973) developed a theory regarding the process of encoding information into memory. This theory is an explicit stochastic model, which describes how input information is stored in more permanent areas of memory storage (*long-term memory*) and how it is subsequently recalled. The authors assert that a simple relationship between the time and input information determines the probability that the input is encoded into memory. This probability follows an exponential distribution with mean a (constant). Put simply, the chances of encoding are continuous and independent (i.e. following a Poisson stochastic process), with actual encoding of inputs occurring at a constant average rate over time. Therefore, $f(t)$ will be the probability density function of forming an association in a specific point in time, as per the following equation:

$$f(t) = \frac{1}{a} e^{-\frac{t}{a}} \qquad \textbf{\textit{Equation 1}}$$

$p(t)$ is the probability of forming an association by time t:

$$p(t) = 1 - e^{-\frac{t}{a}}$$ ***Equation 2***

When an input has *n* pre-existing associative links, the number of associations formed *(k)* is binomially distributed with the parameters $1 - e^{-\frac{t}{a}}$ and *n*. This means that the *n* links affect the chances of an association being formed. The mean number of associations formed will be equal to the following equation:

$$E(k) = n\left(1 - e^{-\frac{t}{a}}\right)$$ ***Equation 3***

Regarding the time necessary to encode *n* associations *T(n)*, the probability density function will depend upon: (i) intra-association intervals being exponentially distributed and independent; (ii) the sum of the inter-association intervals being gamma distributed.

In order to allow predicting the frequency at which a subject will form associations, the parameter *a* should be randomly distributed, as a mean value of sample heterogeneity, following an exponential distribution of the rate *1/a*, of mean *b*:

$$f\left(\frac{1}{a}\right) = \frac{1}{b} e^{-\frac{1}{ba}}$$ ***Equation 4***

Consequently, the probability density *f(a)* of *a* can be determined as a change of variable as follows:

$$f(a) = \frac{1}{ba^2} e^{-\frac{1}{ba}}$$ ***Equation 5***

With the expected value of the probability distribution *p(t)* of forming an association in *t* seconds being:

$$E[p(t)] = \frac{tb}{tb + 1}$$ ***Equation 6***

This formula is a special case of a beta distribution:

$$f(x) = \frac{\Gamma(\alpha + \beta)}{\Gamma(\alpha)\Gamma(\beta)} x^{\alpha - 1}(1 - x)^{\beta - 1}$$ ***Equation 7***

This means that the mean rate of encoding will increase when this distribution is close to 1 (as *tb* increases), i.e. the probability of forming an association will increase, and so will encoding time *t* and the mean encoding rate.

Besides predicting how associations are formed (encoded) in memory, the model also provides a tool for predicting recall patterns. More specifically, it theorises that the average recall pattern of any piece of information is determined by the probability of encoding a particular *k* of the *n* associations *Q (k,n)*:

$$Q(k,n) = \int_0^1 p(t)^k \left[1 - p(t)\right]^{n-k} f\left[p(t)\right] dp(t) \quad \textbf{\textit{Equation 8}}$$

Calculating this probability implies first determining the probability $p(t)^k \left[1 - p(t)\right]^{n-k}$ of forming a particular *k* association for each value of *p(t),* weighting such probability by the probability density function *f[p(t)]* at a certain *p(t)* value and integrating this over the 0 to 1 range of *f[p(t)]* as follows:

$$Q(k,n) = \frac{1}{tbn+1} \prod_{i=1}^{k} \frac{tbi}{tb(n-1)+1} \qquad \textbf{\textit{Equation 9}}$$

Whenever the mean rate of encoding *tb* will be close to 1, the mean probability to form an association will be 0.5 and the probability of forming *n* associations greater than 0.5^n.

Therefore, in summary, this stochastic model explains that there are empirically measurable patterns in the encoding and recall of information in memory and provides clear guidelines to anticipate the chances to recall a concept from long-term memory (as depending upon the formation of a number of associations in memory) (Anderson & Bower 1973). In particular, this model accurately captures the differences in potential activation and processing of information and therefore in the formation of ‘all-or-none’ discrete associations in memory, which are substantially equivalent to a memory association between brands and attributes in consumer memory. Importantly, the probability of these discrete connections being formed was proven to be independent within a memory network, but covarying across various networks (Anderson & Bower 1973). These characteristics of Anderson and Bower’s model match perfectly Romaniuk’s approach (2013), as the Dirichlet distributions are fitted to a set of brand statistics derived from ‘all-or-none’ discrete brand attribute associations collected through consumer surveys (whereby consumers are presented with a list of brands and attributes from a given product category and asked to provide as many associations as they wish).

