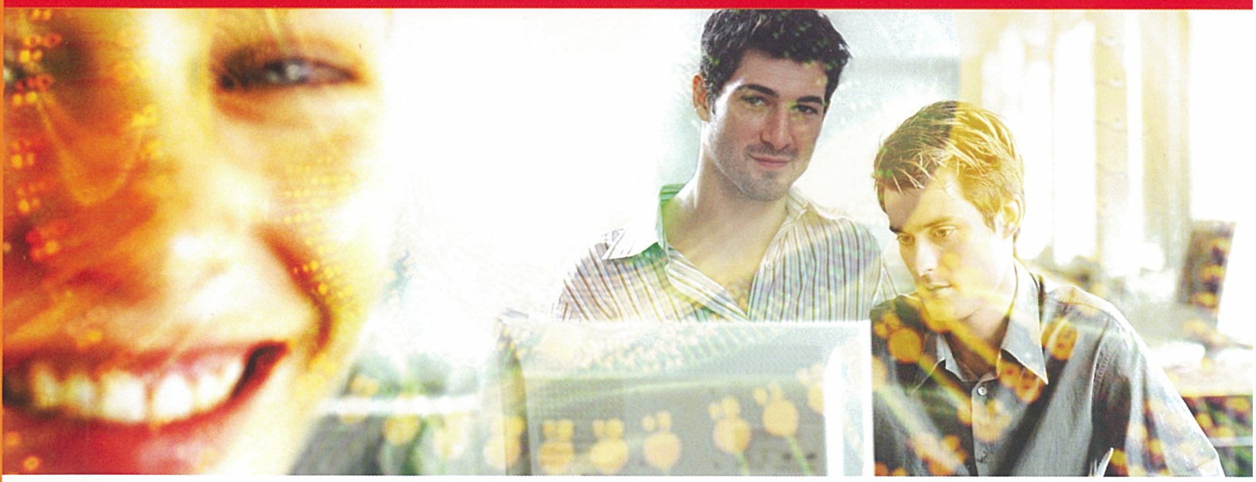
3TU.School for Technological Design

STAN ACKERMANS INSTITUTE

Criteria for assessing a technological design

Kees van Hee and Kees van Overveld

March 2010



TU Delft
Delft University of Technology

TU/e Technische Universiteit Eindhoven University of Technology

UNIVERSITY OF TWENTE.

Criteria for assessing a technological design

Kees van Hee and Kees van Overveld

March 2010

0.Foreword

This report has been finalized by the Steering Committee SAI on March 1st 2010 after elaborate discussions with the management of the 11 design programs within the Stan Ackermans Institute, 3TU.School for Technological Design. The Steering Committee consists of prof.dr. K.M. van Hee, prof.ir. H. Leegwater, prof.dr.ir. M.J.W. Schouten, prof.dr.ir. A.W.M. Meijers, prof.dr. H.J.P. Timmermans, ir. P.L.J. Swinkels, prof.dr.ir. J.I.M. Halman en mr.drs. B.C. Donders.

1. Introduction

In the context of the Graduate Schools of TU/e, TUD and UT, to be installed, the position of the PDEng designers programs in the Stan Ackermans Institute (SAI) is subject for reconsideration. At TU/e, the opinion holds that the designers programs will be part of the graduate programs. This means that designing will take a prominent position, next to research and education. An important aspect of the foreseen integration will be the quality assurance of PDEng programs, both regarding the educational part and the design assignment. Education and research are presently covered by regular surveys[1]; CCTO certification deals with the PDEng programs. One issue is the quality of design assignments. Therefore the need is felt to have clear and assessable criteria for design assignments. Various sets of criteria have been proposed in the past, but these need updating. Here, the authors concentrate on criteria for assessing the *design* rather than the design *process* or the *course elements* in PDEng programs. For assessing the full PDEng programs, all three elements are crucial. For instance, a sound design could have been produced where the trainee got significant help from the supervisor. In this case, the design may receive a high ranking whereas the process is ranked less.

Design criteria can be used for many different purposes. First to distinguish valid design assignments from invalid ones. This is important in selecting and formulating design assignments – prior to their actual execution. One crucial aspect is, that there should be a clearly identifiable *artefact* that is to be designed or re-designed. Second, the criteria can be used in assessing the design assignment after *completion.* As an interesting side-effect, criteria will help shape the actual *realization* of design assignments since, during the design process, both the designer and the supervisors will aim for high scores in the upcoming assessment.

