A Fitness-Utility Model for Design Science Research

T. Grandon Gill and Alan R. Hevner

Information Systems and Decision Sciences, College of Business,
University of South Florida, Tampa, FL 33620
{grandon, ahevner @ usf.edu}

Abstract: Current thinking in design science research (DSR) defines the usefulness of the design artifact in a relevant problem environment as the primary research goal. Here we propose a complementary evaluation model for DSR. Drawing from evolutionary economics, we define a fitness-utility model that better captures the evolutionary nature of design improvements and the essential DSR nature of searching for a satisfactory design across a fitness landscape. We conclude with a discussion of the strengths and challenges of the fitness-utility model for performing rigorous DSR.

Keywords: Design science research, design evaluation, usefulness, utility, fitness, evolutionary economics

1 The Dependent Variable in Design Science Research

Current thinking in design science research (DSR) defines *utility* as the primary research goal (e.g. [1, p. 80]). In this context, the close relationship of utility to practical *usefulness* is emphasized. The choice of usefulness as the pre-eminent dependent variable for DSR ties it to earlier MIS research exploring appropriate dependent variables for information systems [2, 3]. It also establishes a clear relationship between DSR and the influential technology acceptance model (TAM) for information systems, where usefulness plays a pivotal role in motivating use [4]. Given these strong connections to existing well established research streams, does it even make sense to question if usefulness should *always* be our central criteria for evaluating design?

Being contrarians, we do feel the search for the dependent variable in DSR requires some rethinking. Here, we consider a pair of alternative dependent variables: *design fitness* and *design utility*. In the case of fitness, we particularly focus on its biological meaning—the ability of an entity to reproduce itself and evolve from generation to generation. In the case of utility, rather than viewing it as being roughly equivalent to usefulness, we focus on its meaning in fields such as economics and decision sciences, where it serves as the basis for ranking decision alternatives. Naturally, usefulness plays an important role in determining both fitness and utility. Neither of these variables, however, is solely determined by usefulness. Indeed, we believe that understanding the relationship between the three variables via a new fitness-utility model complements current thinking and provides important insights into the nature of design science.

We begin by clarifying the frequently misunderstood concept of the *design artifact*. We then explore the nature of our two proposed dependent variables, fitness and utility, as they are defined in biology, economics, and in the emerging interdisciplinary field of evolutionary economics. Subsequently, we consider how these concepts can be employed in the context of artifacts and designs. We then consider how the guidelines of design science research may be better understood in the context of the fitness-utility model. Finally, the specific benefits and challenges of applying the fitness-utility model for DSR are discussed.

2 Design Artifacts

Central to the notion of DSR is the concept of a design artifact. IT artifacts are broadly defined as constructs (vocabulary and symbols), models (abstractions and representations), methods (algorithms and practices), and instantiations (implemented and prototype systems). [1, p. 77] More generally, artifacts can be viewed as the symbolic representation or physical instantiation of design concepts. Even within a discipline such as MIS, they are not necessarily limited to information systems. Rather, MIS artifacts include organizational designs, process designs, and other intentionally constructed entities relating to information systems.

Conceptually, we can view the design process as a series of layers, as seen in Figure 1. The top layer, the design space, can be viewed as the collection of all possible designs and requirements. Obviously, its contents “exist” in abstract terms only since such a complex design space is infinite. Conceptually, then, we can imagine that the space is partitioned between a few known and many unknown designs. The design process begins with a search of this space in order to identify a particular position, which can be referred to as a *design candidate*.

Between the design space and use artifact layers we find the design artifact layers, of particular significance to DSR. Once a design space candidate has been chosen, we can begin to develop artifacts. As previously noted, these may be symbolic or physical representations of our selected location in the design space. These artifacts may serve a variety of purposes:

1. Providing evidence of design feasibility - Can the proposed design be implemented and does the proposed design meet the requirements? Building feasibility artifacts moves designs across the unknown/known partition.
2. Providing evidence of the value of the design - Does the design offer benefits unmatched by competing design candidates? Here the objective becomes to establish an ordinal valuation that can be used to rank candidate designs.

3. Determining the most effective representation of the design – How can we best communicate the intricacies of the design to the implementators (e.g. architects, programmers).
4. Constructing the actual use artifacts - A blueprint is a construction artifact that serves to guide the physical construction of a house; source code is a construction artifact that serves to generate the programs that are distributed to users.