Anderson and Bower’s model does not, however, explicitly include any description of the stochastic processes involved in information processing, i.e. the activation of concepts in memory and *working memory* performance, i.e. the performance of the more transient or non-lasting area of memory storage. These aspects were covered in later models documented in the psychological literature, which are reviewed next.

Second layer of stochastic processes in memory: information activation

The Active Control of Thought (ACT) theory formalised by Anderson (1976) describes the activation of information in memory and working memory performance. The central assumption of this theory is that the total amount of activation of a memory network of associated concepts is usually equal to the sum of the activation carried by all the components of *declarative* information stored within this network (whereby *declarative* implies ‘pertinent to any given cognitive task’, such as judgements and decision-making). This is calculated accounting for both *limited cognitive capacity* (i.e. only a limited amount of cognitive activity can occur at the same time in working memory) and heterogeneity of working memory (i.e. individuals do not support identical cognitive loads) (Anderson 1976).

Anderson (1993) has successively developed this theory into the ACT-R theory. The ACT-R theory assumes that each memory network of associated concepts has a base level of activation equal to zero and also considers the role of external cues, which will determine the probability of retrieval from long-term memory (thus the ‘R’ in the theory’s acronym). More specifically, it theorises that the activation A_i of any declarative knowledge component *i* of a memory network depends on the semantic similarities held with an external cue, as per the following equation:

$$A_i = \sum_i W_j S_{ji} \qquad \textbf{\textit{Equation 10}}$$

where W_j is the cognitive prominence of the declarative memory component *j* determined by the semantic similarity with the external cue and S_{ji} is the strength of the association of that component to the memory network of associated concepts. This approach formalises the activation of the declarative components’ influence on retrieval, given the compatibility with external cues. Moreover, W_j determines the probability of retrieval from long-term memory.

The model also explains that there is a constant amount of information that can be activated at the same time (known as *limited cognitive capacity*), and thus assumes the following:

$$\sum_j W_j = K \qquad \textbf{\textit{Equation 11}}$$

with *K* being a constant value, often assumed equal to 1 in empirical studies (Daily et al., 2001).

The model’s last key assumption is that activation patterns are often *non-linear*, due to some latency of activation within the network.

This theory has been the subject of numerous subsequent experiments and resulting variants. First, Reder, Park and Kieffaber’s (2009) most recent revision of the ACT-R theory takes into consideration the level of *familiarity* of concepts (i.e. frequency of activation or processing fluency) and the *decay of activation* over time. For instance, repeated exposure to a task implies that the declarative knowledge components pertinent to this task will be activated quite frequently, and thus strengthened (Reder et al. 2009). Over time, however, this strength will decrease, dropping off from the baseline of the last activation. This is explained by this formula (Reder et al. 2009):

$$B = B_w + cN\sum t_i^{\;-dN} \qquad \textbf{\textit{Equation 12}}$$

where B_w is the base level of activation, cN and dN are constants (dN stands for decay of activation) and t_i is the time passed since the last instance of activation ith.

Regarding the strength $S_{s,r}$ of the link between any two declarative knowledge components s and r pertinent to a specific task (Reder et al. 2009):

$$S_{s,r} = cL\sum t_i^{-dL} \qquad \textbf{\textit{Equation 13}}$$

t_i is the time passed since the last instance of association between two concepts; cL and dL are constants (dL stands for decay of link strength).

When a certain degree of '*fan effect*' (i.e. the activation 'loss' due to a number of competing declarative knowledge components) is also considered, then the incremental amount of activation above baseline that can be carried by a specific declarative knowledge component is (Reder et al. 2009):

$$\Delta A_r = \sum\left(\frac{A_S \times S_{s,r}}{\sum S_{s,I}}\right) \qquad \textbf{\textit{Equation 14}}$$

where ΔA_r is the amount of activation above baseline, A_s is the activation of the declarative component s, $S_{s,r}$ is the strength of the link between any two declarative knowledge components s and r, and lastly, $\sum_{s,I}$ is the sum of the strength of all declarative components linked to a specific task (Reder et al. 2009).

Prior to the work of Reder, Park and Kieffaber (2009), Lovett, Reder and Lebiere (1997) developed a version of the ACT-R model that is documented to provide more accurate predictions of information retrieval. More specifically, this earlier (and arguably less utilised) configuration of the ACT-R model presents two advantages. First, it explicitly accounts for the individual differences in working memory performance, i.e. it accounts for the fact that cognitive capacity does vary from individual to individual and people show different levels of cognitive performance. This was achieved by including in the ACT-R model an underlying parameter capturing heterogeneity in working memory performance. This parameter changes in a random fashion, thus follows a probability density distribution. More specifically, this parameter follows a density distribution that, when normalised, is centred at 1 (the original value of K). Second, besides including the heterogeneity parameter, this variation of the ACT-R model also provides a non-linear formula of the base-level activation B_i:

$$B_i = \log\left(\sum t_j^{\;-d}\right) \qquad \textbf{\textit{Equation 15}}$$

with t_j being the time lag since the last activation j^{th} and d being the decay rate.

