Benchmarking: An International Journal

A design science research approach to website benchmarking

Leonie Cassidy John Hamilton

Received 9 July 2014
Revised 15 December 2014
Accepted 27 December 2014

A design science research approach to website benchmarking

Leonie Cassidy and John Hamilton
College of Business, Law and Governance, James Cook University, Cairns, Australia

Abstract

Purpose – Literature-identified website benchmarking (WB) approaches are generally time consuming, survey based, with little agreement on what and how to measure website components. The purpose of this paper is to establish a theoretical approach to WB. A comprehensive design science research methodology (DSRM) artifact facilitates the evaluation of the website against the universal set of benchmark components. This knowledge allows managers to gauge/reposition their websites.

Design/methodology/approach – DSRM establishes a website analysis method (WAM) artifact. Across six activities (problem identification, solution objective, artifact design/development, artifact demonstration, artifact evaluation, results communication), the WAM artifact solves the DSRM-identified WB problem.

Findings – The WAM artifact uses 230 differentiated components, allowing managers to understand in-depth and at-level WB. Typological website components deliver interpretable WB scores. Website comparisons are made at domain (aesthetic, marketing, technical) and/or functional levels.

Research limitations/implications – New/emergent components (and occasionally new functions) are included (and redundant components removed) as upgrades to the DSRM WAM artifact's three domains and 28 functions. Such modifications help keep latest benchmarking comparisons (and/or website upgrades) optimized.

Practical implications – This DSRM study employs a dichotomous present/absent component approach, allowing the WAM artifact's measures to be software programmed, and merged at three different levels, delivering a useful WB tool for corporates.

Originality/value – DSRM identifies the benchmarking problem. Rough-cut set-theory and mutual-exclusivity of components allow the causal-summing of typological website components into an objective WAM artifact WB solution. This new, comprehensive, objective-measurement approach to WB thus offers comparative, competitive, and website behavioral implications for corporates.

Keywords Benchmark, Transaction, Behaviour, Causation, Artefact, Set-theory, Website analysis method, Competitive

Paper type Research paper

Introduction

Websites are assets that can deliver high dollar return-on-investment – especially when fully incorporated into a corporate's business model (Simmons *et al.*, 2008). Although the global e-commerce sales through websites are to exceed $1.5 trillion in 2014 (eMarketer, 2014), at the national and corporate level, business websites generally do not capture the quality, the innovation, and the competitive positioning benchmarks (O'Cass and Ngo, 2012) expected when the corporate is in pursuit of high financial returns, yet they provide a key 24/7/365 connecting environment to online consumers, and visitors, and inquirers.

Benchmarking: An International Journal
Vol. 23 No. 5, 2016
pp. 1054-1075

1463-5771
DOI 10.1108/BIJ-07-2014-0064

Website benchmarking (WB) approaches

Currently well over 150 website components have been identified (Cassidy and Hamilton, 2011b; Olsina and Rossi, 2002; Stepchenkova *et al.*, 2010), but few studies have attempted to use these large numbers of measures when comparing or benchmarking websites. In addition there is little consistency or consensus on what and how to benchmark. However, some recognize WB should deliver a process that is both easy-to-implement and one that deliverers timely, cost-effective, and interpretable (internal and external) results (Boisvert and Caron, 2006).

From a corporate perspective, although websites do change, WB is often assessed irregularly, and inconsistently. This assessment has involved specific scoping (Lee and Kozar, 2009), and subjective comparisons of convenient components at a selected point-in-time, or the occasional comparison against a selection of competitor websites.

WB employing surveys is complicated because respondents are often restricted by their level of understanding regarding the website's marketing, aesthetics and design, technical parameters, language terminologies, or its social and interactive components. Also when large amounts of benchmarkable data are to be collected (Krosnick *et al.*, 2002) the respondent often suffers from task-time fatigue (Lee *et al.*, 2011) and their benchmarking decisions show cognitive inconsistencies (Krosnick *et al.*, 2002).

Other corporates choose to derive their WB metrics through computer scanning approaches (Calero *et al.*, 2005; Olsina and Rossi, 2002) and accept comparative ratings where the website is deemed more innovative, or economic, or strategic (Zhu *et al.*, 2009). As most computer scanning, WB approaches incorporate degrees of human subjectivity in their design or interpretation ratings, these approaches are limited. For example, most are point-in-time developed and scoped by the programming team – who set which website components are to be evaluated and interpreted, and then decide on the intelligence of the software, its comparative mechanisms and its associated metrics.

Although a raft of WB survey software and scanning approaches has emerged each displays limitations in scope, or in the capture of new technologies or features, or in inherent program and reporting subjectivities.

Study motivation

This study is motivated by an interest in e-business, and information systems research, along with concerns for corporates and their websites when competing globally. Further motivation emanates from the relatively slow and often small advances, deployed in business website comparison approaches, and that to-date WB has lacked both a theoretical approach and an approach to capture relevant analytics. Another motivation is to offer corporates a simple, comprehensive, and time-efficient way to interpret their website components and so overcome managerial reluctance to instigate website change(s). A final motivation emanates from the knowledge that when a corporate, through its website, instigates such experimentation, exploration, and exploitation approaches, it can release new competitive and positioning opportunities (Adler *et al.*, 2009). An extensive literature study also identifies (and supports) a need for a different and non-subjective universal approach to business WB. Hence we draw from information systems and look to the design science paradigm as the philosophical basis (Fink and Nyaga, 2009) to our non-subjective WB approach.

Design science research (DSR)

This study adopts a DSR approach to design and develop a new WB approach for business:

> Design science research is a research paradigm in which a designer answers questions relevant to human problems via the creation of innovative artifacts – thereby contributing new knowledge to the body of scientific evidence. The designed artifacts are both useful and fundamental in understanding that problem (Hevner and Chatterjee, 2010).

Under DSR we apply its seven guidelines to develop our technology-based solution (artifact) to solve this WB issue (Hevner *et al.*, 2004). From the literature, we first explain the gaps that exist, and that no theory shows overall relevance to this WB problem. The artifact can be a construct, or a model, or a method, or an instantiation. In this study we engage the artifact as a website analysis method (termed WAM). The WAM artifact is supported by clear, verifiable, and rigorous methodology, and is applied at the construction and evaluation stages of the artifact. The WAM artifact also offers an effective and understandable approach for its relevant audiences (Hevner *et al.*, 2004).

The following sections discuss existing literature approaches to WB and the design science research methodology (DSRM) employed within this study. The application of set-theory; Aristotle's theory-of-causation; planned behavior, motivation, consumption, and gratification theories are further considered for applicability to WB. The study then describes the deployment and evaluation of the WAM artifact, and concludes with an assessment of implications and other areas befitting future research.

Background

Current literary approaches to business WB include theoretical approaches such as subjective impressions of website consumers (Barnes and Vidgen, 2002; Io Storto, 2014), perceived qualities (Rashid and Othman, 2010; Yoo and Donthu, 2001), theory-of-organizational-trust (Fang *et al.*, 2014), and usability. Other WB studies adopt consumer satisfaction and literary assessments (Elling *et al.*, 2007), or prior research across information and service qualities (Webb and Webb, 2004). Loiacono *et al.* (2007) engage the theory-of-reasoned-action and technology acceptance models (TAMs) for marketing, while other WB approaches synthesize the expectation-disconfirmation paradigm and create theories around consumer satisfaction (Ghasemaghaei and Hassanein, 2013; McKinney *et al.*, 2002). Still others study how the visual complexity and ordering of components within a website affect the consumer (Deng and Poole, 2010). The above approaches to WB are based around consumer opinion; hence they remain open to subjectivity constraints.

Some WB researchers adopt narrowly focussed, survey-based approaches, and for a specific purpose, generally employ selected survey components (rather than considering a broad-ranging assessment of the website). These study-specific website aspects include: effects of privacy and perceived security on user trust levels (Chang and Fang, 2013; Flavian *et al.*, 2006), influences of online trust (Chen and Barnes, 2007), e-service quality (Udo *et al.*, 2010), information flows between websites (Nel *et al.*, 1999), and consumer website satisfaction (Luo *et al.*, 2012). Such WB studies are restrictive and are designed to assess specific components, and so cannot provide comprehensive solutions.