As a fortunate coincidence, parallel to the present exercise , criteria are sought in 3TU and KNAW-context to help identify designing as a scientific activity. In particular when considering a PhD graduation on design[2] (formally allowed since the 1908 Higher Education Law!) such criteria are essential.

[1] Dutch: 'visitaties'

[2] Dutch: 'promotie op Proefontwerp'

PhD graduation on design is primarily a TU/e endeavour. It is sad that Technical Universities in general seem to be reluctant to accept designing as an independent scientific activity.

Universities have three prominent responsibilities: *education*, *research*, and *knowledge valorisation*. For *technical* universities, designing technological artefacts is an excellent way to treasure knowledge. Furthermore, designing is an essential vehicle for research: designing technological artefacts leads to discoveries unlikely to occur otherwise. Newly developed principles can only be validated by application in real designs. Designing forms the *experimental* component of the constructing sciences; a generally applicable, innovative artefact is a scientific result in its own right!

Apart from their role in PDEng programs and PhD graduations on design, the criteria to be presented can act in the assessment of scientific staff. Indeed: when classifying some forms of design as scientific activities, it is essential for measuring their scientific outputs to have operational criteria. This implies documenting and registering designs. Part of this may be effectuated by publishing in the standard media for scientific output or by patenting; in some cases this is rendered impossible due to secrecy considerations. A repository of designs, reviewed by independent experts, containing public summaries and review reports could provide a solution.

Both ranking scientific work and scientists by means of quality criteria meets with growing interest. At present, however, such rankings rarely address the qualitative contents of scientific work. Rather, ranking takes place using *derived* features of scientific labour. The first derivative is the *number of publications* in renowned journals or proceedings; the second derivative is the number of *citations* of these publications. A third derivative, the *Hirsch-index* is gaining popularity. Although such methods may facilitate ranking scientists, it is questionable if the actual quality of scientific work is measured.

For our purposes, such attempts are less interesting. We want to be able to judge whether an individual design (both *ex ante* and *post ante*) has sufficient level. A two-point scale {“Yes”,”no”} could suffice. For instance to label a completed design assignment with ‘cum laude’, however, we need more room for nuances. Therefore we propose ordinal scales sampled on 5 values each (see Appendix).

The present report aims to contribute to the development of clear and operational criteria for design assignments, primarily in the context of PDEng programs, secondly for use in PhD graduations on design, and thirdly for assessing design as the fruit of scientific work.

To develop these criteria, we must address the following questions:

- What *kinds* of *artefacts* are being considered?
- What comprises the *design* of such an artefact?
- How is such an artefact *documented*?
- Which *indicators* represent the quality of a design?

We address these issues in the sequel of this report. First we review relevant literature; we conclude with an Appendix containing scales for all proposed indicators. These scales are mere examples: for specific disciplines other choices are possible, but we recommend ordinal scales that are as objective as possible.

As mentioned before, despite the importance of judging the *design process* as well, this is not addressed in the current report.

2. Relevant literature

We consulted 5 sources:

1. Report of the Royal Academy of Engineering
2. CCTO Regulations
3. SAI TU/e Handbook
4. Schouten Report
5. Literature on ‘designing as scientific activity’

A report of the Royal Academy of Engineering, entitled ‘Measuring Excellence in Engineering Research’ [1] appeared in 2000, but this does not specify requirements for a scientifically valid design.

Similarly, it is remarkable that the CCTO regulations [2] lack a description of a design assignment. The *tasks* of a designer are described, though, and in this context we read that designing is about ‘technical solutions for *products* and *systems*, starting from functional requirements and market demands, fitting with the general societal context (attention for environment, safety, re-usability, et cetera)’. This is very much in line with the SAI Handbook [3]. The goal of the program is formulated there as ‘technological solutions in either (1) products and constructions, (2) processes for the manufacturing of such products, (3) systems for transport of persons, information and goods, and (4) management systems both for production and transport’. And: ‘this includes the design of instruments (…)’. This formulation shows signs of struggling with a precise demarcation of the field. Requirements to the *development* of the design are articulated, but requirements to the design proper are lacking.