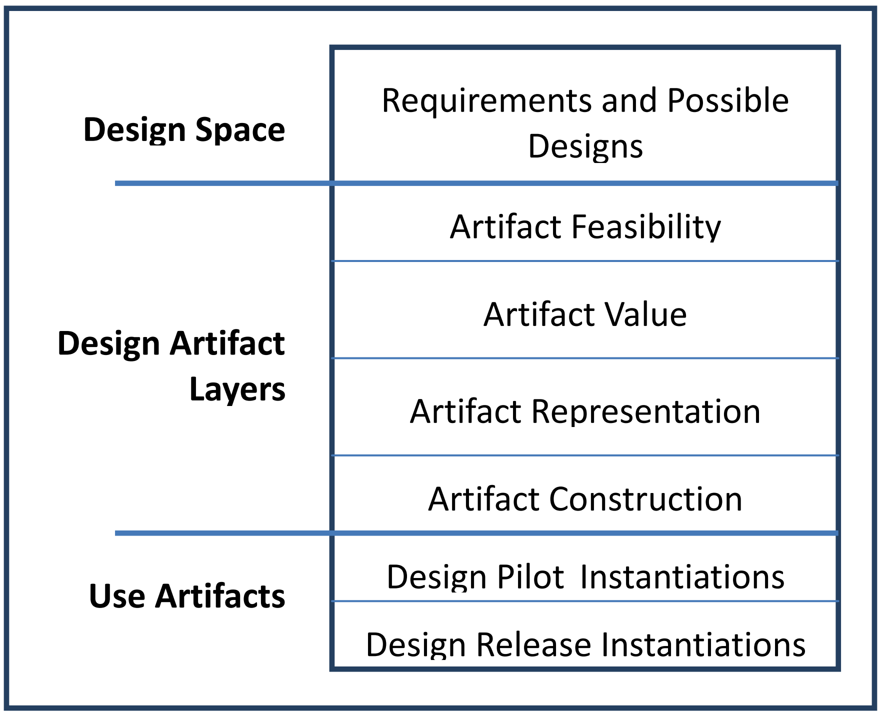

Fig. 1. Design Artifact Abstract Layers

The use artifacts are divided between pilot test instances—for which returning to the design cycle is intentionally left open as a possibility—and release use instances, for which further redesign is not anticipated. While this conceptual scheme obviously maps directly to IT artifacts such as software, it should be recognized that organizations frequently employ a phased roll out of non-technology artifacts, such as organizational structures or incentive plans, with the same notion that the design may later be tuned based upon early experience.

The particular significance of design artifact layers to DSR stems from their nature. As noted previously, the design space itself is too amorphous to be investigated directly. We need a physical representation or symbolic description of a particular design candidate—in other words, an artifact—if we are to conduct meaningful research. The investigation of use artifacts, on the other hand, is largely the domain of behavioral research. Inasmuch as they have already been constructed, the principles incorporated in their design are likely to be of less interest than the principles determining how their use impacts the entities (e.g. organizations) in which they are embedded. Nevertheless, it is certainly possible—indeed probable—that important principles that may guide future design can be acquired by observing constructed instances in use. This highlights the complementarity and need for communication between design science and other research paradigms.

3 Fitness and Utility

Based on our understanding of the layers of the design artifact, we can now move to an exploration of how better to understand and evaluate the artifact in DSR. Two concepts from other disciplines for this task are fitness (biology) and utility (economics).

1 Fitness

To understand fitness, it is useful to begin by proposing two alternative definitions of the fitness of an organism:

Fitness Definition #1: The fitness of an organism describes its ability to survive at a high level of capacity over time.

Fitness Definition #2: The fitness of an organism describes its ability to replicate and evolve over successive generations.

Which definition of fitness you prefer likely depends on your perspective. If the individual in mind were our personal physician, we would strongly prefer he or she focus on definition #1. Terms such as physical fitness, mental fitness and emotional fitness all correspond to this general class of definition. If, however, the individual were an evolutionary biologist, definition #2 would be overwhelmingly preferred. An organism lacking the capacity to reproduce and evolve rapidly goes extinct. What is important about the distinction between definition #1 and definition #2 is that their outcomes are not necessarily correlated. This is graphically illustrated by the experience with human populations, as discussed in Example 1.

Example 1: Two Versions of Fitness in Populations

At the end of the 18th century, Thomas Malthus proposed that any increases in the individual fitness (definition #1) of human populations would lead to a rapid increase in reproductive rate (a contributor to definition #2) that would quickly erase the gains in individual fitness and would, in the long run, reduce individual fitness since gains in food supplies tended to be arithmetic whereas changes in reproductive rates tended to be geometric [5, p. 6]. What has actually happened, however, is in stark contrast to predictions. After a period of adjustment, as individual fitness increases, evolutionary fitness (definition #2) has actually declined.

To illustrate this phenomenon, it is useful to consider two measures: life expectancy (a proxy for definition #1) and fertility rate (a proxy for definition #2). In an organism that employs sexual reproduction, fertility rate represents the number of children each female of the species produces over her lifetime. In human populations—where the number of male babies is slightly higher than the number of female babies—a stable population requires a value is slightly over 2. In much of the industrialized world, this value has fallen far below that stable value. For example, the 2006 U.N. Economic and Social Affairs agency estimated Japan's 2000-2005 fertility rate at a shockingly low 1.29. During the same period, the U.S. had an estimated value of 2.04. Based on definition #1, the fact that Japan has the highest life expectancy in the world among major industrialized nations would imply high fitness. With respect to definition #2, on the other hand, such low birth rates suggest a population that is decidedly unfit from an evolutionary standpoint.