Survey-based WB tools

Several WB tools are actually survey-based methods. These studies typically target specifically chosen website components. Such approaches include WebQUAL 4.0

(Barnes and Vidgen, 2002), WEQ (Elling *et al.*, 2007), SITEQUAL (Yoo and Donthu, 2001), eTailQ (Wolfinbarger and Gilly, 2003), SiteQual (Webb and Webb, 2004), MUSA+ (Grigoroudis *et al.*, 2008), e-SQ (Hu, 2009), and PEEIM (Fang *et al.*, 2014). These approaches employ questionnaires and each measures less than 58 website components on a five or seven point Likert type scale.

Although questionnaires remain popular for data capture, several psychological and human respondent issues warrant consideration prior to survey implementation. First, when researchers use exchange theory to improve the WB motivation for survey completion they must account for respondents who weigh-up the reward/benefit they receive when completing a survey against their cognitive cost (or effort required) (Albaum *et al.*, 1998; Evangelista *et al.*, 2012). Here, and in self-reported questionnaires, respondents weigh a question's focus and then attempt to anticipate the answer being sought (Cao *et al.,* 2005). Thus, bias can emerge in exchange style approaches.

Second, as survey-length increases, response-time increases. Here, some respondents answer each survey question to the best of their ability, and some may experience survey fatigue (Nathan and Yeow, 2011) when weighing their personal cognitive cost against time to select from near-identical questions (Krosnick, 1991). In contrast, less rigorous respondents may just select "don't know/neutral," or select randomly across the scale (Krosnick *et al.*, 2002), while others may select an answer that appears acceptable, but not necessarily optimal in their view (Krosnick and Alwin, 1987; Klandermans, 1984). Hence, carefully prepared and balanced surveys are imperative, but the number of questions asked of respondents remains numerically restrictive when compared to the full range of components available for WB.

Third, "response-order-effects" such as question wording, complexity of rating scale, or order of answers (Krosnick and Alwin, 1987; Krosnick, 1991; Weijters *et al.*, 2010) can directionally lead responses (Schwarz, 2007; Weijters *et al.*, 2010).

Fourth, many past studies utilize student respondents (Loiacono *et al.*, 2007), but such convenience sample approaches lacks validity (Wells *et al.*, 2011) and transferability when applied to consumer diversity within global web communities (Lee *et al.*, 2012; Yang *et al.*, 2012; Xu *et al.*, 2013). These studies may also introduce unexpected (or unexplained) biases (Steelman *et al.*, 2014).

Hence, although well-executed, surveys generally account for the majority of psychological and respondent issues (Krosnick *et al.*, 2002); they remain constrained by the above. In addition, most WB studies are lacking in a strong theoretical base (Io Storto, 2013; Lee and Kozar, 2009; Yoon and Kim, 2009), with some applying various theories to benchmarking tools – which they then adapt, and apply to websites.

WB using modified business tools

Some WB approaches have modified existing business application tools. Kaplan and Norton's (1992) balanced scorecard (BSC) approach has been modified in several tourism and travel WB studies. These studies typically use commercially available software to measure technical performance of the website (Choi and Morrison, 2005; Lee and Morrison, 2010). However, they show considerable variation in their use of the remaining three BSC dimensions – with Choi and Morrison (2005) using consumer, marketing-effectiveness, and/or business perspectives, Lee and Morrison (2010) adopting marketing, consumer, and internal-critical-success-factors, and Kline *et al.* (2004) adding consumer friendliness, site attractiveness, and marketing effectiveness.

Dichotomous (present/absent) WB studies (Kline *et al.*, 2004; Choi and Morrison, 2005; Lee and Morrison, 2010) provide yet another modified BSC approach, however these again show inconsistencies and differing degrees of detail, but they do extend WB studies into classifications that engage between 40 and 150 components.

TAM variants applied to WB have added online shopping perceptions across 45 Likert scale web components of e-service quality and perceived measures of service value, trust, and usefulness (Lee and Wu, 2011), others used 50 Likert scale web components to consider trust and intention to purchase (Gefen *et al.*, 2003). While in tourism accommodation, TAM is modified using semi-structured interviews across 36 website components (Herrero and San Martin, 2012). Again WB variants are purpose specific and inconsistent in approach.

Component-based WB tools

From a WB tools perspective, the key-quality-factors approach rates 28 components as present, or absent, or not assessed and graphs comparisons (Cox and Dale, 2002). In pursuit of potential website performance improvement areas, function-benchmarking captures 91 present or absent business components under 18 separate groups, and rates a level-of-presence for each group (Boisvert and Caron, 2006). Again WB tool approaches lack consistency.

Content-analysis WB approaches also vary with 53 components across six service quality domains (each containing closely related components) (Nusair and Kandampully, 2008), or with website-usability-factors tracked using 38 Likert scale consumer-needs components (Nathan and Yeow, 2011).

Expert reviewers of websites often display differing views of website component groupings (Cassidy, 2010). To improve consistency Stepchenkova *et al.* (2010) use two experts over 11 weeks to build their 99 component and four domain WebEVAL tool.

The above WB approaches are inconsistent, lack precise point-in-time comparison capabilities, are respondent completion-time restricted, are survey scope-and-size restricted, and at some point each WB approach relies on human involvement and subjective judgment. Hence, automated WB tools are developed.

Automated WB tools

Automated WB tools such as WebTango, crawl a website, and compute metrics on 157 components (Ivory and Hearst, 2002), but this tool is limited to website design (Ivory and Megraw, 2005). WebQEM is a hyper-document WB tool tracing 150 direct (or indirect) quality aspects (Olsina and Rossi, 2002) but overlap between the WB measures arises, and its restrictive scoping leaves out other WB measures.

Hypertext markup language (HTML) validation programs such as Weblint, are best used in combination with other HTML checking tools (Bowers, 1996). For example, EvalIris evaluates the markup of webpages, identifying accessibility failures through HTML components (Abascal *et al.*, 2004); however some accessibility issues still require manual checking. HTML tools focus on the technical domain, and so offer narrow WB solutions.

KWARESMI extracts, structures, and organizes web usability guidelines toward an automated evaluation (Beirekdar *et al.*, 2002), but only does so one page at a time. GIST collects information on users; then infers website behavior and attitudes; then identifies and tests new nano-segments; and tracks the gaps across chosen WB segments (Albert *et al.*, 2004). As per previous approaches, these automated WB tools also lack

completeness, requiring degrees of human consideration (developer or respondent), and are often costly to implement.

Thus, a significant WB gap remains in the literature, and we now present our proposal for an artifact to solve this business problem. This artifact – termed WAM (our website analysis method) draws together existing web components under three domains (marketing, aesthetics, and technical) that are associated with a website.

Methodology

DSR

This study follows a DSRM (Peffers *et al.*, 2007), and applies DSR guidelines (Hevner *et al.*, 2004) to the practical study of WB. Table I summarizes the six DSRM activity stages. The above introduction and literature review completes DSRM activity stages one and two, identifies WB approaches, identifies the limitations in the literature, identifies and discusses problems for organizations, and then defines objectives for the solution to the identified WB business problem.

This study now discusses activities three, four, and five of the DSRM process. It then considers the implications of this research, suggests what path future research may take, and lastly provides conclusions to the research community at this point-in-time.

Design and development

Mutually exclusive set-theory and causation theory. The measureable components sourced across the literature each represent a different, and mutually exclusive, part of the website. Hence, using mutually exclusive set-theory we can sum the components present on a website. We can then compare this internal WB score, to the WB scores obtained from other (external) websites.

