Towards Evidence-based Marketing:
The Case of Childhood Obesity

Martine Stead, Deputy Director
Laura McDermott, Research Officer
Gerard Hastings, Director

Institute for Social Marketing
University of Stirling
Stirling
FK9 4LA

1. Martine Stead - Tel: 01786 467387, Email: martine.stead@stir.ac.uk
2. Laura McDermott - CONTACT AUTHOR - Tel: 01786 466456, Email: laura.mcdermott@stir.ac.uk
3. Gerard Hastings - Tel: 01786 467393, Email: gerard.hastings@stir.ac.uk
Fax: 01786 466449
Towards Evidence-based Marketing: The Case of Childhood Obesity

KEY WORDS

Systematic review, food promotion, children, policymaking
ABSTRACT

Contentious commodities such as tobacco, alcohol and fatty foods are bringing marketing under scrutiny from consumers and policymakers. Yet there is little agreement on whether marketing is harmful to society. Systematic review (SR), a methodology derived from clinical medicine, offers marketers a tool for providing resolution and allowing policymakers to proceed with greater confidence. This paper describes how SR methods were applied for the first time to a marketing problem – the effects of food promotion to children. The review withstood scrutiny and its findings were formally ratified by government bodies and policymakers, demonstrating that SR methods can transfer from clinical research to marketing.
PROBLEMS OF EVIDENCE

From the earliest days of the discipline, marketers have been interested in the impact that their ideas and activities might be having beyond the firm, on society more generally. Wilkie & Moore’s (2003) extensive review of the origins of marketing thought shows how ‘marketing and society’ has been a recurring theme; as long ago as the 1920s and 30s concerns about such issues as unfair pricing, pushy salesmanship and emotional advertising encouraged the development of a consumer movement. Over time, debate about these and other controversial marketing practices spread from the public domain to the marketing literature, featuring in leading marketing publications and core textbooks (Andreasen, 1997; Greyser, 1997). In turn, these outlets welcomed commentary on marketing’s wider impact on society (Wilkie & Moore, 2003).

This interest has developed a sharper focus in the last decade, with increasing concerns being expressed – typically from outside the discipline - about the impact of marketing on the consumption of contentious commodities such as tobacco, alcohol and, latterly, energy dense foods (eg. Ellickson et al, 2005; Halford et al, 2004; Pierce et al, 2002; Pollay, 2000; Seiders & Petty, 2004). The World Health Organization, for example, has coined the phrase ‘hazard merchants’ to describe the marketers of such products (eg. World Health Organization, 1999). In the case of tobacco this rhetoric has been matched with muscular action: the recently agreed Framework Convention on Tobacco Control is the first international treaty with the specific aim of curtailing marketing activity (World Health Organization, 2003); to date 168 countries have signed the Convention, and 63 proceeded to full ratification (Framework Convention Alliance, 2005). Some thirty countries worldwide have now
instituted severe restrictions on tobacco marketing and, in the US at least, litigation against tobacco marketers has become commonplace (Hurt & Robertson, 1998).

However, even in the field of tobacco these controls have been slow to emerge. In the UK, for example, restrictions on tobacco advertising were debated for over twenty years before they reached the statute in 2003 (eg. Hastings et al, 1993; Hastings et al, 1994; McDonald et al, 1993). This delay was caused by a combination of vested interest and dubiety in the evidence base. For each study showing that advertising did influence children to smoke, for example (eg. Aitken et al, 1991; Alexander et al, 1983; Goddard, 1990), another could be produced showing it did not (eg. Mizerski, 1995). Time has shown this to be a spurious debate, and that the tobacco industry exploited it with both energy and success (Bitton et al, 2002; Gilmore & McKee, in press; Neuman et al, 2002). The delay brought real and very considerable social costs: when the advertising ban was finally introduced, the UK Minister of Health argued that it would save 3000 lives a year (Milburn, 2001); arguably therefore, every year of delay had cost the same number of lives.

The precision of this calculation underlines the policy maker’s desire to proceed on a sound scientific basis. Where policy-relevant research findings can influence decisions “involving millions of people and billions of dollars” (Franke, 2001), assessing the robustness of those findings is crucial. Evidence-based public policy has become a clarion call for reasons of both accountability and legality - decisions have to be justified to both the electorate and the courts (eg. Robinson et al, 2005). Voters will not respond well to having their freedom limited by, for example, seat belt legislation, unless there is convincing evidence to show that benefits will result. For evidence to be convincing, a consensus among experts is crucial;
recent experiences in the UK of childhood immunisation have shown just how damaging even a tiny number of dissenting voices can be (Heller et al, 2001, Wroe et al, 2005).

Similarly, corporations will not willingly accept restrictions on their activities if they feel the evidence base does not support them, as witnessed by two current law cases – one in Ireland, the other the UK – brought by tobacco companies to challenge restrictions on point of purchase marketing (PJ Carroll & Co Ltd and others v Minister for Health & Children and others; British American Tobacco UK Ltd and others v The Secretary of State for Health, 2004). Solesbury (2001) underlines the point when he argues that the standard of evidence required in political policymaking circles is now on a par with that required for criminal convictions. Again, consensus built on a rigorous evidence base would ease this problem.