We will henceforth *always* refer to the second definition of fitness when we use the unqualified term. There are two reasons for this. The first is that the population-focused view of fitness is generally more sensible when long term systems, such as information systems, are studied. Second, as we shall see later, it would relatively easy to treat variables such as system use or usefulness as a proxy for fitness according to definition #1, implying that little benefit is likely to be derived from advocating definition #1 in place of currently popular dependent variables. Because of the tension between definitions #1 and #2, already noted in the population example, we would expect that examining definition #2 might offer new insights.

Prior to leaving the subject of fitness, it is useful to introduce a model used by evolutionary biologists, that of the fitness landscape. Such a landscape represents a functional relationship between individual attributes, such as specific genes, and the fitness of an organism. For example, if an organism's fitness were determined by N attributes, x_1 through x_N, its fitness landscape would be described as: Fitness = $f(x_1, x_2, \ldots, x_N)$. Fitness landscapes change over time as a result of forces such as the organisms' collective impact on the environment, the impact of co-evolution of other organisms, and the impact of unpredicted events that occur entirely outside of the systems being studied, popularly referred to as black swans [6].

2 Utility

Similar to fitness, the term utility is used in a number of ways. When we consider the utility of a tool, we are normally referring to its usefulness. As currently used in the context of DSR, that is the prevailing meaning. Hevner, et al. [1, p. 83] state: "The utility, quality, and efficacy of a design artifact must be rigorously demonstrated via well-executed evaluation methods." This implies utility to be a characteristic of the design and its intended application context.

Economists, on the other hand, employ the term utility in a different way. Specifically, they posit each individual to have a utility function that can be used to rank choices in the context of decision-making. The assumption that individuals seek to maximize utility is, in fact, foundational to the field of economics. In early economic theory, the assumption was made that utility was determined by current consumption. More recently, however, it has been generally recognized that many factors contribute to economic utility beyond direct consumption, such as relative income, expectations, social context, and goals [7].

To distinguish between the two usages of the term utility, we will refer to the first as *usefulness*. In this context, we apply the broadest meaning of the term—including factors such as efficacy in performing the task (including performance), range of task cases performed, ease of use, ease of learning, and cost-benefit in the performance of a task. Essentially, we assume that any artifact characteristic that impacts task performance directly can be classified under the usefulness category. Presumably, if our choice of a tool was dictated strictly by usefulness, as just defined, then there would be little reason to distinguish between the two meanings of utility. When we employ the term utility in the rest of the paper, however, we assume its economic meaning and further assume that it represents a complex function that is not adequately described by the single usefulness dimension.

3 Evolutionary Economics

Evolutionary economics is a field that examines economic systems from the perspective of evolution. As it happens, the foundational assumption of the field ties the notions of fitness and utility together. At the risk of oversimplifying, the basic concept that drives the field is that, as humans, our utility function has evolved as a response to the fitness landscapes we have faced and, as a consequence, is tuned towards maximizing fitness. The rationale is stated as follows by Gandolfi et al. [8, p. 97]: "Given the logic of natural selection, it is difficult to conceive how, for any living entity, a preference for maximizing fitness could fail to evolve."

The argument is based upon the fact that, on a static landscape, high fitness individuals will tend to crowd out lower fitness individuals. Because fitness landscapes are themselves subject to change (as noted previously), traits that promote diversity—e.g. an urge, on the part of some individuals, to seek out new peaks—are also likely to survive over time in some percentage of a population, described as *evolutionarily stable strategies* (ESS). The percentage of a population described by a particular niche ESS may grow after sudden shifts in fitness that increase in the value of the strategy, while niche strategies may well decline in percentage during long periods of stability in which highly visible high fitness peaks draw an increasing portion of the population. Regardless of where an entity exists on the fitness landscape, however, utility will tend to drive it toward local peaks.

A sensible argument can be made that our current utility preferences do not map well to fitness. Numerous researchers have demonstrated that we, as human beings, are far from rational in our processes of choice [9]. There are a number of ways to respond to this. First, evolution is slow and—particularly over the past 250 years—changes in the environments we face as a consequence of the industrial and information ages have been so rapid that it would be inconceivable that our utility preferences could have kept up. Fortunately, our built-in genetic utility function also imbues most of us with a desire to learn and, as a consequence, our utility function can adapt to our changing environment through that mechanism, as opposed to natural selection. Second, it is actually very rare that we encounter tasks with such well-defined inputs (e.g. where probabilities are fully known) in our day-to-day life. Indeed, it is often the case that when we attempt to quantify such values in order to make our decision making more precise, we are in fact deluding ourselves. Third, even individuals who have done extensive research into the "irrationality" of our decision rules acknowledge that there are many contexts where these decision rules prove to be beneficial [10].