As set-theory houses mutual exclusivity we collate each website component present into just one mutually exclusive subset of similar items – termed a function. Each function, and its small cluster of similar website components, is grouped with other like

Process Iteration

Activity 1. Problem identification and motivation	• Approaches in literature: time consuming, generally survey based, lack of agreement on components (and how measure), limited components and/or area measured • Organizations need: easy-to-implement, all-encompassing tool that benchmarks internally (and externally) and offers multi-level comparisons – and easy-to-interpret, efficient and cost effective, and capable of delivering interpretable solutions for website improvement
Activity 2. Define objectives for a solution	• Extensive literature review and investigation of research developed tools • Develop comprehensive and integrated framework to evaluate websites. One allowing managers to strategize and possibly reposition their website (possibly to better fit user expectations)
Activity 3. Artifact design and development	• DSRM, set-theory, causation theory, planned behavior, motivation theory, gratification theory, consumption theory • WAM artifact = multi-level, hierarchical, 234 literature identified, mutually exclusive components into 28 mutually exclusive functions into 3 mutually exclusive domains. Measure components present = 1 or absent = 0 = software programmable easy-to-interpret scores at level, cost effective
Activity 4. Demonstration	• Implement WAM during the design stages of www.therideguide.com.au • A purposed-built publicly accessible tourism website built in 5 monitored stages – each with more components added to gauge if component additions influence traffic on an active website and if these additions can be deemed website improvements
Activity 5. Evaluation	• Google advanced analytics and server statistics track consumers. Data analyzed and results compared with objectives (est. June 2013) • Statistical analysis and interpretation
Activity 6. Communication	• Journal articles and conferences • *DSR* and *IS* journals, and *Benchmarking*, *An International Journal*, and ACIS conferences

Source: Adapted from Peffers *et al*. (2007)

Table I.
Design science research methodology

functions (each made up of their own mutually exclusive set of similar components). We then group these mutually exclusive functions into one of three mutually exclusive domains. The sum of all the components present across the three domains of a website constitutes the total score of all the components present on the website.

When measured against the sum of all the components possible for any website, a benchmark score for this website can be generated, and website comparisons at component, functional, and domain level are achievable. Thus, a benchmarking process is theoretically available for websites. The difficulty is – how can this be done, and can the existing website components be meaningfully interpreted?

To understanding the possible website components through mutually exclusive set-theory approach we define the WAM artifact as the universal set of all possible website components. We note that within a domain, the full subset of literature-established possible website components making up a function can be different in number when directly compared to another function's full subset of components. Hence, to mathematically compare functional contributions to a website – a functional scaling process is also required (Figure 1).

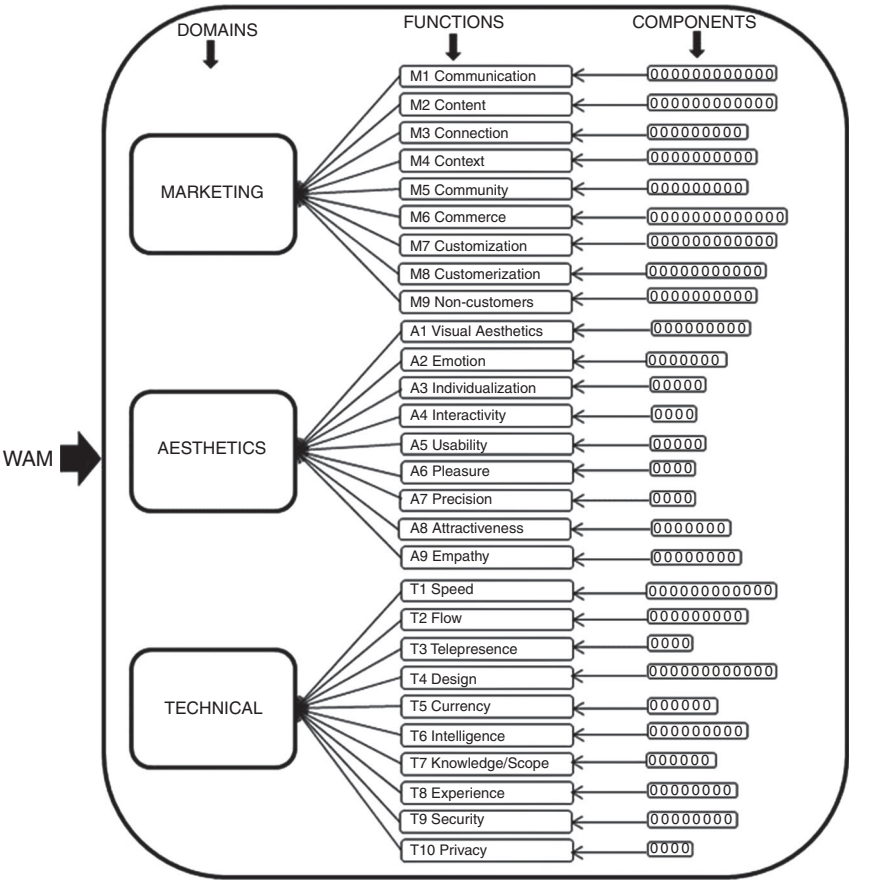


Figure 1.
WAM

Source: Developed from Cassidy and Hamilton (2011a)

Next we combine mutually exclusive set-theory with Aristotle's four-step theory-of-causation (material cause, formal cause, efficient cause, and final cause) (Falcon, 2011), and build the following linkages:

(1) Material cause: the WAM artifact universal set is defined. The WAM artifact captures all possible literature-supported website components.

(2) Formal cause: all components are linked and compared through typologically collated subsets and sets. The presence (or absence) of components within a subset or website set is determinable.

(3) Efficient cause: the existing components of a website are collected and compared to the WAM artifact universal set of all possible website components. This allows internal benchmarking at the component, the function, the domain, and at the overall website level. It also allows external benchmarking against other chosen (or competitive) websites.

(4) Final-cause: the universal set components not accounted for in a website are presented as both improvement considerations and as possible future inclusion components. This can be useful when strengthening weaker functions or poorly covered domains within a website.

Rough-cut mutually exclusive set-theory and causation theory. The WAM artifact is dynamic and can grow. For example – as the literature advances, and technology advances, new website components continually emerge, and each of these new components potentially adds to a function, then to a domain, and then to the website's subset score of the WAM artifact universal set. Alternatively, if the new component represents a breakthrough technology (Cassidy and Hamilton, 2011b), another new function (or even a new domain) may start to emerge for inclusion in the WAM artifact universal set. Likewise, obsolete components (such as obsolete website software-related technical components) are considered for deletion.

Thus, the WAM artifact universal set typifies an infinite number of mutually exclusive website components (Vilenkin, 1986), and at any point-in-time, functional and domain subset changes may arise. Hence, the currency of the WAM artifact universal set engaged is normally a very close approximation to the actual scope of universal set components available for websites.

Over time, component, functional, and/or domain changes are possible, and these changes affect the WAM artifact universal set. Hence, this study engages a rough-cut (Zhu, 2007), inclusion-exclusion principal of set-theory, and the causal summing of components into a rough-cut, mutually exclusive, causal-summing, set-theory approach when benchmarking websites via their components.

The WAM artifact universal set (U) contains all available website components at the point-in-time. C_p is components present on the website. The website's benchmark score (W) compares the existing components against the possible components of the universal set as follows:

$$W = \frac{\sum \text{components present}}{\sum \text{components possible}} = \frac{\sum (C_p)}{\sum (U)}$$

where: $C_p \in U$

This study's multi-level benchmarking of websites offers further finely focussed assessments, and more detailed internal or external comparison, and it also delivers an interpretable benchmark score (Boisvert and Caron, 2006).

To statistically compare components present on the website against the WAM artifact universal set each component present is recorded as "1," and each component missing is scored as "0." Components in a function are summed into a functional score. Similar functional scores are then summed into a domain score. The functional and the domain scores allow finer internal (or external) website comparisons. A functional or a domain score below its maximum exhibits a degree of "greyness" (or the extent of missing components) (Deng, 1989; Ho and Wu, 2006) in the score, and these missing components can then be scheduled for inclusion at a chosen point-in-time.

Each function captures a different aspect of the website, and the number of its literature-established components can vary. Such variations may be exacerbated when new components are added to the website. Hence, to reduce component masking, currently only 13 components are allowed per function, and to ensure sufficient triangulation, the function must possess at least four components.

Hence, the theoretical approach of rough-cut, mutually exclusive, set-theory plus the causal summing of the website components, enables current benchmarking scores and comparisons to be established.

The relative contribution of a function (F_j), from all its components (C_{ij}), is described as the maximum number of components (x_j), present in a given function, with "10" norming the components (i), into a comparable subset solution. Thus, a function (F_j) is benchmarked as follows:

$$F_j = \frac{10}{x_j} \sum_{i=1}^{x_j} C_{ij}$$

A comparable domain level (D_{jk}), benchmark score is derived with (x_{jk}) as the maximum number of components present in each function (j), of the domain (k), and y_k is the maximum number of functions in this domain:

$$D_{jk} = \sum_{j=1}^{y_k} \frac{10}{x_{jk}} \sum_{i=1}^{x_{jk}} C_{ijk}$$

A low scoring function contains less than its possible number of components, and where some missing components are deemed important by management, these can be scheduled for point-in-time addition. High scoring functions may be checked by management against industry (or competitor) norms, thus ensuring each existing component remains appropriate. Whichever decision management may make, missing website components within the WAM artifact universal set are highlighted for point-in-time attention. This logic again applies at the domain level, highlighting weaker domains for potential internal WB improvement by management.