The SAI Handbook contains criteria for PhD graduation on design. Although a design for PhD-defence is assessed more strictly, the criteria listed form a useful start for the present purpose.

‘The design for PhD graduation includes:

- The description (in any adequate format) of the designed artefact (product or process);

Any artefact is described in morphological, functional and procedural terms. The morphological *description regards the construction of the artefact as an assembly of components, the artefact as a part of its environment. The* functional *description addresses the specification of the artefact in terms of its performance, its effectiveness. The* procedural *description refers to the relation of the artefact and the actors contributing to the decision making process, and the manufacturability of the artefact in the form of an instruction.*

- The description of the initially planned design process, and its actual realization;

The realized design process is described, where the designer logs all design steps and decisions. Considerations regarding possible alternatives need to be recorded to facilitate a scientifically founded judgement on the final solution. The generation of alternative solutions, optimisation, and a possible simulation to support choices need to be described. (Assessors may realize, however, that design may be assessed differently from scientific research with respect to the aspect of reproducibility.)

- (if applicable) the description of the research, carried out to obtain new knowledge needed for realizing the design, as well as the description of the results of this research.

It is also possible to publish performed scientific research and its results via the standard routes (in scientific journals, et cetera). For the realization of the design, it is only required that necessary knowledge is available – irrespective whether this was obtained by new research or otherwise."

This last point is of no concern in a PDEng assignment.

Further, the SAI handbook gives requirements regarding scientific standards: *intersubjectivity* (interpretation of the artefact must not depend on the interpreter), *reliability* (the artefact must function the same under different circumstances, provided that these circumstances fall within the envelope of the predetermined specifications) and *assessability* (defined as 'the description of the artefact in terms of the rationales for its decisions'). Finally, there are (somewhat vague) requirements regarding novelty and completeness, and the PhD defence proper.

The SAI Handbook explains how a design is to be described – albeit in the context of a PhD graduation on design. This is a valid starting point, though, therefore we copy it for our purposes:

In addition to a specification or analysis of the design assignment and a report of the assessment of the design, the following 12 *aspects* or *properties* of the design need to be described::

1. *Functional* properties of the artefact (such as working, effectiveness).
2. Properties regarding the *manufacturability* of the artefact.
3. *Non-functional* properties of the artefact (such as expected lifetime, reliability).
4. *Economical* properties (such as marketability).
5. *Societal* properties (including control and management).
6. *Morphological* properties (actually the architecture of the artefact).
7. *Static and dynamical* properties (regarding changeability of the artefact).
8. *Organisational issues related to the artefact* such as production, use, maintenance and disassembly.
9. *Scientific relevance* such as the application of existing knowledge (including: *which* existing knowledge) and the inspiration to new knowledge.
10. *Aesthetic* properties: the artefact as a work of art.
11. *Ecologic* properties: relations between the artefact and the environment.
12. *Historic* en *geographic context*: relation to other cultural entities in society.

Some of these aspects are somewhat vague; others overlap or may be difficult to make operational; not all apply to each artefact. Nevertheless, the list is a rather complete enumeration of relevant aspects..

A more recent, and more concrete approach is found in the Prof. Jeu Schouten Report (May 2009 edition) [4]. This identifies two groups of quality aspects: *scientific quality* listing 8 aspects, and *societal quality* listing 4 aspects.

Scientific quality

1. Intersubjectivity (as explained above).
2. Assessability (as explained above).
3. Knowledge increase: the design contributes to scientific knowledge in the relevant area.
4. Effectiveness: the design efficiently satisfies the design requirements.
5. Concreteness: the design can be presented, e.g. as prototype or drawing.
6. Relation to research: the design involves the integration of design, development and scientific research.
7. Methodical approach.
8. Novelty: demonstrable innovation, originality, and inventiveness.

Societal quality

1. Manufacturability.
2. Economic benefit.
3. Social and cultural relevance.
4. Professionalism.

It seems possible to provide instruments for measuring these criteria - but this has not yet been done. Furthermore, these criteria are somewhat of a mixture of criteria for the method of working (e.g., ‘methodical approach’), the assessment of the result (‘intersubjectivity’), and the artefact proper.