The key point here is that utility can be treated as the mechanism by which we make choices when confronted with a fitness landscape. Obviously, it is not perfect. Rather, it represents an estimate-of-fitness that we can apply to make decisions presented by such a landscape. With this principle in mind, we turn to its specific application in the domain of DSR.

4 Relationship of Fitness, Utility, and Usefulness in Design Science

The concepts of fitness and utility can readily be applied to the design of systems. If we revisit Figure 1, it should be evident that the design space is an example of a fitness landscape, with each design candidate being an entity that can be located on that landscape. Design artifacts perform two key roles in the design search process:

1. They provide evidence that a particular design candidate is feasible, has value, can be effectively represented, and can be built. This serves to help us better understand the shape of the design fitness landscape, moving combinations from the unknown to the known category.
2. Through careful evaluation, they provide a basis for choosing between alternative designs.

The first of these directly impacts our knowledge of design fitness. The second refines our estimate-of-fitness that is a basis for choice; it therefore involves changing our utility function through learning.

Where design systems differ from biological evolution is in the role played by intentionality. The mechanisms of evolutionary change—such as production of new gene combinations through sexual reproduction and mutation—are posited to exert their influence with considerable randomness. While survival rates serve to cull the low fitness organisms from the population, the actual construction of such organisms is unguided. In the design space, on the other hand, designers intentionally concentrate on areas of the design fitness landscape where promising candidates have been identified. What that means is that while utility serves as an estimate-of-fitness for design artifacts, it also feeds back into the fitness landscape itself since a low fitness evaluation for a particular design candidate will discourage further investigations into nearby regions of the design landscape. This, in turn, reduces the fitness of those regions since placing less effort into building artifacts based on a particular design will necessarily reduce the flow of future artifacts based on that design (which is how we define fitness). Moreover, the shape of the utility function is likely to be guided by two forces: the nature of the evaluation artifacts being studied *and* by actual experience from artifacts developed for use. Thus, the experience of artifacts placed in practice has the ability to impact the design fitness landscape just as evaluation artifacts do. Thus the new fitness-utility model can re-frame DSR as follows:

The goal of DSR is to impact the design space so as to ensure a continuous flow of high fitness design artifacts. This impact is accomplished in two ways: through the production of artifacts that demonstrate the feasibility of new designs and through improving the utility function that we use to assess the fitness of evaluation artifacts.

This definition, of course, represents a type of artifact, one where fitness and utility replace potential usefulness as dependent variables. What we shall now do is to identify ways in which this approach differs from our prior understanding of DSR. We do this by examining the seven DSR guidelines proposed by Hevner et al. [1].

5 Fitness and Utility Goals in DSR Guidelines

How does the fitness-utility approach to DSR differ from the existing paradigm? In this section we concentrate on how fitness defined in terms of reproductive efficacy and utility defined in terms of a choice frontier alter our perspective on DSR.

1 Guideline #1: Design as an Artifact

The fact that DSR is constrained to deal with the concrete by this guideline is important in distinguishing it from behavioral research. Moreover, the original guidelines are sufficiently broad in their definition that meta-design findings (e.g. a list of attributes that contribute to design quality) would, itself, constitute an artifact and would therefore—quite rightly—fall under the DSR heading. Under the fitness-utility approach, the term “produce” would be too limiting, however, since research leading to changes to the design utility function would fall under the approach. Thus, research can radically change the design space without necessarily producing a design artifact. A radical restatement of this guideline would be as follows:

Guideline #1: The objective of the fitness-utility model of DSR is to impact the design space through the creation and evaluation of design artifacts.

2 Guideline #2: Problem Relevance

While we agree with the continued importance of relevance, the problem with the existing statement of guideline #2 is that, from a practical standpoint, it tends to constrain the time horizons for design research. We often cannot foresee what problems will be relevant for the future of IT. The challenge this unpredictability presents to DSR is that if you try to anticipate the important long term problems that a design will solve, it will be nearly impossible to get them right. Thus, being overly problem-focused demands a shorter term outlook. Another way of looking at the issue is to use the analogy of constructing a puzzle. At the beginning of a puzzle, as in a design process, you have a collection of pieces that can only be put together in certain ways. True “problems”—in the form of missing pieces--tend to be discovered near the end of the assembly, when the gap is identified. If we require that our design science solve an important problem, we may need to wait until we know what is missing. With this caveat, we do not recommend any changes to DSR Guideline #2.

Guideline #2: The objective of DSR is to develop technology-based solutions to important and relevant business problems.

3 Guideline #3: Design Evaluation

The fitness-utility model recognizes a large number of characteristics that could potentially be used to assess design fitness. These are illustrated in Figure 2. The area within the fitness ellipse outside of the intersection with the usefulness ellipse reflects characteristics that can impact fitness that are not a direct a result of usefulness (although they may be correlated with it). Those characteristics listed in Figure 2 are intended to serve as an incomplete list of examples that will now be discussed. We begin, however, by revisiting usefulness.