To enable external WB, domain scores are collated into an overall website score (W) using the following formula (adding k domains, and z as the maximum number of domains):

$$W = \sum_{k=1}^{Z} \sum_{j=1}^{y_k} \frac{10}{x_{jk}} \sum_{i=1}^{x_{jk}} C_{ijk}$$

Thus, from a typological perspective, rough-cut set-theory (Zhu, 2007) identifies missing website components, and so highlights gray (or missing component) areas (Deng, 1989) within the function. But although the rough-cut, mutually exclusive, set-theory approach deploying causal-summing of WB components delivers the benchmarks required, it also needs to give effective solutions.

Cause-and-effect considerations. Benchmarking should contribute toward a superior state of affairs – one consistent with the corporate's purpose to exist, to survive and hopefully to grow (Moriarty, 2011). This is a cause-and-effect process. The cause satisfies knowledge about something as related to material components, to formal principles, to efficient actions, or to a reason to apply something (and so change a state of affairs). The effect builds on the antecedents (which themselves cause change), or the opportune/chance alterations, or the efficiencies that are generated (Hamilton, 2010). Therefore, WB also permits management to reflect, and to then change a process, or to then choose to build a different future perception.

We visualize this effect recognition process as Figure 2. In three-dimensional space, the front pale domain score of all its components within a website can be visually benchmarked for major weaknesses against the back, darker, WAM artifact universal set for each websites domain. The same process applies to the overall website vs the WAM artifact universal benchmark set. Such visual benchmarking comparisons (Boisvert and Caron, 2006) also align with our above theory. However, Aristotle's "final" cause also includes consideration around responsive decision making. Hence, we introduce cognitive theory to understand consumer decisions for visiting a website, and consider how consumer expectations can be met by management.

Cognitive theory. Ajzen's (1991) theory-of-planned-behavior suggests a consumer intends to do something – such as visiting a website. From motivational-theory, some website consumers respond to their intrinsic needs and anticipate personal internal benefits yet to be obtained from a website – such as pleasure, satisfaction, or enjoyment (Lin and Lu, 2011; Luo *et al.*, 2011; Wakefield *et al.*, 2011). Others are motivated to then pursue their extrinsic needs, and to seek useful benefits – such as obtaining information, social interaction, or casually passing time (Luo *et al.*, 2011). Thus, consumer motivation to engage a website leads to consumer consumption of some offering on the website, and this is captured as consumption theory. Further, a level of reflective consumer gratification about the website follows consumption. This is captured as user-gratification theory.

At a chosen website, the successes in motivating a consumer's consumption (LaRose and Eastin, 2004) reflects on the website's ability to intrinsically and/or extrinsically meet

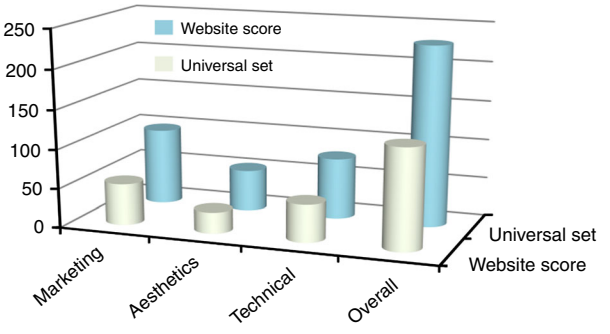


Figure 2.
Effects recognition in website benchmarking

each consumer's cognitive goals (often gauged against this consumer's acceptable cognitive cost) (Kim *et al.*, 2007). Hence, when the net motivation-to-consumption-to-gratification outcomes pathways are positive (Hamilton and Tee, 2013), then consumers likely display a greater tendency to reuse the website (Lin and Lu, 2011; Wakefield *et al.*, 2011).

Thus, managers can design their website by recognizing their consumers' likely motivations. Past studies suggest that a sequencing of the consumer expectations/motives-to-consumptive-acquisitions-to-gratifying/revisit intentions (Hamilton and Tee, 2013) may increase revisiting. Hence a stronger suite of engaging and consumer-targeted behaviorally focussed website offerings may then draw additional consumers into using (and reusing) the website.

DSRM demonstration and evaluation of the WAM artifact

To demonstrate the WAM artifact universal set as suitable for WB, standards can be developed across chosen industry-sector websites – but commercial IP issues may hinder access to such websites. A second WAM artifact universal demonstration can see a few similar business websites undergoing sequential changes in specific component areas – but obtaining permission to instigate such ongoing changes remains difficult.

This study opts to demonstrate the WAM artifact universal set through the build of a commercial (and consumer-targeted) website. The study adopts a tourist information website, specifically built for the region's niche market of motorcycle tourism. This sequential build adds consumer-engaging components stage-by-stage. Each build stage remains commercial, fits the chosen existing marketplace, and has sufficient consumer reach. Also, at the appropriate development stage all normal website tools (SEO, Facebook, Twitter, Flickr, YouTube, and Forums) are deployed.

This study employs Google advanced analytics server statistics (and other supporting web-analytics services) to monitor six consecutive build stages of www.therideguide.com.au. This study was conducted over 100 weeks between mid-2012 and mid-2014. Each website stage is live, and is tracked for at least 12 weeks, and each stage incorporates additional components – ones designed to advance the website's overall capabilities, and to grow its consumer base.

Homepage screen captures of each website stage are shown at Figure 3 (and all screen capture pages are retained). Once the five website stage developments conclude the website ownership is transferred to a national web-marketing company and is to undergo further regional and national motorcycle ride-tourism development.

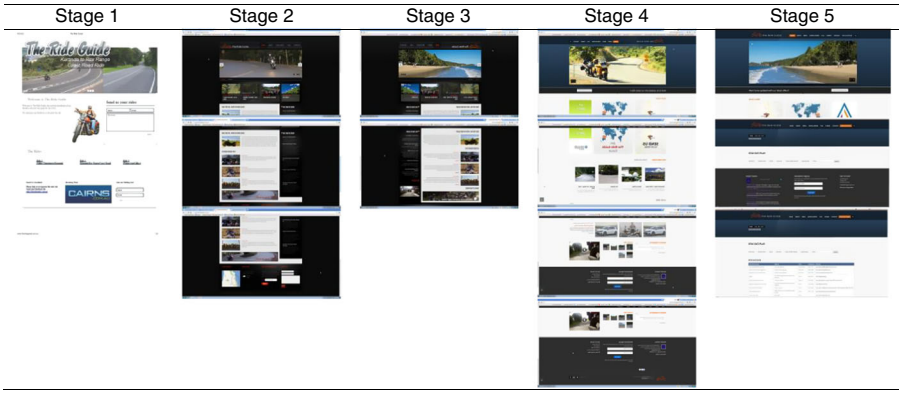


Figure 3.
The ride guide homepage at each change

Stage one is a static one page "brochure" of limited information. This represents the simplest website presence for www.therideguide.com.au. Stage two is a static website of five pages – with basic ride information, photos, static-weather and news, frequently asked questions, an about-us section, and a contact-us section. Stage three adds the basic motivations of communications, static maps, additional ride details, and a YouTube channel.

Stage four commences the build that caters for additional expectations. It includes the interactive consumptive value additions of live map variations (day, street and 3D) and basic social media (twitter, Facebook, forums), along with embedded videos, and a photo gallery. The gratification component of the website includes a "send us your rides" and a forum section. These additions are designed to improve the website's value and to grow consumer traffic.

Stage five houses a fuller suite of consumptive interactive components, plus a significant rider-selection database of activities, ideas, information, and accommodation – designed to add to the motorcycle rider's experiences. Again the study maps consumer traffic connections and the cycles of actions they pursue. Hence, we test the WAM artifact using staged component additions across the development cycle of this commercial website. We stepwise expand the motivation, consumption, and user-gratification component areas that likely appeal to motorcycle riders.