In several scientific disciplines, designing is studied as a scientific activity. We give some examples. First, the classical study of Herbert Simon [5] devotes a chapter to design science. Another example is [6], where design is compared with natural science and behavioural science; a design is seen as the proof of a solution for a given problem. Designing is depicted as an exploratory voyage, discovering new problems and assessing which methods work and which don’t. The paper gives guidelines for such research. Our deliverables and aspects for assessment are coherent with its findings.

A third example is [7], where a philosophical analysis of artefacts is presented. Furthermore, there is extensive literature on measures for complexity for artefacts, in particular for software systems.

In [8], the design process is analysed, focussing on the distinction between art and skill. Also, [5] devotes a chapter to complexity. We come back to the issue of measurements for complexity in the Appendix, when we present scales for our indicators.

The subsequent table – several versions of which exist – summarizes the distinction between research and design.:

Aspect	**Research**	**Design**
Questions	Why?	How?
Starting point	Empirical data	Requirements
Result	Knowledge or theory	Artefact
In search of	Truth	Value
Thinking terms	Invariants	Variants

3. What is a technological design?

First and foremost we want to define what it is that is being designed. This we call an *artefact*. An artefact can be either a *product* or a *process* that brings forward products. A product can be a physical product, but also a software product or even a service. An artefact can be a living organism: similar to the design of molecules on the basis of specifications in chemistry; micro-organisms can be designed that meet with certain requirements.

Artefacts as we see them serve an economical or societal purpose. We are interested in artefacts that rely on scientifically founded technologies, and whose quality can be analysed methodically. We give 13 examples to sketch the range of what we consider as artefacts.

(1) artificial materials, (2) machines, (3) instruments, (4) (chemical) installations, (5) built structures, (6) vehicles, (7) complex consumer products, in particular using electronics , (8) software systems, (9) networks for telecommunication, (10) assembly or production facilities, (11) logistic organisations, (12) complex controllers, procedures or protocols (13) a bacteria that is capable of consuming waste and transform this into harmless or even beneficial substance, as well as the process to grow such bacteria.

So we do not consider paintings, musical pieces, sculptures, advertisements and film scripts. We also exclude technological and organisational consultancy. (Obviously a design can be used as consultancy, but not every consultancy is a design).

Designing a law could fall under our definition, provided that formal proof of the consistency with other legislature is given, and provided that the effects of the new law are quantifiable. In that case designing a law falls under (12). Computer games are also included, provided that predetermined formal specifications exist; in that case they are in instance of (8).

Our concern is the *design* of such an artefact.

A design can occur in various modalities. One is an abstract representation of the artefact in the form of a *symbolic* model – either an *informal* model (text, images) or a *formal* model (mathematical model). Another is a *physical model* in the form of a *prototype*. A third is a *computer model*.

In all these cases the design serves the following purposes::

- *Communication* between stakeholders (commissioners, users, developers, …).
- *Analysis* of the properties of the artefact-to-be-developed. These properties often relate to the behaviour of the artefact in its foreseen context. The analysis can regard the (measurable) *performance* and the *conformance*, that is: showing the presence or absence of certain behavioural features of the artefact. Prototypes and computer models are used for *experimental* analysis of such behaviour. Formal models are used for mathematical analysis, i.e., formal proofs, about the behaviour.
- *Construction* of the artefact. In that case the design is a blueprint for the production of the artefact, perhaps to aid mass production.
- *Documentation* of the artefact, for instructing users, installation, maintenance and future modification.

When designing an artefact, we usually deal with a collection of models, including the various modalities. Safeguarding consistency among these models (models shall not contradict each other) is an art in itself. Models are certainly not realized all at once. A design process is characterized by much trial-and-error, leading to iterations, plagued by incomplete and inconsistent models. Once the design is completed, however, there has to be a collection of consistent models that together give an adequate and complete account of the artefact. 'Complete' means that enough information is present to assess and realize the artefact. Notice that a design assignment often regards only *part* of an artefact.

An important category of design assignments relates to *re-design*. Re-design is often initiated by the observation that an existing artefact no longer meets current requirements – which may differ from the ones at the time of the original design. Identifying the part of the artefact that needs replacement or modification may be far from trivial. Theoretically, design therefore can amount to a minor alteration that is very difficult to spot.

From now on, 'design' will mean the collection of models together representing the artefact. The activity leading to these models is called modelling, which is a form of synthesis.