As illustrated in Figure 2, the potential usefulness of a design artifact still plays a key role in assessing fitness, as it did in the original model. What the figure also suggests, however, is that there may be times when a design artifact becomes so useful that it actually inhibits further improved designs—much the way increased life expectancy (i.e. fitness definition #1) has become associated with below replacement fertility rates (i.e. fitness definition #2). In fact, the tendency of organizations to stick with designs that have proven useful is a well-documented phenomenon known as the Innovator's Dilemma [11].

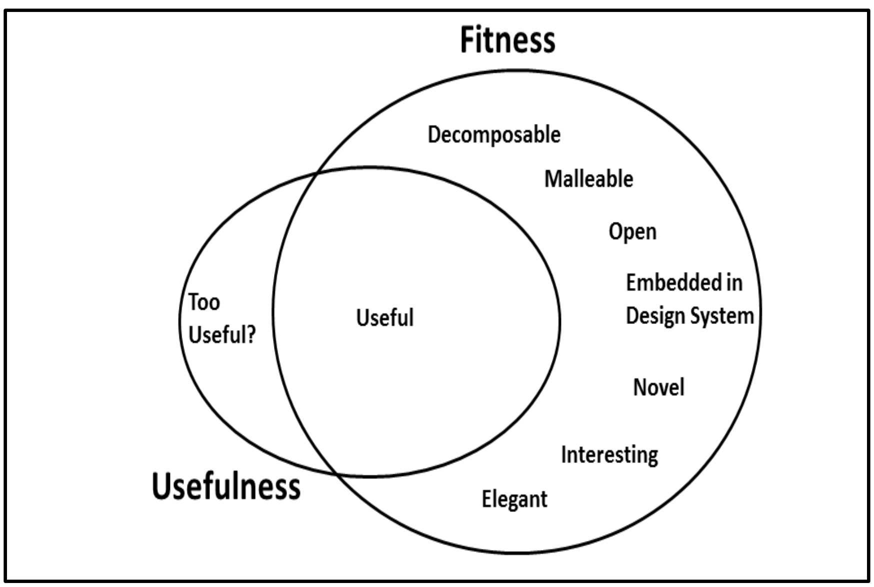

Fig. 2. Design Candidate Fitness Characteristics and Usefulness

The other key characteristics of artifact fitness are briefly discussed here:

Decomposable. The seminal work that launched the study of design science is Herbert Simon's *The Sciences of the Artificial* [12]. The second half of the book is largely devoted to explaining why systems tend to evolve from nearly decomposable subsystems. Indeed, even under the existing design science goals, decomposability is likely to exert a strong influence on design quality and would therefore be evaluated as part of the design. In addition, such systems tend to be easier to construct, since work on individual components can be conducted separately. The particular difference that the fitness-utility approach would engender involves the reproduction and evolution of partial designs. Where a design cannot be decomposed into nearly independent subsystems, evolution of the design would tend to be a matter of all-or-nothing. Where a design is built upon separable systems or constructions, on the other hand, pieces of the design—strands of *design DNA* to use a biological analogy—may exhibit high fitness and evolve rapidly while others may remain static or be discarded.

Malleable. Related to decomposability, the malleability of an artifact represents the degree to which it can be adapted by its users and respond to changing use/market environments [13, 14]. MIS research has demonstrated that users frequently employ tools for unintended purposes. We would expect that such adaptation would allow designers to evolve artifacts to support these uses more effectively.

Open. Another characteristic that has the potential to impact design fitness is the degree to which artifacts are open to inspection, modification, and reuse. Openness tends to encourage design evolution by making it easier both to see how an artifact is designed and to modify existing components of the artifact. For example, an information system created as an open source application has a significant advantage over a proprietary design in terms of its ability to evolve rapidly.

Embedded in a Design System. We would expect design artifacts that are the product of a sustainable design system environment to evolve more rapidly than artifacts that are produced in a context where design is an unusual activity. This particular source of fitness can sometimes act as a counterweight to openness, as organizations with highly effective research and development activities may be reluctant to open up their designs and may use legal measures—such as patents and copyrights—to discourage unauthorized parties from evolving the original designs. An effective design system can produce a stream of design artifacts, however, even without the financial rewards that comes from transforming these into use artifacts.

Novelty. A design may be considered novel if it originates from an entirely new region of the design space. Once such a design candidate has proven viable, other design candidates from the same region are likely to follow in an attempt to locate the local peak on the fitness landscape. A particular challenge that novel design artifacts present is that the creative process through which they are envisioned may not meet the criterion of rigor suggested by the original guideline and the potential benefits of the design may be hard to evaluate.

Interesting. Normally, a design artifact is created in order to explore or demonstrate some specific purpose. From time-to-time, however, an artifact may demonstrate unexpected emergent behaviors that are worthy of subsequent investigation and the creation of subsequent artifacts. Social scientists (e.g. [15]) have long asserted that research which largely conforms to existing expectations yet also incorporates an unexpected element is most likely to interest other researchers.