Lastly, stage six moves the website beyond this study and into a commercial management and national reach status – where future competitive analysis is limited.

At each stage of www.therideguide.com.au's development its function, domain, and website scores are calculated using the above formulae (from the rough-cut mutually exclusive set-theory and causation theory section). Table II shows the nine aesthetic domain scores (A1-A9) grouped across the five stages of our planned website development (as the number of components possible per function, the number of components present per function, the normed score per function and the domain totals).

Table III provides the three website domain scores at each stage of development for TheRideGuide.com.au, along with the corresponding benchmark score and the corresponding percentage of the total score possible.

The visual effects recognition of this TheRideGuide.com.au website summarizes the internal website analysis against the WAM artifact universal set. This website (www.therideguide.com.au) enlists fewer components than available in the WAM

		Present components					Normed out of 10 component score				
Function	Possible components	Stage 1	Stage 2	Stage 3	Stage 4	Stage 5	Stage 1	Stage 2	Stage 3	Stage 4	Stage 5
A1	9	3	8	8	8	8	3.33	8.89	8.89	8.89	8.89
A2	7	4	5	5	6	6	5.71	7.14	7.14	8.57	8.57
A3	5	0	0	0	0	0	0	0	0	0	0
A4	4	1	2	2	3	3	4	5	5	7.5	7.5
A5	5	3	3	3	4	4	6	6	6	8	8
A6	4	0	0	0	0	0	0	0	0	0	0
A7	4	2	4	4	4	4	5	10	10	10	10
A8	7	1	3	4	4	4	1.43	4.29	5.71	5.71	5.71
A9	8	0	0	0	0	0	0	0	0	0	0
Total	53	14	25	26	29	29	25.47	41.32	42.74	48.67	48.67

Table II. Aesthetic domain calculations

artifact universal set. After the WAM artifact comparison, and focus group feedback, management then selects which missing components are to be specifically added to the website, and at what stage. The focus groups include commercial website developers, website marketing executives, motorcycle riders, and academics. Hence, management-guided changes to the website, while introducing subjectivity, do add website components, and each of these changes yield higher WB scores (Tables II and III, and Figures 4 and 5).

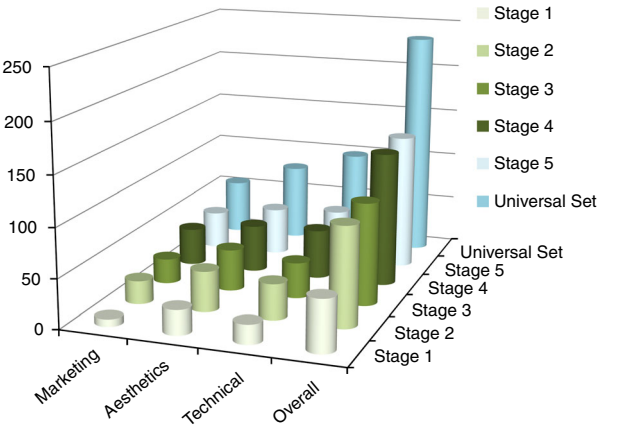
Figure 5 summarizes the above website change process as a timeline against first-time-consumer/user traffic. Stage one's single page "Brochure" shows visitors but soon loses its appeal. Stage two's static five page approach doubles stage ones first-time-consumer/user traffic, however again quickly loses most of its peak consumer market. Stage three's interactive five page approach wins, and likely retains more consumers. It adds the basic motivations of communications, static maps, additional ride details, and a YouTube channel.

After stages one, two, and three, consumer numbers initially peak each time with more components winning more consumers, but with no ongoing change in the website these numbers decline. However, as the 12 weeks per stage study continues an increasing, but minimal number of consumers still continue to visit the website. Consumer traffic shows this is due to the website's growth in the breadth of its domestic and international market reach over time.

Stage four sees the introduction of broad consumer interactivity. Here inward links from appropriate forums and a Facebook site rapidly build the consumer market to a

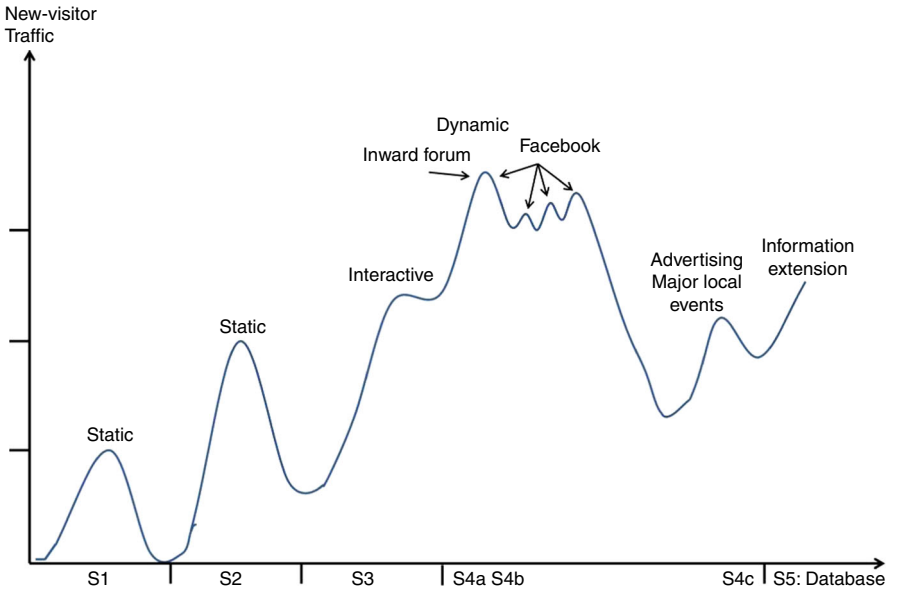
Domain	Stage 1	Stage 2	Stage 3	Stage 4	Stage 5
Aesthetic	25.47	41.32	42.74	48.67	48.67
Technical	19.85	36.59	36.59	50.72	52.66
Marketing	7.57	23.69	25.6	38.33	38.33
Benchmark score	52.89	101.60	104.93	137.72	138.55
Possible score %	18.89	36.28	37.48	49.19	49.48

Table III.
Domain and stage calculations



Source: www.therideguide.com.au

Figure 4.
Effects recognition



Note: www.therideguide.com.au's website changes

Figure 5.
Longitudinal tracking

new peak (more than triple first-time-consumers at stage one). In stage four there is a lot of activity. At 4a traffic growth occurs quickly, however without continued Facebook postings, the traffic again begins to fall. Three subsequent and closely-spaced Facebook postings at 4b generate three small peaks in traffic – with each peak higher than the previous one. This indicates that continual closely-spaced new Facebook postings do continue to draw new consumers to the website. Once again with no additional links, forums or Facebook activity, first-time-consumer numbers again decline rapidly. Thus, social media communication approach can be used to engage with like-minded (or similarly behaving) consumers, and it can be used to drive additional consumer traffic to the website.

The recognizable peak in the data at 4c (prior to stage five update) is produced by visitor "overflow" from several internationally-advertised major sporting events occurring in this study's region. This indicates new consumer traffic can result from regional activities – especially ones that appear to be complementary, even without building behavioral connections (such as showing these consumers that this website provides great regional images and alternate ways to see some regional attractions). Hence, when a major event aligns to the website's target market, the website should cross-promote its complementary offerings toward the major event's consumers, and should pursue winning their positive word-of-mouth comments, thus capturing additional website traffic.

Lastly stage five adds a useful database of accommodation, eateries, and attractions, but without advertising this is slowly sourced by consumers, and so the growth in new visitors is slower but still significant.

Hence, the addition of website components, specifically designed to target a chosen consumer group, results in an increase in first-time-consumer traffic over time. This traffic can be increased when more components and more personal interactivity are

both added to the website. However, when forums, inbound links, and social media (Facebook) are each behaviorally focussed at this website's consumer group, traffic can rapidly increase.

Implications of research

Theoretical implications

This study addresses benchmarking theory, and an online benchmarking gap in the literature. Through the design science paradigm (Fink and Nyaga, 2009) it solves the WB problem using Peffers *et al.*'s (2007) DSRM approach. This DSRM approach incorporates rough-cut (Zhu, 2007) mutually exclusive set-theory (Boisvert and Caron, 2006) and Aristotle's four-step theory-of-causation (Falcon, 2011) with a typological collation of literature-described components – that are point-in-time, multi-level, and objectively (present/absent) based.