Greater clarity and consensus about the impact of business on society would also have benefits for corporations themselves. Uncertainty about policy decisions and the possibility of legal threats can affect share prices as well as election results. Strategic planning is also much easier in a predictable and consistent business environment (Wilson & Gilligan, 1998). Building brands, consumer loyalty and stakeholder relationships is more difficult if key arms of marketing, such as advertising or sales promotion, are under constant threat or the reputation of business – or a particular business sector – is being undermined. If marketing is to make a more sophisticated and useful contribution to the debate about its impact on society, a more objective and explicit process for assessing this impact is needed. In the arenas of social welfare and public health it could move from reactively defending itself against accusations of being part of the problem, to proactively contributing to the solution.
In short, when business and social interests appear to be in conflict, everyone – citizens, policy makers and the business community - would benefit from a more rigorous, transparent and consensual way of measuring the impact of marketing on society.

**THE RISE OF EVIDENCE-BASED DECISION-MAKING**

Medicine has faced similar dilemmas. Doctors have to make decisions about what works and what side effects are acceptable; they also have to liaise with policy makers to manage their interface with society. The concept of evidence-based decision-making, with its emphasis on rigorous methodologies for sifting, prioritising and interpreting evidence, has arisen in response to this need (Mulrow, 1994). It is defined as "the conscientious, explicit, and judicious use of current best evidence in making decisions about the care of individual patients” (Sackett et al, 1996).

In the UK this has had profound effects on practice. During the early 1990s, local health authorities were required to justify investment in medical interventions on the basis of both effectiveness and cost-effectiveness (Harrison, 2002). To support this kind of decision-making, a formal infrastructure for producing and disseminating effectiveness research to the National Health Service was developed. Most notably, the Cochrane Centre at the University of Oxford [http://www.cochrane.co.uk] was established to facilitate the preparation and maintenance of authoritative reviews of the best evidence (Light, 2003). Over the past ten years, the Centre has produced an impressive body of literature extending across a range of health topics (Boaz et al, 2002), from the treatment of depression following stroke (Hackett et

At the heart of this process lies the ‘systematic review’ (SR) (Boaz et al, 2002). This is ‘a review of the evidence on a clearly formulated question that uses systematic and explicit methods to identify, select and critically appraise relevant primary research’ (Khan et al, 2001). It is a method for identifying and synthesising the findings from all relevant studies on a given topic. Where a traditional literature review may be limited or skewed by reliance on a narrow pool of evidence, or by a reviewer’s natural tendency to favour some studies over others, a systematic review makes explicit the criteria by which studies were found and selected (Petticrew, 2001). Thus, whilst the concept of review is not new; systematic review is different in that it demands adherence to agreed standards and the adoption of transparent and replicable procedures (Boaz et al, 2002). Because systematic review is rigorous and less prone to bias than a traditional literature review, it provides a quality-controlled overview of all the existing research on which to inform public policy debates (Baldwin et al, 2002).

A detailed ‘protocol’ for carrying out the review is developed and agreed in advance. This specifies four fundamental procedures: (i) the research questions and objectives to be addressed; as with any sound research, this ensures that the objectives rather than the data drive the review process (CRD, 2001); (ii) explicit inclusion / exclusion criteria for the subject matter of the papers to be covered by the review (Khan et al, 2001); (iii) the search strategy. This is usually designed to identify both published and unpublished research, and specifies precise details of the sources and databases to be consulted, along with a list of the search terms to be used. It is typically very broad; as the aim in SR is to cast the net as widely as possible to improve the chances of capturing all relevant studies; and (iv) quality control.
Systematic reviews critically appraise the methodological quality of research (or reviews of research) to ensure that only the best work is included (Light, 2003). Methodological criteria are specified (for example, ‘only randomised controlled experimental studies will be included’), and studies which are judged to fall below this level are excluded.

The protocol is a formal document and so can be peer reviewed to maximise rigour. It also forms a blueprint against which the fidelity and findings of the final review can be checked, both by the review team and anyone else who wishes to scrutinise their work. Throughout the review process a detailed log is kept of every search undertaken, including the date it was conducted, the search term or phrase used, the search field in which the term or phrase was used, any applied limits and the number of hits generated by each search. This log can also be scrutinised, criticised and even replicated by independent third parties.

This means that any criticisms have to be precise and explicit. Any rival reviews that come to different conclusions have to be able demonstrate precisely how the methods they used differ from - and improve upon - those used in the original review. And independent third parties can do the same analysis.

In this way both the protocol and the review process deliver what is perhaps the most valuable single quality of SR: transparency.
FROM MEDICINE TO MARKETING

Systematic review is now well established in the UK, and there are increasing signs of interest in the concept of evidence-based decision-making further afield (Solesbury, 2001). For example, the recent EC White paper on Governance proposes that the legislative decision-making process becomes more transparent and evidence-based (European Commission, 2001).

However, the broader field of public policy is more nebulous than clinical medicine. Comparing alternate treatments for a heart condition lends itself to clear cut and tightly controlled experimental studies. The randomised controlled trial is commonly viewed as the ‘gold standard’ in research design and this in turn makes SR a more straightforward process. Measuring the impact of complex health policy interventions which operate in real world settings is on the other hand a much messier challenge (McKinlay, 1993; Stead et al, 2002; Tones, 2000). Furthermore, policy decision-making is a complex process, often driven by factors other than the nature of the evidence itself. Politics often play an important role, and tensions exist between knowledge and power in the shaping of policy (Solesbury, 2001).