Elegant. In many areas of design, such as architecture, consumer products and apparel, there is an ongoing tension described as form versus function. Function relates to practical usefulness. Form, in contrast, describes aesthetic elements such as appearance that do not necessarily serve a useful purpose, yet nevertheless increase the user's utility. The characteristic of an MIS design artifact that corresponds to form might best be referred to as *elegance*. Like quality, elegance is hard to define in a rigorous manner and yet characteristics that might be associated with it—such as compactness, simplicity, transparency of use, transparency of behavior, clarity of representation—can all lead to designs that invite surprise, delight, imitation, and enhancement.

If the fitness-utility approach is taken to DSR, then the evaluation criteria are where the utility function is to be shaped. Further thinking and research are needed in order to propose methods for formulating the utility model in specific DSR projects. Thus, we would require a restatement of the original design evaluation guidelines along the following lines:

Guideline #3: The fitness of a design artifact must be estimated using a utility function that considers the full range of characteristics that can impact the likelihood that the artifact will further be reproduced and evolve.

5.4 Guideline #4: Research Contribution

With respect to this guideline, the fitness-utility approach and the original approach are relatively similar. As originally stated, however, it is not clear that research that leads to better understanding of utility (i.e. estimating the fitness of a design artifact) would be included under the design heading. For this reason, a preferable rewording might be:

Guideline #4: Effective DSR impacts the design space through contributions in the area of the design artifact, design fitness, design foundations and theories, and/or design methods.

5.5 Guideline #5: Research Rigor

A particular challenge associated with the use of the term rigor is that it is perceived to be generally "understood" but is rarely defined. One definition that has been proposed (e.g. [7]) treats research rigor as consisting of three related elements: 1) the investigation is systematic, 2) a thoughtful balance is struck between the risk of accepting that which is false (Type 1 error) and rejecting that which is true (Type 2 error), and 3) challenging questions are posed. By this definition, the current guideline would tend to place considerable obstacles in the way of early stage design artifacts, inasmuch as: i) systematic search of the design space is generally impossible, ii) current standards of empirical research in the social sciences tend to lean heavily towards avoiding Type 1 error [16] making rejection of novel ideas more likely, iii) early stage design artifacts often leave challenging questions—such as scalability and relative benefits compared to alternative designs—largely unanswered. Rather than abandoning rigor altogether, the guideline could be revised as follows:

Guideline #5: DSR requires that the construction and evaluation of design artifacts be investigated employing a level of rigor appropriate to the nature and stage of design.

5.6 Guideline #6: Design as a Search Process

There is little need to change the spirit of this guideline, which captures perfectly the process of search in a fitness landscape. A slight modification to the wording is desirable, since the fitness-utility model assumes we are searching for high fitness artifacts in a design space.

Guideline #6: The search for high fitness design candidates and artifacts requires utilizing available means to reach desired ends while satisfying laws in the design space.

5.7 Guideline #7: Communication of Research

This guideline once again illustrates the preference for late-stage design research in the original conception of DSR. Management-oriented audiences, in particular, are unlikely to be impressed by designs whose usefulness has not been demonstrated. The fitness-utility approach would take an entirely different perspective. Where the goal is to exert impact on the design space (which is a fitness landscape), what makes sense is to target those communities most likely to initiate the next iteration of the design process through supplying resources, which would naturally include time, intellectual effort, facilities, and money.

Guideline #7: Design research must be communicated to those communities most likely to supply the resources required for future design using communication channels appropriate to each community.

6 Discussion - Pros and Cons of the Fitness-Utility Model

With its focus on reproductive fitness (i.e. definition #2) rather than individual artifact fitness (i.e. definition #1), the fitness-utility model offers both strengths and weaknesses when contrasted with the existing DSR paradigm, which is why we view it as a complement rather than as a competitor to the existing approach. We discuss five advantages of the new model followed by three challenges.

6.1 Makes the Researcher an Active Participant in the Design System

Because developers often publish research relating to the artifacts they are creating, a great deal of design research in IT is already action research. Under the fitness-utility model, however, even the non-technical researcher strives to play an active role in the design system through impacting fitness values in the design space. Successful research will, as a matter of definition, lead to either an increase or decrease in the production of new artifacts based upon the specific design candidate or candidates investigated. The fitness-utility model would also be predicted to maximize the potential impact of individual research contributions by focusing on early stage design. Thus, if the researcher's goal is to impact the design space, consistent with the goals of the fitness-utility model, the earlier the artifacts evolving from a particular design candidate can be identified, the better.

6.2 Provides an Alternative Basis for Evaluating Research Impact

Today, within the MIS research discipline, the impact of research is generally measured through the estimated quality of the publication outlet and through subsequent citations by other researchers. The fitness-utility approach offers another alternative: chart the evolution of subsequent artifacts contrasted with the findings of the research. If the artifact continues to evolve and incorporate design DNA deemed favorable by the research, then impact—in the truest sense of the word—has been achieved. The same can be said of research that stifles the further evolution of design DNA deemed detrimental to fitness. For example, if particular design practice (e.g. allowing the user to enter free form text into a textbox that is then used to query a database) leads to a security threat (e.g. malevolent SQL injection), impactful DSR that identifies this as a low fitness practice should reduce the frequency of the occurrence in later artifacts.