Components are likeness-grouped first into functional subsets, and then into the larger domain subsets of the WAM artifact. Thus, by identifying, investigating, and then choosing to add new website components, researchers can further develop the WAM artifact (and its subsets of website components). The artifact can also be used to generate component-recognition algorithms, and researchers can investigate these for specific internal or external industry requirements.

One function represents less than 4 percent of the benchmarking score, and each is of equal weighting. The three domains also equally contribute to the website score. Website function scaling does introduce minor subjectivity, but this approach allows direct comparisons between websites. It also controls subjectivity between functions and their components. Future studies may offer refinements in this area.

Practical implications

A business, benchmarking against its competitors, tends to advance its website only sufficiently to cover its comparative weakness (Kim and Mauborgne, 2005). To be competitive, a corporate's website requires ongoing development, and so must deliver a suitable return-on-investment. However to understand what and how to benchmark first the website must be accurately scored, and then interpreted to competitively draw its consumers into making behavioral decisions regarding the corporate's transaction offerings.

The WAM artifact has a solid theoretical underpinning and holds its existing components within 28 functions and three domains of the website. As each component is programmable a "desired level of completeness" as a WB software solution can be established against the available universal set of components (or against other chosen external websites). This deployable software approach also allows researchers and management to use the WAM artifact components to reposition their website well beyond those of their competition – and into an untapped or "blue ocean" market space (Kim and Mauborgne, 2005). This shift is applied at activity four of the DSRM.

Future research

Measurement aspects

The WAM artifact contains 230 differentiated component measures that currently exist within the literature. As new measures emerge the WAM artifact adapts to include them. This comprehensive collation allows managers to understand WB in-depth and to new levels.

With continual literary updates arising, the universal set of WAM artifact components requires constant monitoring to remove obsolete components, and to add

new components (such as socially interactive measures develop and new social domain), and also to promote the connectivity of certain components when they are deemed as being of higher importance to the consumer. To ensure the WB algorithm remains current researchers must incorporate these alterations into the WAM artifact.

The WAM artifact is designed to permit regular inclusion/removal updates when new literature-identified components emerge; inclusion of new functions is also possible if/when needed. As research into social networking gains momentum the possibility of a fourth domain may occur.

When new disruptive or radical change technologies emerge (Benner and Ranganathan, 2012; Manyika *et al.*, 2013), the WAM artifact approach can also incorporate such technologies by the inclusion of a measurement suite of the newly identified components and even functions. Thus, WAM artifact additions and removals ensure the WB algorithm remains current.

Theoretical aspects

Websites differ in purpose; hence, not every WAM artifact component is necessary, suitable, or required, for every website. As new technologies emerge the components of the universal WAM artifact require updating. This is an ongoing area of research, and of component, function and domain classification considerations. Such research can also incorporate unique industry-focussed component and universal set refinements.

Prior to this study, industry-specific WB remained inadequately and inconsistently researched. This study's WAM artifact approach can now be deployed to provide multi-leveled, industry-specific, targeting (or niche) comparison studies. These are rich areas for future and/or refined WB industry studies.

Management aspects

Previous WB studies typically engage limited numbers of website components, typically concentrating on specific website areas (Kim *et al.*, 2009). Other WB approaches do not deliver timely results – forcing management to implement changes *post hoc* (Pang *et al.*, 2009).

Being adaptable to an industry's target requirements, researchers can use the WAM artifact as an efficient and/or effective, manual and/or software, monitoring tool. This objective WB approach suggests researchers and managers of the corporate may find further new solutions (as scenarios or choice-options). The WAM artifact is also a usable management tool when considering/framing the early development stages of a website.

Conclusion

As a valuable global reach resource the website should be a consumer-connecting window, helping to maintain a corporate's competitiveness. Previous studies have considered various groupings across a combined total of well over 150 component measurements. This study resides within the design science paradigm of information systems (Fink and Nyaga, 2009) and engages the Peffers *et al.* (2007) DSRM framework. However, outside this DSRM study, no previous approach has engaged both a theoretical and a universal set of website components that currently exist within the literature.

The WAM artifact, developed through DSRM, is comprehensive, with a robust literature-basis, combining rough-cut set-theory with mutual exclusivity, and applied at a specific point-in-time. When combined with the theory-of-causation, website components are literature-linked and grouped by presence thus creating function,

domain, and website scores at level which can be rated against the current universal set of WB components. Hence, an in-depth website analysis method (WAM) is delivered.

This DSRM WAM artifact is programmable (present/absent), adaptable (business/consumer specific), and expandable (new component inclusions). It solves the identified WB problem for industry, fills the gap in WB literature, and provides a theoretical framework for comprehensive WB assessments. It can also be further linked to relevant information systems and to management studies.

The WAM artifact's WB scores can be employed, and tactically improved, when pursued in conjunction with theories encompassing planned behavior, motivation, consumption, and user gratification. These approaches offer new behaviorally-emphasized components as new website development combinations which can be selectively developed and placed into appropriate functions and domains as new differentiators against competitive websites.

References

Abascal, J., Arrue, M., Fajardo, I., Garay, N. and Tomás, J. (2004), "The use of guidelines to automatically verify web accessibility", *Universal Access in the Information Society*, Vol. 3 No. 1, pp. 71-79.

Adler, P.S., Benner, M., Brunner, D.J., MacDuffie, J.P., Osono, E., Staats, B.R., Takeuchi, H., Tushman, M.L. and Winter, S.G. (2009), "Perspectives on the productivity dilemma", *Journal of Operations Management*, Vol. 27 No. 2, pp. 99-113.

Ajzen, I. (1991), "The theory of planned behavior", *Organizational Behavior and Human Decision Processes*, Vol. 50 No. 2, pp. 179-211.

Albaum, G., Evangelista, F. and Medina, N. (1998), "Role of response behavior theory in survey research: a cross-national study", *Journal of Business Research*, Vol. 42 No. 2, pp. 115-125.

Albert, T.C., Goes, P.B. and Gupta, A. (2004), "GIST: a model for design and management of content and interactivity of customer-centric web sites", *MIS Quarterly*, Vol. 28 No. 2, pp. 161-182.

Barnes, S.J. and Vidgen, R.T. (2002), "An integrative approach to the assessment of e-commerce quality", *Journal of Electronic Commerce Research*, Vol. 3 No. 3, pp. 114-127.

Beirekdar, A., Vanderdonckt, J. and Noirhomme-Fraiture, M. (2002), "KWARESMI – knowledge-based web automated evaluation with reconfigurable guidelines optimization", 9th International Workshop on Design, Specification, and Verification of Interactive Systems DSV-IS2002, Springer-Verlag, London.

Benner, M.J. and Ranganathan, R. (2012), "Offsetting illegitimacy? How pressures from securities analysts influence incumbents in the face of new technologies", *Academy of Management Journal*, Vol. 33 No. 1, pp. 213-233.

Boisvert, H. and Caron, M. (2006), "Benchmarking web site functions", *Benchmarking: An International Journal*, Vol. 13 Nos 1/2, pp. 174-189.

Bowers, N. (1996), "Weblint: quality assurance for the world wide web", *Computer Networks and ISDN Systems*, Vol. 28 Nos 7/11, pp. 1283-1290.

Calero, C., Ruiz, J. and Piattini, M. (2005), "Classifying web metrics using the web quality model", *Online Information Review*, Vol. 29 No. 3, pp. 227-248.

Cao, M., Zhang, Q. and Seydel, J. (2005), "B2C e-commerce web site quality: an empirical examination", *Industrial Management and Data Systems*, Vol. 105 No. 5, pp. 645-661.

Cassidy, L. (2010), "Website benchmarking: evaluating scaled and dichotomous approaches", unpublished honours thesis, James Cook University, Cairns.

Cassidy, L. and Hamilton, J. (2011a), "Website benchmarking: evaluating scaled and dichotomous approaches", in Rotchanakitumnuai, S. and Kaewkitipong, L. (Eds), *The 11th International Conference on Electronic Business, Borderless E-Business for the Next Decade*, Thammasat University, Bangkok, pp. 408-412.