4. Which are the relevant aspects of an artefact and a design?

An important condition for a design assignment in a post-masters program holds that there shall be sufficient scientific and technological level. Sometimes an artefact is designed as a solution for an entirely new situation; more often there is an existing artefact that shall be improved, to meet use of new technology, or to meet with changed user requirements. Most often, similar artefacts exist, which allows comparison with the state-of-the-art.

The following aspects of an artefact need consideration:

1. *Functionality*: Which are the functions to be fulfilled by the artefact, and how effective shall it be? Most often, these requirements are initially vague; the role of the artefact in its context is usually described in a merely global way. Together this forms the set of requirements. To a large extent, the designer determines the functionality in the form of specifications, staying within the envelope determined by the requirements. In case of a re-design, finding the part that needs adjustment may form the main challenge.
2. *Impact*: What is the economical and societal relevance of the artefact? Which revenues are expected, and for whom? What is the societal purpose of the artefact.? Which risks are implied by the production, use and disassembly of the artefact? In what respect does the artefact contribute to sustainable society?
3. *Realisation*: Can the artefact be constructed and deployed in an efficient manner? This regards quantities such as cost of development and use (total cost of ownership). This is usually expressed in money – an other possibility is *resource consumption* such as energy.
4. *Inventiveness*: To what extent is the solution novel? ‘Novel’ may mean the deployment of a novel technology, or an innovative combination of existing technologies. In both cases, there can be the case of a *creative invention*; it can also be a trivial compilation of existing elements. Inventiveness is therefore partially determined by the complexity of the artefact.
5. *Complexity*: Designing a complex artefact requires the knowledge of methods and techniques from various disciplines. A truly complex artefact, in general, will be only realizable by a design team.

6. *Elegance*: Any artefact usually has aspects of aesthetics – although this does not have to relate to its visual appearance. The structure or architecture of a design can enchant due to its simplicity, transparent subdivision of components, use of symmetries, et cetera..
7. *Genericity*: To what extent does the chosen solution allow generalisation to other artefacts? This does not always have to be the case. A high score on this aspect is a necessary condition for PhD graduation on design, or the recognition as scientific output, though. Genericity is an important condition for knowledge increase.
8. *Methodical approach*: Has a scientifically valid method been followed, including state-of-the-art techniques for design? Contrary to research, where experimenting with the latest craze is part of the game, design is all about the final result. So proven technology is sometimes preferred instead of the latest new technology, although it could also be a goal of the project to test and improve the latest technology. A solid design method at least contains a requirements engineering step, which gives ample freedom to the designer. Further, it contains steps for checking the design for internal consistency of the various models involved, and checking the requirements that should be met by the artefact. These checks can be embodied by formal techniques (so called verification) or informal and experimental techniques (so called validation).
9. *Documentation* and *presentation*: Is the description of the artefact sufficient to check that the design has been carried out according to the rules; are the models sufficient to demonstrate essential features of the artefact?

The 12 criteria for quality in the Schouten report can be found in these 9 aspects of design, with the exception of ‘relation to research’. Complexity and elegance don’t occur in the Schouten report.

The subsequent table gives the relation between the Schouten report and our aspects.

Aspects of design vs. Schouten Report	Functionality	Impact	Realisation	Inventiveness	Complexity	Elegance	Genericity	Methodical approach	Documentation and presentation
Intersubjectivity									
Assessability								X	X
Knowledge increase				X			X	X	
Effectiveness	X								
Concreteness									X
Relation to research									
Methodical approach								X	
Novelty				X					
Manufacturability			X						
Economic benefit		X							
Social and cultural relevance		X							
Professionalism								X	

5. How is a design to be documented?

Documentation of a design is crucial, both for assessment and for knowledge transfer regarding construction or use of the artefact. We assume the following deliverables to be available as documents::