These two modifications to the ACT-R model provide robust predictions of the odds of information retrieval (retrieval propensities), especially in any task-accomplishment that is based on cognitive processes (Lovett et al. 1997). As such, Lovett, Reder and Lebiere's (1997) version of the ACT-R theory is relevant to the application of the Dirichlet model to brand associations, since the ultimate goal of fitting the Dirichlet distributions to brand associations is to provide a robust prediction of a brand's retrieval propensities resulting from: (i) the network of brand associations held in memory by consumers determines; and more importantly (i) reaching a sufficient level of activation within the network.

Nonetheless, although non-linear functions such as Lovett, Reder and Lebiere's (1997) version of the ACT-R theory provide robust predictions of retrieval propensities, they do not cater for the fact that retrieval propensities are further affected by the following characteristic of ANT-like memory structures. That is, the output (retrieved) information is not directly proportional to the input (activated) information. Later models documented in the psychological literature predict this discrepancy between input and output information through non-linear models, and describe it as *latency of retrieval* and a final 'layer' of stochastic processes characterising retrieval propensities in ANT-like memory structures.

Third layer of stochastic processes in memory: latency of retrieval

Daily, Lovett and Reder's (2001) model explains the concept of *latency of retrieval*, as the amount of information above a certain threshold of activation (i.e. the difference between effectively retrieved information and activated information), which determines the odds of retrieval, and the decay of activation over time.

According to this model, given the total level of activation A_i of a memory network of associated concepts i, the probability of retrieving any information from the network is:

$$\Pr(retrieval) = \frac{1}{1+e^{-(A_i-\tau)/S}} \qquad \textbf{\textit{Equation 16}}$$

where τ is a constant threshold of activation necessary for retrieval and S is a constant measure of interference or 'noise'. If the total activation of a network of associated concepts is above τ, latency of retrieval will be the total amount of information activated above the threshold (Daily et al. 2001):

$$L(retrieval) = Fe^{-fA_i} \qquad \textbf{\textit{Equation 17}}$$

where F and f are constants, 'transforming' the total level of activation into actual retrieved information.

Put simply, latency of retrieval is the difference between the total level of activation A_i and the threshold τ, and A_i moderates the chances of retrieval. Therefore, solving the equations of this model, it is possible to estimate retrieval propensity. Higher values of W_i (salience of declarative knowledge components of the network i) also imply a higher A_i, which may impact speed and ease of retrieval. Accordingly, this model predicts that individuals with higher W_i are able to retrieve information more quickly and accurately than individuals with lower W_i. This does not mean that the relationship between W_i and the accomplishment of a task given a certain amount of activated declarative knowledge components is linear: small changes in W_i linearly affect A_i, but changes in A_i have exponential effects on retrieval.

Considering this model in relation to brand associations is important, since a brand's retrieval propensity will effectively depend on latency of retrieval as well, and should not be described by taking into account only encoding and activation chances.

The combination of the three layers of stochastic processes

From reviewing the psychological literature, it is clear that retrieval propensities are concurrently determined by: (i) the probability of establishing an association in memory, which follows a Beta Binomial Distribution – see Anderson and Bower (1973); (ii) the non-linear relationship with the total level of activation of a focal memory chunk, which results in retrieval chances being exponentially distributed – see Lovett, Reder and Lebiere (1997); and (iii) latency of retrieval (Daily et al. 2001). These are summarised, for simplicity, in Table 1.

Table 1. Stochastic memory theories and relative statistical distributions

	Phenomena	**Statistical distributions**
Forming associations in memory	The probability that a piece of information is encoded in memory	It follows exponential distribution with constant mean, which implies chances of encoding being continuous and independent (i.e. as per a Poisson stochastic process)
	The number of associations formed in memory (or total number of encoded pieces of information) given *n* pre-existing associated links	It is Binomially distributed
	The time needed for associations to be formed in memory	Inter-associations intervals are exponentially distributed and the sum of inter-associations intervals is gamma distributed
	Chance to form an association	It follows a beta distribution
Activating information	Activation patterns	They are non-linear, with as-if-random variation and exponential effects on information retrieval
Retrieving information	Information retrieval propensities	They depend on exceeding thresholds of activation patterns, thus are exponentially distributed

These three processes generate three sub-sets of latent variables that concurrently shape retrieval propensities, which have been conceptualised as being *as-if-random* independent variables by Anderson and Bower (1973). Accordingly, the sub-sets of latent variables created by each layer must be multiplicative and positive, since they are concurrently shaping another set of variables. Thus, the degree of variance shown in these multiplicative and positive variables should display *geometric means*.