Cassidy, L. and Hamilton, J. (2011b), "Website benchmarking: theoretical and measurement aspects", in Rotchanakitumnuai, S. and Kaewkitipong, L. (Eds), *The 11th International Conference on Electronic Business, Borderless E-Business for the Next Decade*, Thammasat University, Bangkok, pp. 413-416.

Chang, Y.-S. and Fang, S.-R. (2013), "Antecedents and distinctions between online trust and distrust: predicting high and low risk internet behaviors", *Journal of Electronic Commerce Research*, Vol. 14 No. 2, pp. 149-166.

Chen, Y.-H. and Barnes, S. (2007), "Initial trust and online buyer behavior", *Industrial Management and Data Systems*, Vol. 107 No. 1, pp. 21-36.

Choi, S. and Morrison, A.M. (2005), "Website effectiveness for bricks and mortar travel retailers", *International Journal of Tourism and Hospitality Research*, Vol. 16 No. 1, pp. 63-78.

Cox, J. and Dale, B.G. (2002), "Key quality factors in web site design and use: an examination", *International Journal of Quality and Reliability Management*, Vol. 19 No. 7, pp. 862-888.

Deng, J. (1989), "Introduction to Grey system theory", *The Journal of Grey System*, Vol. 1 No. 1, pp. 1-24.

Deng, L. and Poole, M.S. (2010), "Affect in web interfaces: a study of the impacts of web page visual complexity and order", *MIS Quarterly*, Vol. 34 No. 4, pp. 711-730.

Elling, S., Lentz, L. and De Jong, M. (2007), "Website evaluation questionnaire: development of a research-based tool for evaluating information websites", *Electronic Government 2007 proceedings of 6th International Conference, Berlin, Heidelberg*, pp. 293-304.

eMarketer (2014), "Global B2C e-commerce sales to hit $1.5 trillion this year driven by growth in emerging markets", available at: www.emarketer.com/Article/Global-B2C-Ecommerce-Sales-Hit-15-Trillion-This-Year-Driven-by-Growth-Emerging-Markets/1010575 (accessed June 4, 2014).

Evangelista, F., Poon, P. and Albaum, G. (2012), "Using response behaviour theory to solicit survey participation in consumer research: an empirical study", *Journal of Marketing Management*, Vol. 28 Nos 9/10, pp. 1174-1189.

Falcon, A. (2011), "Aristotle on causality", in Zalta, E.N. (Ed.), *The Stanford Encyclopedia of Philosophy*, available at: http://plato.stanford.edu/archives/fall2011/entries/aristotle-causality (accessed June 2, 2013).

Fang, Y., Qureshi, I., Sun, H., McCole, P., Ramsey, E. and Lim, K.H. (2014), "Trust, satisfaction, and online repurchase intention: the moderating role of perceived effectiveness of e-commerce institutional mechanisms", *MIS Quarterly*, Vol. 38 No. 2, pp. 407-428.

Fink, D. and Nyaga, C. (2009), "Evaluating web site quality: the value of a multi paradigm approach", *Benchmarking: An International Journal*, Vol. 16 No. 2, pp. 259-273.

Flavian, C., Guinaliu, M. and Gurrea, R. (2006), "The role played by perceived usability, satisfaction and consumer trust on website loyalty", *Information and Management*, Vol. 43 No. 1, pp. 1-14.

Gefen, D., Karahanna, E. and Straub, D.W. (2003), "Trust and TAM in online shopping: an integrated model", *MIS Quarterly*, Vol. 27 No. 1, pp. 51-90.

Ghasemaghaei, M. and Hassanein, K. (2013), "Consumers' satisfaction with online information quality: the moderating roles of consumer decision-making style, gender and product involvement", *2013 Proceedings of the 21st European Conference on Information Systems, Utrecht*, pp. 1-7.

Grigoroudis, E., Litos, C., Moustakis, V.A., Politis, Y. and Tsironis, L. (2008), "The assessment of user-perceived web quality: application of a satisfaction benchmarking approach", *European Journal of Operational Research*, Vol. 187 No. 3, pp. 1346-1357.

Hamilton, J.R. (2010), "Rating the website: an objective approach", *International Journal of Electronic Business*, Vol. 9 No. 4, pp. 1-24.

Hamilton, J.R. and Tee, S. (2013), "Understanding social network site consumer engagements", *Proceedings of the 24th ICIS Conference, Melbourne, December 4-6*, Vol. 24, No. 1, pp. 1-10.

Herrero, A. and San Martin, H. (2012), "Developing and testing a global model to explain the adoption of websites by users in rural tourism accommodations", *International Journal of Hospitality Management*, Vol. 31 No. 4, pp. 1178-1186.

Hevner, A. and Chatterjee, S. (2010), *Design Research in Information Systems: Theory and Practice*, Springer, Dordrecht.

Hevner, A.R., March, S.T., Park, J. and Ram, S. (2004), "Design science in information systems research", *MIS Quarterly*, Vol. 28 No. 1, pp. 75-105.

Ho, C.-T. and Wu, Y.-S. (2006), "Benchmarking performance indicators for banks", *Benchmarking: An International Journal*, Vol. 3 Nos 1/2, pp. 147-159.

Hu, Y.-C. (2009), "Fuzzy multiple-criteria decision making in the determination of critical criteria for assessing service quality of travel websites", *Expert Systems with Applications*, Vol. 36 No. 3, part 2, pp. 6439-6445.

Io Storto, C. (2013), "Evaluating ecommerce websites cognitive efficiency: an integrative framework based on data envelopment analysis", *Applied Ergonomics*, Vol. 44 No. 6, pp. 1004-1014.

Io Storto, C. (2014), "Benchmarking website performance in the public sector: a non parametric approach", *Journal of Computers*, Vol. 9 No. 3, pp. 636-643.

Ivory, M.Y. and Hearst, M.A. (2002), "Improving web site design", *IEEE Internet Computing*, Vol. 6 No. 2, pp. 56-63.

Ivory, M.Y. and Megraw, R. (2005), "Evolution of web site design patterns", *ACM Transactions on Information Systems*, Vol. 23 No. 4, pp. 463-497.

Kaplan, R. and Norton, D. (1992), "The balanced scorecard – measures that drive performance", *Harvard Business Review*, Vol. 70 No. 1, pp. 71-79.

Kim, C. and Mauborgne, K. (2005), *Blue Ocean Strategy*, Harvard Business School Press, Boston, MA.

Kim, H.-W., Chan, H.C. and Gupta, S. (2007), "Value-based adoption of mobile internet: an empirical investigation", *Decision Support Systems*, Vol. 43 No. 1, pp. 111-126.

Kim, J., Kim, M. and Kandampully, J. (2009), "Buying environment characteristics in the context of e-service", *European Journal of Marketing*, Vol. 43 Nos 9/10, pp. 1188-1204.

Klandermans, B. (1984), "Mobilization and participation: social-psychological expansions of resource mobilization theory", *American Sociological Review*, Vol. 49 No. 5, pp. 583-600.

Kline, S.F., Morrison, A.M. and St. John, A. (2004), "Exploring bed and breakfast websites: a balanced scorecard approach", *Journal of Travel and Tourism Marketing*, Vol. 17 Nos 2/3, pp. 253-267.

Krosnick, J.A. (1991), "Response strategies for coping with the cognitive demands of attitude measures in surveys", *Applied Cognitive Psychology*, Vol. 5 No. 3, pp. 213-236.

Krosnick, J.A. and Alwin, D.F. (1987), "An evaluation of a cognitive theory of response-order effects in survey measurement", *Public Opinion Quarterly*, Vol. 51 No. 2, pp. 201-219.

Krosnick, J.A., Holbrook, A.L., Berent, M.K., Carson, R.T., Hanemann, W.M., Kopp, R.J., Mitchell, R.C., Presser, S., Rudd, P.A., Smith, V.K., Moody, W.R., Green, M.C. and Conaway, M. (2002), "The impact of 'no opinion' response options on data quality: non-attitude reduction or an invitation to satisfice?", *Public Opinion Quarterly*, Vol. 66 No. 3, pp. 371-403.

LaRose, R. and Eastin, M.S. (2004), "A social cognitive theory of internet uses and gratifications: toward a new model of media attendance", *Journal of Broadcasting and Electronic Media*, Vol. 48 No. 3, pp. 358-377.