Despite these challenges, the principles of evidence based decision making and SR are increasingly being taken beyond medicine, into fields such as crime and criminal justice, social welfare and health education, where practice and policy options are also actively debated and the balance between professional and public interest has to be determined. For example, a recent ESRC-funded study used SR to assess the effectiveness of financial ‘safety net’ products designed to protect mortgagors against the risk of arrears and repossession (Baldwin et al, 2002). As with medical research, institutions have been established to guide
and progress the preparation and dissemination of research used to guide social policy. In 2002, the Campbell Collaboration was established to build on the experience of the Cochrane Collaboration and carry out reviews of interventions in the fields of education, criminal justice and social work [http://www.campbellcollaboration.org]. Similarly, the Centre for Evidence Based Policy and Practice [http://www.evidencenetwork.org] was established to produce and disseminate social science research and publications relevant to policy and practice among a variety of stakeholders including the research community, central and local government, and industry. Reviews now exist on a disparate range of topics including teaching approaches (Higgins et al, 2005), social care (Turner et al, 2005), and crime reduction (Petrosino et al, 2002). In the UK this growing interest in evidence based decision making is reflected at the highest levels of Government (Cabinet Office, 1999), where commitments have been made to use research evidence to inform policy and practice and great emphasis is put on establishing ‘what works’ (Boaz et al, 2002) to ensure efficient use of taxpayers’ money.

Similar thinking can be usefully applied to marketing, especially where there is the potential for conflict between business and public interest. In many areas of marketing, the process of evaluating marketing’s potential harmful impact on society is challenging and complex (Polonsky et al, 2003). Previous efforts to navigate through contested claims and evidence have been hampered by a tendency to draw selectively on research conducted by key interest groups such as industry associations or lobby groups (eg. Maubach & Hoek, 2005). A key site of particularly topical conflict is the debate concerning the promotion of food to children, and its possible contribution to rising levels of childhood obesity in the USA, Europe and the UK. Scrutiny is increasingly being turned towards the role of food advertising by international advisory bodies such as the World Health Organization (Hawkes, 2004; World
Health Organization, 2004), governments (eg. Department of Health, 2004; National Institutes of Health, 2004; Zinn, 2003), special interest groups (eg. The Kasier Family Foundation, 2004; Sustain, 2005; The Coalition on Food Advertising to Children, 2003) and academic commentators (eg. Eagle et al, 2004; Nestle & Jacobson, 2000). However, the argument that food advertising influences children’s diet is hotly contested by industry and other commentators who claim that “there is no serious and methodologically sound evidence that shows that food advertising leads to an increase in the consumption by children of whole categories of foods” (Young, 2003 as cited in Livingstone, 2005).

For policymakers seeking to decide appropriate responses - for example, to intervene to restrict food promotion, to increase promotion of healthier foods as a counterbalance, to opt for clearer food product labelling, to increase children’s ‘advertising literacy’, or to do nothing - it is crucial that a way is found through these contesting claims for the evidence (Livingstone, 2005). This paper demonstrates the potential of SR as a methodological tool for providing resolution and allowing policymakers to proceed with greater confidence. As well as informing practice surrounding the use of marketing, SR poses great potential to academics who can use its rigorous and transparent methods to develop and advance our understanding of key marketing concepts and theories. The SR approach could be used, for example, to synthesise and make sense of disparate evidence on the effectiveness of different segmentation or pricing strategies and approaches.
SYSTEMATIC REVIEWS IN MARKETING: THE CASE OF FOOD PROMOTION TO CHILDREN

Background to the Problem

In the UK, as in the US, great concern has been expressed about the increasing problem of obesity and overweight, especially among children (Department of Health, 2003). The proportion of overweight children aged between 6 and 15 years increased by 7% between 1996 and 2001 (Department of Health, 2003), and levels of obesity rose by 3.5%. Obesity is of grave concern given the significant risks it poses for the long-term physical and mental health of young people. In terms of physical health, obesity is associated with a range of chronic diseases such as coronary heart disease, Type II diabetes, and cancer (National Audit Office, 2001). Its consequences can even extend to social and psychological problems. For example, obese children or teenagers may be subject to discrimination or prejudice, or may suffer from low self esteem (POST, 2003).

This phenomenon is known to be multi-factoral, with both diet and level of activity contributing to the problem. One particularly contentious issue on the diet side of the debate is the role, if any, that food advertising and promotion has on consumption, especially by children. To resolve this, in 2002 the Food Standards Agency (the UK Government body charged with issues of food safety and nutrition) commissioned a review of the existing research evidence on:

- the extent and nature of food promotion to children
the effect, if any, that this promotion has on their food knowledge, preferences and behaviour.

Because of the controversy surrounding the topic it was crucial that the review was as rigorous and transparent as possible. A team of researchers from four leading UK universities\(^1\) was selected through a peer review process to carry out a systematic review, which took nearly 18 months to complete (Hastings et al, 2003).

**Review Methods**

Systematic review methods were used to ensure replicability and transparency. The generic framework for undertaking systematic reviews was adapted for this marketing review. The key stages in the review process are shown in Figure 1. First, a preliminary literature search and analysis was undertaken to provide information on the potential nature, size and quality of the evidence base and to aid the development of the research questions and review methods (Stage 1). Once the methods for the review were developed and refined (Stage 2), searching for literature began (Stage 3). The search strategy was broad and comprised four methods: (i) searches of electronic databases (ii) searches of grey literature (iii) personal contact and (iv) examination of the reference lists of key studies. These searches yielded 29946 potentially relevant titles and abstracts. There was a huge amount of overlap in the results. This was, in part, attributed to similarities between different searches undertaken within each database and an overlap between the databases themselves.
A set of relevance criteria were then developed to help filter the titles and abstracts (Stage 4). Primary research studies or reviews published in English since 1970 were eligible for inclusion. Studies had to address directly the extent and nature of food promotion to children, and/or its effects on their food knowledge, preferences and behaviour. Where mentioned, the terms ‘children’ ‘food’ ‘promotion’, and ‘knowledge’, ‘preferences’ and ‘behaviour’ had to correspond with agreed definitions developed specifically for the purpose of the review. For example, ‘food’ was defined as all foods and non-alcoholic drinks, and ‘promotion’ was defined as any form of commercial promotion including advertising, branding, packaging, merchandising and in-school marketing. From this, a total of 201 articles were considered relevant to the review and the full text of these papers were retrieved (Stage 5).