In probability theory the Central Limit Theorem (CLT) states that when a group of random variables show geometric means, a Gaussian distribution is a valid mathematic approximation for these variables (Johnson et al. 1994). As exemplified in Figure 1, since this distribution is technically close to a gamma distribution, a key component of the Dirichlet model, this confirms that the combination of these distributions is very similar in nature (if not *technically* identical) to those included in the Dirichlet model.

For these reasons, it is plausible to conclude that the Dirichlet theory is suitable not only to describe and predict retrieval propensities, but also to approximate the various underlying stochastic processes that underpin retrieval: from the establishment of associations in memory (encoding), to the subsequent activation of relevant information and the actual act of 'bringing back to mind'. This has important implications for the use of the Dirichlet theory in relation to consumer memory, as discussed next.

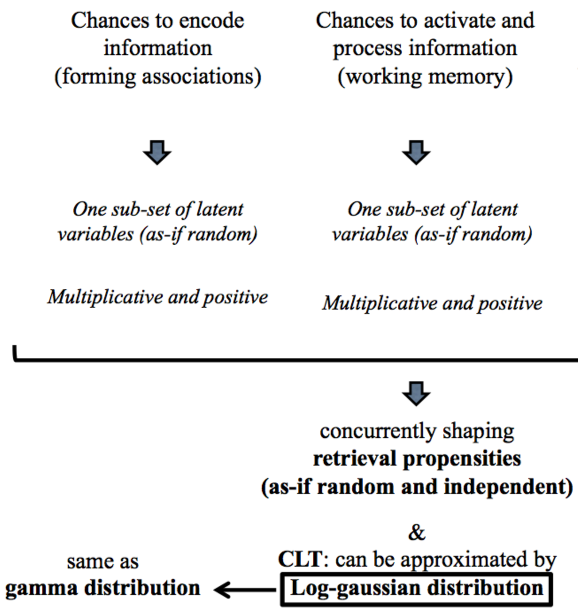

Figure 1. Memory being 'Dirichlet-like'

So what…

One may question the value of digging into psychological literature in order to unpack the details of multiple stochastic models describing consumer memory to back up a key analytical assumption that is supported *a priori* by empirical evidence. Wright and Kearns (1998) provide a simple answer to such a question. Whilst Romaniuk's application of a fundamental empirical generalisation such as the Dirichlet model to consumer memory is surely robust and fascinating, the aim of this research note was 'to determine which type of understanding produces the best results' (Wright and Kearns 1998, p.4). Accordingly, by 'baking-off' the Dirichlet theory application (and thereby the empirical evidence offered) and the evidence offered by psychological research, this research note has confirmed that the Dirichlet model is suitable and possibly more parsimonious for describing brand retrieval (as the model encompasses a single stochastic process shaping retrieval propensities, thus somewhat 'summarises' the three underpinning processes that are theorised separately by psychological research).

The scope of empirical generalisations in marketing is to provide a comprehensive description of complex phenomena, enabling managers to understand and track relevant outcomes of marketing practice (Ehrenberg 1995). To this extent, the Dirichlet theory is a vital diagnostic and prescriptive marketing tool. Therefore, confirming that the application of this model to a crucial aspect of consumer behaviour such as brand memory associations is robust is very important to marketing practice. More specifically, the present research has confirmed that the Dirichlet model is a suitable way of approximating an otherwise complex and unobservable set of cognitive processes that characterise the way the consumer memorises and retrieves brands. Therefore, it is now possible to comfortably draw further and more detailed inferences on consumer memory. For instance, besides predicting market statistics, it is now possible to analyse another key output of the Dirichlet model: the parameters of the distributions forming the theory.