Lee, J. and Morrison, A.M. (2010), "A comparative study of web site performance", *Journal of Hospitality and Tourism Technology*, Vol. 1 No. 1, pp. 50-67.

Lee, S.-M., Ungson, G.R. and Russo, M.V. (2011), "What determines an engaging website? An empirical study of website characteristics and operational performance", *Journal of High Technology Management Research*, Vol. 22 No. 1, pp. 67-79.

Lee, Y. and Kozar, K.A. (2009), "Designing usable online stores: a landscape preference perspective", *Information and Management*, Vol. 46 No. 1, pp. 31-41.

Lee, Y., Chen, A.N.K. and Ilie, V. (2012), "Can online wait be managed? The effect of filler interfaces and presentation modes on perceived waiting time online", *MIS Quarterly*, Vol. 36 No. 2, pp. 365-394.

Lee, F.-H. and Wu, W.-Y. (2011), "Moderating effects of technology acceptance perspectives on e-service quality formation: evidence from airline websites in Taiwan", *Expert Systems with Applications*, Vol. 38 No. 6, pp. 7766-7773.

Lin, K.-Y. and Lu, H.-P. (2011), "Why people use social networking sites: an empirical study integrating network externalities and motivation theory", *Computers in Human Behavior*, Vol. 27 No. 3, pp. 1152-1161.

Loiacono, E.T., Watson, R.T. and Goodhue, D.L. (2007), "WebQual: an instrument for consumer evaluation of web sites", *International Journal of Electronic Commerce*, Vol. 11 No. 3, pp. 51-87.

Luo, J., Ba, S. and Zhang, H. (2012), "The effectiveness of online shopping characteristics and well-designed websites on satisfaction", *MIS Quarterly*, Vol. 36 No. 4, pp. 1131-1144.

Luo, M.M., Chea, S. and Chen, J.-S. (2011), "Web-based information service adoption: a comparison of the motivational model and the uses and gratifications theory", *Decision Support Systems*, Vol. 51 No. 1, pp. 21-30.

McKinney, V., Yoon, K. and Zahedi, F.M. (2002), "The measurement of web-customer satisfaction: an expectation and disconfirmation approach", *Information Systems Research*, Vol. 13 No. 3, pp. 296-315.

Manyika, J., Chui, M., Bughin, J., Dobbs, R., Bisson, P. and Marrs, A. (2013), "Disruptive technologies: advances that will transform life, business, and the global economy", McKinsey Global Institute, May, pp. 1-23, available at: www.mckinsey.com

Moriarty, J. (2011), "A theory of benchmarking", *Benchmarking: An International Journal*, Vol. 18 No. 4, pp. 588-612.

Nathan, R.J. and Yeow, P.H.P. (2011), "Crucial web usability factors of 36 industries for students: a large-scale empirical study", *Electronic Commerce Research*, Vol. 11 No. 2, pp. 151-180.

Nel, D., Niekerk, R. V., Berthon, J.-P. and Davies, T. (1999), "Going with the flow: web sites and customer involvement", *Internet Research: Electronic Networking Applications and Policy*, Vol. 9 No. 2, pp. 109-116.

Nusair, K.K. and Kandampully, J. (2008), "The antecedents of customer satisfaction with online travel services: a conceptual model", *European business Review*, Vol. 20 No. 1, pp. 4-19.

O'Cass, A. and Ngo, L.V. (2012), "The antecedents of customer satisfaction with online travel services: a conceptual model", *Industrial Marketing Management*, Vol. 41 No. 1, pp. 125-135.

Olsina, L. and Rossi, G. (2002), "Measuring web application quality with WebQEM", *Multimedia IEEE*, Vol. 9 No. 4, pp. 20-29.

Pang, M.-S., Suh, W., Kim, J. and Lee, H. (2009), "A benchmarking-based requirement analysis methodology for improving websites", *International Journal of Electronic Commerce*, Vol. 13 No. 3, pp. 119-162.

Peffers, K., Tuunanen, T., Rothenberger, M.A. and Chatterjee, S. (2007), "A design science research methodology for information systems research", *Journal of Management Information Systems*, Vol. 24 No. 3, pp. 45-77.

Rashid, M.A.A. and Othman, M.N.A. (2010), "User adoption and use of e-government online service", *Gading Business and Management Journal*, Vol. 14, June, pp. 55-66.

Schwarz, N. (2007), "Cognitive aspects of survey methodology", *Applied Cognitive Psychology*, Vol. 21 No. 2, pp. 277-287.

Simmons, G., Armstrong, G.A. and Durkin, M.G. (2008), "A conceptualization of the determinants of small business website adoption: setting the research agenda", *International Small Business Journal*, Vol. 26 No. 3, pp. 351-389.

Steelman, Z.R., Hammer, B.I. and Limayem, M. (2014), "Data collection in the digital age: innovative alternatives to student samples", *MIS Quarterly*, Vol. 38 No. 2, pp. 355-378.

Stepchenkova, S., Tang, L., Jang, S., Kirilenko, A.P. and Morrison, A.M. (2010), "Benchmarking CVB website performance: spatial and structural patterns", *Tourism Management*, Vol. 31 No. 5, pp. 611-620.

Udo, G.J., Bagchi, K.K. and Kirs, P.J. (2010), "An assessment of customers' e-service quality perception, satisfaction and intention", *International Journal of Information Management*, Vol. 30 No. 6, pp. 481-492.

Vilenkin, N.Y. (1986), *Stories About Sets*, Academic Press, New York, NY.

Wakefield, R.L., Wakefield, K.L., Baker, J. and Wang, L.C. (2011), "How website socialness leads to website use", *European Journal of Information Systems*, Vol. 20 No. 1, pp. 118-132.

Webb, H.W. and Webb, L.A. (2004), "SiteQual: an integrated measure of website quality", *The Journal of Enterprise Information Management*, Vol. 17 No. 6, pp. 430-440.

Weijters, B., Cabooter, E. and Schillewaert, N. (2010), "The effect of rating scale format on response styles: the number of response categories and response category labels", *International Journal of Research in Marketing*, Vol. 27 No. 3, pp. 236-247.

Wells, J.D., Valacich, J.S. and Hess, T.J. (2011), "What signal are you sending? How website quality influences perceptions of product quality and purchase intentions", *MIS Quarterly*, Vol. 35 No. 2, pp. 373-396.

Wolfinbarger, M. and Gilly, M.C. (2003), "eTailQ: dimensionalizing, measuring and predicting eTail quality", *Journal of Retailing*, Vol. 79 No. 3, pp. 183-198.

Xu, J., Benbasat, I. and Cenfetelli, R.T. (2013), "Integrating service quality with system and information quality: an empirical test in the e-service context", *MIS Quarterly*, Vol. 37 No. 3, pp. 777-794.

Yang, T., Linder, J. and Bolchini, D. (2012), "DEEP: design-oriented evaluation of perceived usability", *International Journal of Human-Computer Interaction*, Vol. 28 No. 5, pp. 308-346.

Yoo, B. and Donthu, N. (2001), "Developing a scale to measure the perceived quality of an internet shopping site (SITEQUAL)", *Quarterly Journal of Electronic Commerce*, Vol. 2 No. 1, pp. 31-47.

Yoon, C. and Kim, S. (2009), "Developing the causal model of online store success", *Journal of Organizational Computing and Electronic Commerce*, Vol. 19 No. 4, pp. 265-284.

Zhu, W. (2007), "Topological approaches to covering rough sets", *Information Sciences*, Vol. 177 No. 6, pp. 1499-1508.

Zhu, Y., Basil, D.Z. and Hunter, M.G. (2009), "The extended website stage model: a study of Canadian winery websites", *Canadian Journal of Administrative Sciences*, Vol. 26 No. 4, pp. 286-300.

Further reading

Ivory, M.Y. and Hearst, M.A. (2001), "The state of the art in automating usability evaluation of user interfaces", *ACM Computing Surveys*, Vol. 33 No. 4, pp. 470-516.

Page, C. and Lepkowska-White, E. (2002), "Web equity: a framework for building consumer value in online companies", *Journal of Consumer Marketing*, Vol. 19 No. 3, pp. 231-248.

Corresponding author
Leonie Cassidy can be contacted at: leonie.cassidy1@jcu.edu.au