These 201 articles were then further assessed using more stringent relevance and quality criteria (Stage 6). For example, in terms of methodological quality, articles had to provide information about sample design, data collection methods, and data analysis procedures. At this early stage of assessment, all types of sample design (eg. purposive, quota and convenience) and study design (including experiments, surveys, observation and qualitative methods) were permitted providing that they were clearly described. This process reduced the 201 articles to 50 studies describing the extent and nature of food promotion to children and 32 providing evidence of its effects on their knowledge, attitudes, and behaviour. Data extraction forms (which provided a full but concise description of each study in terms of design, sample, methods and procedures, analysis and results) were then completed for all included studies (Stage 7). Data were extracted by one reviewer, using a pro forma, and checked by a second reviewer.
The included studies were then subject to a final quality rating to gauge their relative quality; this was used to help assess which studies’ findings should be given more weight in drawing conclusions from the evidence (Stage 8). Studies were categorised, on the basis of their rating scores, as high, medium or low quality. For studies examining the extent and nature of food promotion to children, quality criteria included: the sample size, diversity and timing, thoroughness of the analysis, and the clarity and completeness of the data reporting. For studies examining the effects of food promotion to children, quality criteria included: the quality of the exposure measure, the quality of the effect(s) measure(s), the appropriateness of the analysis procedures, the extent and thoroughness of the analysis, and the clarity and completeness of data reporting.

The heterogeneity of the studies in terms of exposure type, subjects, settings and outcomes was too great to permit data synthesis by meta-analysis. A qualitative narrative synthesis, a common technique for systematic reviews where the evidence base is strongly heterogeneous, was therefore conducted (Stage 9).

Throughout the course of the research, every effort was made to maintain the objectivity of the review. As well as adhering to systematic procedures, the work was scrutinised by continuous peer review, from the initial proposal to do the work, during the review process itself right through to report drafting and publication. In total some forty different academics, from a variety of institutions and disciplines, refereed the project, or some stage of the project. In addition, an independent advisory panel, comprising representatives from industry, public health and academia, provided regular guidance and scrutiny.

TAKE IN FIGURE 1
Review Findings

The Extent and Nature of Food Promotion to Children

Fifty studies assessed the extent and nature of food promotion to children. Most of the studies were North American and used content analyses methods. The findings of these studies are summarised in Figure 2.

TAKE IN FIGURE 2

In short, the review found that television advertising is by far the most frequently used medium to promote foods to children and that the advertised diet is inherently unhealthy, dominated by foods high in sugar, fats and salts. Themes of fun and fantasy or taste, rather than health and nutrition are used to promote foods to children.

Effects of Food Promotion on Children’s Food Knowledge, Preferences and Behaviour

Thirty-two studies examined potential causal links between food promotion and children’s food knowledge, preferences and/or behaviour. Figure 3 provides a brief overview of each study’s characteristics and summarises their findings in terms of five main outcomes: nutritional knowledge, food preferences, food purchasing and purchase-related behaviour, food consumption, and diet and health.

TAKE IN FIGURE 3

(i) Nutritional Knowledge
Eight studies provided modest evidence that exposure to food promotion can influence nutritional knowledge. Of the eight studies, four found that exposure to food promotion had a significant impact on or was associated with significant changes in children’s nutritional knowledge and perceptions, usually in the direction of greater inaccuracy or increased confusion. Three studies found that exposure to food promotion had no significant impact on or was not associated with significant changes in children’s nutritional knowledge and perceptions. In the eighth study, the findings were inconclusive.

(ii) Food Preferences

Fourteen studies investigated whether food promotion influenced children’s food preferences and found strong evidence of an effect in this domain. Two studies measured but did not report data on the effect of food promotion on degree of liking for foods (Jeffrey et al, 1982 Study 1; Jeffrey et al, 1982 Study 2/Fox, 1981). Of the twelve studies that did report results, seven found that exposure to food promotion had an impact on, or was associated with significant changes in, children’s food preferences in the direction of the advertised foods. Usually these were foods high in salt, sugar or fat, but where the advertised foods were healthy, effects in the desired direction were also found (Norton et al, 2000).

(iii) Food Purchasing and Purchase-related Behaviour

Seven studies examined the impact of food promotion on children’s food purchasing and purchase-related behaviour (eg. purchase influence behaviour or ‘pester power’). The studies took different measures of purchasing and purchase-related behaviour. One study used actual sales of snacks from a school vending machine as a measure of purchase behaviour (French et al, 2001), while another relied on self-reported purchase of specific cereal brands (Goldberg et al, 1990). When assessing purchase influence behaviour, three studies actually
observed children’s behaviour in the supermarket (Galst & White, 1976; Reeves & Atkin, 1979; Stoneman & Brody, 1982), while others relied on either mothers’ (Taras et al, 1989) or children’s (Atkin, 1975) reports of purchase influence attempts. Overall, the SR found strong evidence that food promotion influences children’s food purchasing and purchase-related behaviour. Both the methodologically stronger and less strong studies found evidence of effects. In all except one study, the effect was in the direction of increasing purchase requests for foods high in fat, sugar or salt.