6.3 Aligns with Dynamic Environments

A central premise of this paper is that over time the evolutionary fitness of design artifacts becomes far more interesting than the use fitness of a particular artifact. The validity of this premise is likely to depend on the environment in which it is situated. For very static environments, for example, a particular use artifact may exist for a very long time. In such a world, the use fitness of the artifact is a matter of considerable interest. In a highly dynamic environment, on the other hand, the artifact's potential to evolve needs to be given much greater weight. Our belief is that such dynamism describes most environments facing IT designers today, and that forces such as globalization, social media, and advances in telecommunications will likely serve to increase environmental turbulence.

6.4 Recognizes the Inherent Limitations of Intended Usefulness

Our research suggests that while usefulness is likely to be the best single predictor of artifact use (a finding consistent with most TAM research), it is not a necessarily a very good predictor when applied by itself. In fact, a reasonable argument can be made that many of the most *interesting* (see [15]) findings of MIS revolve around examples where an IT artifact's impact was far different from the designer's intended use.

6.5 Encourages Collaboration between MIS Researchers and Designers in Other Fields

The fitness-utility approach specifically targets clients in the design communities supplying the resources necessary for further design evolution. In early stage IT design research these communities will likely contain a preponderance of researchers in technical fields such as computer science including many academics. Thus, we will have a strong incentive to collaborate with these communities if we are to exert impact. Where we may be able to contribute is in our understanding of the potential unintended consequences of artifacts employed in an organizational setting, as previously described. Having observed these consequences in the field and studied them in our literature, we are in a unique position to provide perspective to designers who may otherwise become overly focused on intended use.

6.6 Current Research Standards Do Not Reward Design Impact

Given that researcher rewards, including promotion and tenure, tend to be closely tied to measured research impact based on numbers of quality publications and citations, the fact that the fitness-utility model offers another approach to measuring impact over time—tracing how artifact design DNA changes as a consequence of research findings—may not be appealing to academic researchers. In the absence of institutional change with respect to how impact is defined, it may be hard for the fitness-utility model to gain traction.

6.7 The Framework for Evaluating Design Fitness Is Not Well Researched

Earlier in the paper, we proposed a number of non-use characteristics (Figure 2) that seemed likely to impact design fitness. This list was largely inducted from examples and could in no way be considered complete, rigorously derived, or rigorously supported. Unfortunately, there is little research into the characteristics that provide good estimates of design fitness as we have defined it. Stated another way, our design utility function is largely unexplored. This naturally presents a substantial obstacle to any research that attempts to estimate the fitness of a particular artifact. The largely unexplored forces driving fitness and utility are in stark contrast to the much better established approaches to evaluating design usefulness. Although the field laments its lack of theoretical base, constructs, and generalizability [1, p. 99], it has a plethora of these when contrasted with the fitness-utility model. Add to this the fact that immediate usefulness is likely to seem a more concrete research objective than fitness, and the researcher is likely to have a much easier time designing research under the existing paradigm.

6.8 Building Rigor for Fitness-Utility Research Requires Alternative Research Methods

It may be argued that the last challenge actually represents a considerable opportunity for future research into the factors that lead to fitness. Such research, however, is likely to substantially differ in character from the main body of existing MIS research. To understand fitness, you need to look backward in time in order to trace the evolution of an artifact. Indeed, it may take years to validate the actual fitness of an artifact—a necessary step if the characteristics contributing to fitness are to be identified. Thus, historical research methods are likely to play a much greater role that is the case in most contemporary MIS research (see [17]). In addition, fitness landscapes in general tend to be rugged, meaning that interdependencies between variables prevent decomposability. Such ruggedness can confound traditional statistical techniques. What this means is that data analysis techniques most preferred by MIS researchers may prove largely inapplicable in the analysis of sources of fitness.

7 Conclusions

Several times in this paper, we have posited that the fitness-utility model for design science research is better viewed as a complement to the existing usefulness model, rather than as a competitor. As illustrated in Table 1, which summarizes the analyses presented in this paper, the two models focus on different objectives, are most applicable to different artifacts, tend to examine different units of analysis, are appropriate for different time horizons, are likely to employ different research methods, and will tend to be of greatest interest to different client constituencies. We have already noted that high levels of usefulness may actually inhibit artifact evolution. There is likely to be quite a bit of causality here—an organization making a large investment in designing and deploying a use artifact is unlikely to view the tendency to evolve rapidly as a major benefit. To the contrary, such a manager is most likely to appreciate an artifact that is highly useful and is likely to remain that way as long as possible. For that clientele, maximizing the fitness of design, as we have defined it, is more likely to be more scary than desirable. Moreover, an understanding of the factors contributing to usefulness is central to the fitness-utility model. Many factors outside of usefulness may contribute to fitness, but we expect usefulness will typically prove to be the single most important factor in most design settings.