In stochastic models such as the Dirichlet theory the parameters of the distributions that form the model explain some relevant aspects of the phenomenon described by the theory itself (Sharp et al. 2012). Empirical research on the Dirichlet parameters tends to be focused on three key parameters that carry the most explicit insight about purchase behaviour within a given product category. These are the parameters *M*, *K* and *Phi*. The *M* parameter describes the overall weight of category purchases by all shoppers; typically reflecting seasonality, promotions, stock outs and other demand fluctuations. The *K* parameter is the shape parameter of the gamma distribution underlying the Negative Binomial Distribution and a measure of heterogeneity in the category latent selection rates. As such, it expresses behavioural characteristics of a group of shoppers and the distribution of light and heavy buyers (lower *K* values imply a greater proportion of heavy buyers) or customer concentration. The parameter *Phi* is a transformation of the traditional *S* parameter, $1/(1+S)$, and measures the variance in shoppers' choice probabilities, hence capturing the degree of switching or brand loyalty (a *Phi* of 0 represents no loyalty, while a *Phi* of 1 represents 100% loyalty) and distinguishing between repertoire and subscription markets.

By confirming that the Dirichlet model can robustly depict retrieval propensities and the various sub-layers of cognitive processes that underpin brand salience in ANT-like brand memory structures, it is now possible to advance some more specific hypotheses on the meaning of these parameters in relation to consumer memory for a given product category. More specifically, it is plausible to interpret *M*, *K* and *Phi* as follows (see Table 2).

Accordingly, it is possible to draw the following implications.

- *M* will create a 'ceiling effect' or constrain to each brand's retrieval propensities;
- *K* will effectively provide a guideline to manage (i.e. increase) brand retrieval propensities across different segments of consumers (ideally, just like purchases can be 'nudged' by marginally increasing purchase propensities of non-buyers and light

buyers, a brand's retrieval propensity would increase through 'nudging' cognitive prominence across consumers who cannot retrieve the brand from memory or provided, on average, only very few associations in the survey);
- *Phi* will effectively provide a valuable competitive benchmark helpful to assess the degree of competition for retrieval in the market: the closest to 0, the toughest the competition for retrieval in consumer memory.

Table 2. The meaning of the Dirichlet parameters for brand associations

	Purchases	Brand associations
M	Total level of demand	Total level of associations or 'category knowledge' across all brands held by the group of consumers surveyed.
K	Customer concentration	Proportion of 'high-knowledgeable' customers (i.e. customers who hold large networks of associations across all brands in the product category) in the group of consumers surveyed.
Phi	Loyalty (switching)	How scattered/concentrated the total level of associations or 'category knowledge' is (one brand vs. many brands) for the group of consumers surveyed.

Much more is also possible. For instance, it is now also possible to interpret deviations and exceptions from the Dirichlet norms as a direct reflection of retrieval inefficiencies that can emerge at the encoding, activation or retrieving stage of information processing. Also, it is now feasible to carry out explicit comparisons between consumer memory and consumer buying behaviour by looking into configurations of the Dirichlet model that account for covariates such as Rungie et al.'s approach (2007).

References

Anderson, JR (1976). *Language, memory and thought*, Hillsdale ,NJ, Erlbaum.

Anderson, JR (1993). *Rules of the Mind*, New Jersey, Lawrence Erlbaum Associates Inc.

Anderson, JR & Bower, GH (1973). *Human Associative Memory*, Washington, DC, Hemisphere Publishing Corporation.

Daily, L, Lovett, M & Reder, L (2001). Modeling individual differences in working memory performance: A source activation account. *Cognitive Science*, 25, 315-353.

Ehrenberg, ASC (1995). Empirical Generalisations, Theory, and Method. *Marketing Science*, 14, G20-G28.

Johnson, NL, Kotz, S & Balakrishnan, N (1994). Continuous Univariate Distributions, New York, John Wiley and Sons.

Lovett, M, Reder, L & Lebiere, C (1997). Modeling individual differences in a digit working memory task. Lawrence Erlbaum Associates, 460.

Reder, L, Park, H & Kieffaber, P (2009). Memory systems do not divide on consciousness: reinterpreting memory in terms of activation and binding. *Psychological Bulletin*, 135, 23-49.

Romaniuk, J (2013). Modeling mental market share. *Journal of Business Research*, 66.

Rungie, C. M., Laurent, G. & Uncles, M. 2007. The generalized Dirichlet model: A comprehensive model of repeated purchase with covariates. Adelaide: University of South Australia.

Sharp, B (2010). How Brands Grow, South Melbourne, Oxford University Press.

Sharp, B, Wright, M, Dawes, J, Driesener, C, Meyer-Waarden, L, Stocchi, L & Stern, P (2012). It's a Dirichlet World: Modeling individuals' Loyalties reveals How Brands Compete, Grow, and Decline. *Journal of Advertising Research*, 52, 203-213.

Wright, M & Kearns, Z (1998). Progress in marketing knowledge. *Journal of Empirical Generalisations in Marketing Science*, 3, 1-21.

Dr. Lara Stocchi is a Lecturer in Marketing, School of Business and Economics, Loughborough University