**Food Consumption Behaviour**

Eleven studies found modest evidence of an effect on food consumption behaviour. Effects were sometimes inconsistent and were not found in all the studies, but were found in sufficient studies to suggest that food promotion can, in some contexts, influence children’s food consumption behaviour. For example, in one study, food promotion reduced children’s likelihood of selecting fruit or orange juice, compared to a sweet, for a daily snack (Gorn & Goldberg, 1982/Gorn & Goldberg, 1980b), and in another it increased boys’ calorific consumption from a tray of snack foods (Jeffrey et al, 1982 Study 2/Fox, 1981).

**Diet and Health**

Six studies investigated the effects of food promotion on diet and health. Overall, there were small but significant associations between television viewing and diet, television viewing and obesity, and television viewing and cholesterol. For example, Bolton (1983), a strong study, found that the greater a child’s food advertising exposure, the more frequent his or her snacking and the lower his or her nutrient efficiency. Dietz and Gortmaker (1985) found a significant relationship between television viewing and obesity, and Wong et al (1992) found
a significant relationship between television viewing/video game playing and high cholesterol.

THE CHALLENGES OF SYSTEMATIC REVIEWS IN MARKETING

Conducting the review was also immensely challenging; bio-medical methodologies do not transfer easily to social science. Two issues in particular were challenging:

(i) *Keeping the Evidence in Context*

Adherence to the SR protocol, with its a priori tightly defined research questions and search parameters, has the advantage of establishing a clear focus and minimising bias. However the protocol’s specific focus excludes consideration of potentially interesting data which are not capable of directly answering the review questions but might help provide an intellectual fix on the issue. For example, unlike Young (2003), we could not set the findings in a wider context by examining studies that assessed children’s ability to discriminate advertising from programmes, because ability to discriminate programming from promotion is not directly relevant to the question of whether promotion impacts on knowledge, attitudes and behaviour. This type of inevitable omission opens a SR up to accusations of neglecting the bigger picture. In this sense there is a trade off between precision and context (McDonald, 2003).

(ii) *Deciding Which Types of Evidence are ‘Better’*

A ‘hierarchy of evidence’ has been developed by systematic reviewers working in clinical medicine to assist in the classification and prioritisation of studies (see Figure 4).
TAKE IN FIGURE 4

Underlying this is the principle of validity and the elimination of bias. At the pinnacle of the hierarchy is the meta-analysis of several randomised controlled trials (RCTs) (Jones, 2002). A systematic review or meta-analysis of RCTs uses complex statistical procedures to combine the data from a number of studies into a single numerical estimate (Light, 2003). Below the meta-analysis is the individual RCT, and below this are studies using other experimental designs such as cohort and case control studies. At the bottom of the hierarchy are ‘cross sectional studies’, ‘case series’ and ‘case reports’.

In all SRs, the decision about the cut-off point is both crucial and difficult. If the cut-off point is set too low the results risk being skewed by the inclusion of methodologically flawed studies with limited reliability. Equally there are risks in setting the evidential threshold too high, if this results in such a small pool of studies being included that few conclusions can be drawn (McDonald, 2003). An overly pragmatic approach can lead to unnecessary gaps in knowledge, and telling policymakers that ‘we know nothing’ is a dangerous strategy. We must make the most of the evidence that we do have and recognise the distinction between “not having (experimental) evidence of an effect” and concluding that no (experimental) evidence means that “there is therefore no effect” (adapted from McDonald, 2003).

The challenges posed by the hierarchy of evidence concept are even more marked when the SR methodology is used in a social science rather than clinical medicine context. In our review, there were no existing meta-analyses, and indeed no systematic reviews - the ‘highest’ forms of evidence - on food promotion’s effects on children. Although we did find
19 experimental studies using comparison or control groups, sometimes with random allocation, none applied all three of the common features of the RCT design: random allocation, control group and double-blinding. This left them open to criticisms of imperfect methodology, “External validity has to do with the extent to which results from an experiment can be generalised beyond the specific, and limited, circumstances of the experiment and into the real world. It is on this count that Goldberg’s work is suspect” (Paliwoda & Crawford, 2003, p17). However, there are often valid reasons why social science experiments cannot attain the methodological purity of clinical trials – as Livingstone (2005) points out, it is more difficult in a naturalistic setting to eliminate possible confounding variables. There are also obvious ethical difficulties in exposing children to potentially harmful stimuli, such as adverts for sugary foods, over the long term. In social science, the ‘perfect study’ often simply cannot exist, for technical, ethical or other reasons. Continually stating that no conclusions can be drawn because the perfect study has not yet been conducted defers decision making and means that no progress can be made; at some point policy decisions have to be made on the basis of the evidence which already exists (Livingstone, 2005; McDonald, 2003).

Another related decision posed by the hierarchy of evidence concept is whether to restrict the review to studies which are as homogenous as possible. Homogeneity potentially increases the internal validity of the pool of included studies, and synthesis of the evidence is facilitated by the fact that the similarity of outcome measures and methods permits the use of meta-analytic techniques to compare effect sizes, thus increasing the robustness of the review. However, limiting the review only to homogenous studies such as controlled experiments also opens it to criticisms of a lack of external validity (eg. Food Standards Agency, 2003a; Paliwoda & Crawford, 2003). This is because experimental conditions bear little relation to
how advertising is consumed in the real world – phenomena such as advertising clutter, repetition and saturation are very difficult to control or measure in an experiment. This calls for cross-sectional and observational studies to provide a more naturalistic view. But such studies come at the bottom of the bio-medical hierarchy. By design they are limited in terms of what inferences can be drawn from them. They can suggest associations between an exposure and an effect, but they cannot prove causality. Partial correlations which control for confounding variables help to establish the relationship with more confidence, but the question of causal direction remains problematic.