Nevertheless, we believe the fitness-utility model for DSR is too important to ignore. It is our strong belief that an artifact that continues to evolve will *always* end up outperforming an artifact that fails to evolve, regardless of their respective usefulness at the time they were conceived. This is the core of Christensen's innovator's dilemma [11] and if we do not recognize this process, we are ignoring a major force that shapes today's competitive environment. As we have pointed out, however, such research is likely to adhere to different guidelines (Section 5) and depart considerably from existing DSR practices. It is our goal in this paper to alert researchers and reviewers of these differences and offer some justification as to why they are necessary. In doing so, it is our hope to stimulate future DSR thinking along the lines of the fitness-utility model. We hope to advance these ideas further by describing case studies and performing DSR projects in which evaluation is based on the fitness-utility model.

Acknowledgements. The authors are grateful for the insightful comments by the anonymous reviewers that led to improvements in the paper.

References

1. Hevner, A., March, S., Park, J., Ram, S.: Design Science in Information Systems Research. MIS Quarterly**.** 28(1), 75-106 (2004)

2. DeLone, W., McLean, E.: Information Systems Success: The Quest for the Dependent Variable. Information Systems Research. 3(1), 60-95 (1992)
3. DeLone, W., McLean, E.: The DeLone and McLean Model of Information Systems Success: A Ten-year Update. Journal of Management Information Systems. 19(4), 9-30 (2003)
4. Venkatesh, V., Davis, F.: A Theoretical Extension of the Technology Acceptance Model: Four Longitudinal Field Studies. Management Science**.** 46(2), 186-204 (2000)
5. Gill, R., Glazer, N., Thernstrom, S.: Our Changing Population. Prentice Hall, Englewood Cliffs, NJ (1992)
6. Taleb, N.: The Black Swan: The Impact of the Highly Improbable. Random House, New York, NY (2007)
7. Gill, T.: Informing Business: Research and Education on a Rugged Landscape. Informing Science Press, Santa Rosa, CA (2010)
8. Gandolfi, A.E., Gandolfi, A.S., Barash, D.: Economics as an Evolutionary Science: From Utility to Fitness. Transaction Publishers, New Brunswick, NJ (2002)
9. Ariely, D.: Predictably Irrational: The Hidden Forces that Shape our Decisions. Harper Perennial, New York, NY (2007)
10. Ariely, D.: The Upside of Irrationality: The Unexpected Benefits of Defying Logic at Work and at Home. Harper Collins, New York, NY (2010)
11. Christensen, C.: The Innovator’s Dilemma, Harper Collins, New York, NY (1997)
12. Simon, H.: The Sciences of the Artificial (3^{rd} Edition). MIT Press, Cambridge, MA (1996)
13. Gregor, S., Jones, D.: The Anatomy of a Design Theory. Journal of the AIS. 8(5), Article 2, 312-335 (2007)
14. Williams, K., Chatterjee, S., Rossi, M.: Design of Emerging Digital Services: A Taxonomy. In Hevner, A., Chatterjee, S. (eds.) Design Science in Information Systems Research. Springer, New York, NY. 235-253 (2010)
15. Davis, M.: That’s Interesting! Towards a Phenomenology of Sociology and a Sociology of Phenomenology. Philosophy of the Social Sciences. 1, 309-344 (1971)
16. Ziliak, S., McCloskey, D.: The Cult of Statistical Significance. University of Michigan Press, Ann Arbor, MI (2008)
17. Mokyr, J.: The Gifts of Athena: Historical Origins of the Knowledge Economy. Princeton Univ. Press, Princeton, NJ (2002)

Table 1. Summary of Usefulness and Fitness-Utility Models

Characteristic	Usefulness Model	Fitness-Utility Model
Focus	Useful artifacts	Artifact reproduction and evolution (fitness) and the choice mechanisms guiding artifact design (utility)
Applicable artifacts	Construction and use	Feasibility and evaluation
Unit of study	Entire artifact	"Design DNA" within artifact
Time horizons	Short and medium-term	Long term
Source of rigor	Careful evaluation of intended use and expected performance	Systematic evaluation of non-usefulness factors that may contribute to fitness and the potential for unintended consequences
Most likely external (non-MIS research) clients	Developers and use clients	Researchers outside of MIS and R&D clients
Source of models	Study of current artifacts in the field	Study of historical progression of artifacts based upon a particular design candidate
Particular value offered by MIS research	Understanding the organizational context in which artifact development and use takes place	Understanding the role played by unintended consequences in typical artifact implementation; broad perspective on factors that influence artifact success
Desired impact of research	Improved design and development of useful artifacts and better understanding of the factors that make an artifact useful	Improving fitness of desirable design DNA and suppression of undesirable strands; better understanding of the factors that increase real-world artifact fitness leading to improved choice between alternative design candidates.