Livingstone (2005) argues that in complex fields such as social policy both sorts of evidence need to be considered and combined; correlational evidence can demonstrate a link between exposure and behaviour under naturalistic conditions, and experimental evidence can establish a causal link under controlled conditions. Other researchers agree that different research designs each have a role to play in establishing an evidence base for decision making, “Every research strategy within a discipline contributes importantly relevant and complementary information to a totality of evidence upon which rational clinical decision making and public policy can be reliably based” (Hennekens & Buring, 1994). Clearly SR does not represent a “straitjacket” approach that permits only the consideration of perfect experiments. Systematic reviewers have increasingly recognised the importance of different types of evidence; even qualitative research now has an established role within SR following years of lively debate about how exactly it should be incorporated (Dixon-Woods & Fitzpatrick, 2001).

If evidential criteria and thresholds are appropriately set, and the context and requirements of social science research are recognised as different, SR can still be very useful.
**DID THE SYSTEMATIC REVIEW HELP?**

This was the first systematic review\(^2\) to be conducted on the effects of marketing practice. This combined with the topical and contentious nature of obesity meant that the review also attracted substantial media interest (Carvel, 2003, p11; Elliot, 2003, p7; Frith, 2003, p3; Uhlig, 2003, p1). Additional dissemination took place through a series of sixteen seminar and conference presentations across the UK, involving all the key stakeholders, including the general public. In addition the review as been presented in the USA (Hastings, 2005a), Australia (Hastings, 2004) and New Zealand (Hastings, 2005b). This has exposed the work to examination by journalists of all political persuasions and academics from a range of disciplines. The review has proved equal to this scrutiny, and been shown to be both comprehensive and rigorous: no researcher or study emerged that had been either overlooked or misrepresented.

But perhaps the hardest test of the review has been its capacity to withstand critiques from commercial marketers. These took two principal forms: an alternative review (Young, 2003) and a detailed critique (Paliwoda & Crawford, 2003), both funded by an advertising trade organisation. The alternate review (Young, 2003) used conventional literature review procedures - that is, it was not explicit about how sources had been identified and selected - and reached the opposite conclusion to the systematic review, arguing that there was no evidence of food promotion influencing children. Young’s case partly rested on the contention that nearly all of the existing evidence contains some methodological validity - experiments lack external validity and cross-sectional and observational studies lack internal validity - and therefore no safe conclusions could be drawn. The Food Standards Agency convened an ‘Academic Seminar’ of senior academics to compare the two reviews. The
seminar recognised the fact that the Young review was based on considerably fewer studies, questioned the grounds on which he had selected and assessed studies, and indicated that his position of rejecting both kinds of evidence as unsatisfactory was unhelpful (Food Standards Agency, 2003a). The systematic procedures were deemed to be more reliable than the conventional review methods used by Young, and the systematic review’s findings were strongly endorsed. This adjudication was then published (Food Standards Agency, 2003a):

“On the balance of evidence the Hastings review had provided sufficient evidence to indicate a causal link between promotional activity and children’s food knowledge, preferences and behaviours.”

“There were inconsistencies in the way he [Young] had assessed the results of these studies. As such there were concerns expressed as to whether the conclusions reached by Young could be fully justified.”

The main arguments of the commissioned critique (Paliwoda & Crawford, 2003) were that the SR was “unscientific” and that other factors were more important than promotion in terms of influencing children’s behaviour. The Paliwoda and Crawford critique was also sent by the FSA to peer review. Reviewers rejected it as “very disappointing” and of “low” quality. Its conclusions were dismissed as not being “justified by any ‘findings’”. The reviewers’ adjudication was again published (Food Standards Agency, 2003b):

“One respondent found Paliwoda and Crawford’s comment….to be a facile argument.”
“Paliwoda and Crawford were not clear about how they considered Hastings to be ‘unscientific’. It was noted that Hastings provides extensive detail on the selection processes and criteria it uses.”

The robustness of the SR has subsequently been demonstrated by the response of policy makers. Its findings were formally ratified by the FSA Board and have been accepted by the Ministry responsible for telecommunications (Department for Culture, Media and Sport). Most importantly of all, it has directly informed government policy. The recent Public Health White Paper (the channel by which the UK Government expresses its legislative intent) states that “there is a strong case for action to restrict further the advertising and promotion to children of those foods and drinks that are high in fat, salt and sugar” (Department of Health, 2004). The UK government is also introducing further policies (eg. such as banning unhealthy vending machines from schools in England and Wales) to protect children from the dangers of junk food.

TOWARDS EVIDENCE-BASED MARKETING

The business community needs to be able to speak with authority about the impact of its practices on society. In what are increasingly litigious and accountable times, anything less than the most rigorous evidence base is going to be inadequate. This study demonstrates that the bio-medical concept of evidence based decision making, and the systematic review procedures on which it relies, can transfer across to the field of marketing. As noted above there are problems and discomforts in applying such precise procedures to our field, but the benefits of doing so are considerable.
First and foremost, SR methodology provides a transparent, rigorous and objective summary of the current evidence base upon which an informed policy debate can take place (McDonald, 2003). Challenges to this can, of course, be made but they need to be equally thorough and show how and why they differ from the original; opinion, ad hoc studies and even conventional reviews will simply not withstand comparison. In the food promotion case, for example, peer review repeatedly ratified the findings of the systematic review and rejected the findings of conventionally conducted literature reviews which used selective and partial search and inclusion criteria (Food Standards Agency, 2003b). Furthermore, the transparency of the SR process and its dependence on external review and scrutiny has the benefit of involving other investigators, stakeholders and policy users in the process. Polonsky and colleagues (2003) note that solutions to the ‘harm chain’ caused by certain marketing activities, such as tobacco and food promotion, require “a broader degree of cooperation amongst all stakeholder groups (eg. governments, not-for-profits, consumers and firms)” (p360). Systematic review, with its transparent methods and processes, potentially facilitates this cooperation and provides a focus for dialogue. It is worth noting that even the SR cannot eradicate all doubt or uncertainty - for example, decisions made by the National Institute for Clinical and Health Excellence (NICE) about the effectiveness of different drugs are not always readily accepted by the public and other stakeholders – but it does provide a solid foundation upon which informed debates can occur and hopefully be resolved.

Systematic review also changes the rules of engagement. As noted in section 2 the sheer scale and effort involved, combined with tightly defined procedures which greatly constrain researcher subjectivity, keep the focus on increasing scientific understanding, rather than supporting or attacking a particular perspective. One study uncovered by the review
illustrates the potential this offers. It showed how advertising on vending machines could encourage school children to opt for healthier food items (French et al, 2001). In the process it demonstrated that advertising can influence behaviour, but also that this influence can be beneficial as well as harmful to food choices. Thus the focus shifts from the search for culprits to the identification of solutions.

In the UK the debate is now shifting towards this positive perspective. The acceptance of the review findings and methodology by at least sections of the food industry is enabling them to engage in policy debate as intellectual equals rather than the representatives of an important but partial vested interest.

Systematic review also has the potential to contribute constructively in other contested areas of marketing and public policy. Controversy currently surrounds the issue of direct-to-consumer advertising of prescription drugs. For example, there is heated debate between Australia, which currently disallows it, and New Zealand which permits it, surrounding standardisation of practice between the two countries. While New Zealand is considering banning direct-to-consumer advertising, other countries including Canada and EU countries, are considering relaxing their restrictions. Various strategies have been proffered as a means of resolving the issue, such as harm chain analysis (Polonsky, 2003) and stakeholder analysis (Maubach & Hoek, 2005); both are approaches which have been developed with the aim of resolving differences between stakeholders and reconciling divergent opinions in order to provide a more consensual approach to policy development. Systematic review has the clear potential to complement such approaches.
Accepting a move to evidence based marketing has important implications for academics. There is a need to develop expertise in meta-analysis and the synthesis of evidence. Textbooks on research methodology should describe both the methodological criteria used to systematically evaluate the validity of different types of marketing evidence and the quantitative techniques used for summarizing that evidence. Theoretical and methodological debate is needed among marketing academics concerning the utility and nature of a hierarchy of marketing evidence: what types of evidence should be given most weight in assessing the impact of marketing interventions? Where should the threshold for ‘good enough evidence’ be set? How little or how much heterogeneity of study designs is appropriate in providing a definitive picture of marketing effects?

Peer reviewers will have to start demanding of literature review authors that they articulate how and why they selected particular studies, and how they appraised the evidence. Journals could assist the systematic review approach by adopting a more structured abstract format which incorporates issues of methods and design into the portion of an article the reader sees first, as have many medical journals. Ultimately, some type of Cochrane or Campbell Collaboration for marketing should be developed.

These are prodigious challenges. Furthermore, marketing is ultimately concerned with that most subjective and illusive of phenomena - human behaviour, and consequently, marketing knowledge will ever be elaborate and contested. Arguably this militates against rigour and scientific precision, and suggests that attempts to achieve these will always be found wanting. At the same time, however, the fact that we are dealing with people, and not just human behaviour but the human condition, makes it all the more important to try.
Figure 1: Overview of stages in the review process

Stage 1: Preliminary literature search and analysis

Stage 2: Development of review questions and search strategy

Stage 3: Searches for relevant literature:
- Electronic databases
- Grey literature
- Personal contact
- Examination of bibliographies

Stage 4: Initial relevance assessment (assessing titles and abstracts for relevance to the review)

- Initial relevance criteria
  1. Publication date 1970 onwards
  2. English language study
  3. Is a primary research study or review
  4. Relates directly to the extent and nature of food promotion to children and/or the effects of food promotion on children
  5. Where any of the terms, if mentioned, correspond to agreed definitions of food, children, promotion, food KPB

Stage 5: Retrieval of full text papers

Stage 6: Second stage relevance and quality assessment

Stage 7: Data extraction

Stage 8: Rating of study quality

Stage 9: Data synthesis and analysis

REPORT
Figure 2: The extent and nature of food promotion to children

<table>
<thead>
<tr>
<th>Research Question</th>
<th>Results</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>1. What promotional channels are being used to target children?</strong></td>
<td>All 50 studies addressed this question.</td>
</tr>
<tr>
<td></td>
<td>Television is the principal channel used by food marketers to reach children.</td>
</tr>
<tr>
<td></td>
<td>There is some evidence that the dominance of television has recently begun to wane. The importance of strong, global branding reinforces a need for multi-faceted communications combining television and merchandising, ‘tie ins’, and point-of-sale activity.</td>
</tr>
<tr>
<td><strong>2. What food items are being promoted to children?</strong></td>
<td>41 studies addressed this question.</td>
</tr>
<tr>
<td></td>
<td>Food products dominate children’s advertising and the majority of this promotes the so-called ‘big four’ of pre-sugared breakfast cereals, soft-drinks, confectionary and savoury snacks. In the last ten years advertising for fast food outlets has rapidly increased turning the ‘big four’ into the ‘big five’.</td>
</tr>
<tr>
<td></td>
<td>The advertised diet contrasts sharply with that recommended by the public health community. The recommended diet gets little support.</td>
</tr>
<tr>
<td><strong>3. What are the creative strategies used to target children?</strong></td>
<td>34 studies addressed this question</td>
</tr>
<tr>
<td></td>
<td>Themes of fun and fantasy or taste, rather than health and nutrition are used to promote foods to children. Fast-food advertising tends not to describe the product that is being advertised but focuses on the experience of the meal and the brand.</td>
</tr>
<tr>
<td>Author</td>
<td>Year</td>
</tr>
<tr>
<td>-------------------</td>
<td>------</td>
</tr>
<tr>
<td>Atkin</td>
<td>1975</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td>Bolton</td>
<td>1983</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td>Borzekowski &amp; Robinson</td>
<td>2001</td>
</tr>
<tr>
<td>Cantor</td>
<td>1981</td>
</tr>
<tr>
<td>Clarke</td>
<td>1984</td>
</tr>
<tr>
<td>Coon et al</td>
<td>2001</td>
</tr>
<tr>
<td>Dawson et al</td>
<td>1988</td>
</tr>
<tr>
<td>Study</td>
<td>Year</td>
</tr>
<tr>
<td>-----------------------------------------</td>
<td>------</td>
</tr>
<tr>
<td>Dietz &amp; Gortmaker</td>
<td>1985</td>
</tr>
<tr>
<td>French et al</td>
<td>2001</td>
</tr>
<tr>
<td>Galst</td>
<td>1980</td>
</tr>
<tr>
<td>Galst &amp; White</td>
<td>1976</td>
</tr>
<tr>
<td>Goldberg et al (Study 1)</td>
<td>1978</td>
</tr>
<tr>
<td>Goldberg et al (Study 2)</td>
<td>1978</td>
</tr>
<tr>
<td>Study</td>
<td>Year</td>
</tr>
<tr>
<td>-----------------------------</td>
<td>------</td>
</tr>
<tr>
<td>Goldberg</td>
<td>1990</td>
</tr>
<tr>
<td>Gorn &amp; Florsheim</td>
<td>1985</td>
</tr>
<tr>
<td>Gorn &amp; Goldberg</td>
<td>1980</td>
</tr>
<tr>
<td>Gorn &amp; Goldberg</td>
<td>1980/1982</td>
</tr>
<tr>
<td>Gracey et al</td>
<td>1996</td>
</tr>
<tr>
<td>Heslop &amp; Ryans</td>
<td>1980</td>
</tr>
<tr>
<td>Jeffrey et al (Study 1)</td>
<td>1982</td>
</tr>
<tr>
<td>Study</td>
<td>Year</td>
</tr>
<tr>
<td>-----------------------</td>
<td>---------</td>
</tr>
<tr>
<td>Jeffrey et al., Fox</td>
<td>1981/82</td>
</tr>
<tr>
<td>Kaufman &amp; Sandman</td>
<td>1983</td>
</tr>
<tr>
<td>Peterson et al</td>
<td>1984</td>
</tr>
<tr>
<td>Reeves &amp; Atkin</td>
<td>1979</td>
</tr>
<tr>
<td>Ritchey &amp; Olson</td>
<td>1983</td>
</tr>
<tr>
<td>Ross et al</td>
<td>1980/81</td>
</tr>
<tr>
<td>Study</td>
<td>Year</td>
</tr>
<tr>
<td>-----------------------------</td>
<td>------</td>
</tr>
<tr>
<td>Stoneman &amp; Brody</td>
<td>1981</td>
</tr>
<tr>
<td>Stoneman &amp; Brody</td>
<td>1982</td>
</tr>
<tr>
<td>Winman &amp; Newman</td>
<td>1989</td>
</tr>
<tr>
<td>Wong et al</td>
<td>1992</td>
</tr>
</tbody>
</table>
Figure 4: Hierarchy of evidence

<table>
<thead>
<tr>
<th>Rank</th>
<th>Methodology</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Systematic reviews and meta-analyses</td>
</tr>
<tr>
<td>2</td>
<td>Randomised controlled double blind trials</td>
</tr>
<tr>
<td>3</td>
<td>Cohort studies</td>
</tr>
<tr>
<td>4</td>
<td>Case control studies</td>
</tr>
<tr>
<td>5</td>
<td>Cross sectional surveys</td>
</tr>
<tr>
<td>6</td>
<td>Case series</td>
</tr>
<tr>
<td>7</td>
<td>Case reports</td>
</tr>
</tbody>
</table>

Source: Jones (2002)
NOTES

1. The Universities of Strathclyde, Oxford, York and London City

2. Since the review was completed in Sept 2003, the Cochrane Collaboration has produced a review on tobacco advertising and its impact on young people’s smoking
REFERENCES


*British American Tobacco UK Ltd and others v The Secretary of State for Health* (2004) EWHC 2493 (Admin).


Coalition on Food Advertising to Children (CFAC) (2003) *Children’s Health or Corporate Wealth? The case for banning television food advertising to children (A briefing paper).* Australia: CFAC.