1. Statement of the assignment, including the requirements.
2. Contextual analysis (a.k.a. domain analysis), explaining the context for which the artefact is intended, and which requirements stem from the context and from the stakeholders. This is an essential document with a re-design.
3. Primary functional specifications of the artefact as a whole, answering the questions ‘what does the artefact do, and how does it do it?’
4. Secondary functional specifications of the artefact as a whole. These regard maintainability, reliability, and others. These answer the question *how* the artefact fulfils the primary functional specifications.
5. Architecture, in the form of models representing the components, their connection, and their (primary or secondary) functional specifications. Components may contain other components, in which case their architecture needs to be given as well, up to the level of atomic components of which the construction is a given fact.
6. Prototypes, demonstrators and simulation models that serve in analyses.
7. Analysis of the artefact. This answers the question to the extent in which the artefact satisfies (extra-) functional specifications. This regards both conformance and performance. An analysis can be either formal, informal or experimental.
8. Plan for realisation. This explains how the artefact is to be constructed or manufactured.
9. Plan for implementation. This explains how the artefact shall be installed in its context, how it is to be maintained, and how it will be finally disassembled.
10. Impact analysis. This comprises an economical analysis and a societal analysis, showing the costs and revenues of realisation and implementation. Also the artefact’s role in society, its benefits and risks are outlined.
11. Methodical underpinning of the design. This explains to what extent state-of-the-art methods and techniques were used, and how the artefact compares to existing, similar artefact.

12. Process report. This explains how the design process took place. In particular the (mis)match between planning and realisation; the way contingencies were dealt with, how design choices came about, and the role of stakeholders in the process.

These deliverables show a reasonable match with the deliverables according to the SAI Handbook (numbers refer to items in the Handbook).

SAI Handbook	1	2	3	4	5	6	7	8	9	10	11	12
Statement of the assignment												
Contextual analysis											X	X
Primary functional specifications	X											
Secondary functional specifications			X									
Architecture						X				X		
Prototypes, demonstrators and simulation models												
Analysis of the artefact							X					
Plan for realisation		X						X				
Plan for implementation					X			X				
Impact analysis				X							X	
Methodical underpinning of the design									X			
Process report												

The next section describes the development of indicators for the 9 aspects.

6. How to evaluate a design?

For each of the 9 aspects of quality we give one or more indicators. Indicators should be equipped with a 5-point scale. In the Appendix, we give an example scale for each indicator. Sometimes a particular indicator does not apply, or cannot be measured. For instance if the design assignment only involves part of an artefact. This is no disqualification of the design; in this case the indicator is marked as 'not applicable' (NA).
Indicators are intended to assist the evaluation of the aspect to which they belong. We allow room for adding other indicators. Apart from its indicators, the aspect *as a whole* should also be scored on a 5-point scale:

1. very bad
2. bad
3. marginal
4. good
5. very good

In practice, evaluation for each aspect will be restricted to levels 3, 4, and 5 since the designer will only present his work if all aspects are at least acceptable. Therefore, the following qualifications can be used as well: 3=below expectations, 4=meets expectations, and 5=exceeds expectations.

There will always be a commission of experts to evaluate the design aspects, partially based on the indicators. Evaluation regards the entire *design*. It is important to observe that the overall quality of a design can not be fully determined by the designer: the nature of the assignment as given in the requirements has influence as well. Therefore, the *design process* should also be taken into account when evaluating a *design assignment* – but this falls beyond the scope of the present report.

We recommend to address the criteria in the design programs' curricula to make prospect designers familiar with relevant aspects of a design. By using these criteria, a best practice will develop that perhaps may lead to future adjustments of the criteria.

The assessment form looks as follows:

Project name	
Designer	
Company	
Company supervisor	
University supervisor	

Aspect	Indicator	Value	Judgement
Functionality	satisfaction		
	argumentation		
	ease of use		
	ease of administration		
Impact	damage per incident		
	chance of incident		
	sphere of impact		
	economic value		
	sustainability		
	Impact on well-being		
Possibility of realisation	technical		
	economical		
Inventiveness	indicator of time		
	indicator of surprise		
Complexity	structural complexity		
	functional complexity		
	reduction of complexity		
Elegance	identified features		
	argumentation		
Genericity	re-usability of artefact		
	re-usability of best practice		
Methodical approach	methodical modelling		
	methodical analysis		
Documentation	systematic organisation		
	completeness		
	correctness		
	accessibility		

1. Functionality

We start from a description of the *context* and the *requirements.* Context occurs in two varieties: an environment where a new artefact should be realized (green field), or a context containing an existing artefact that needs modification or extension (brown field). The second case regards a re-design. This could entail the optimization of an existing system, where a new configuration for parameters needs to be designed. Generally, the brown field variety requires identification of components to be modified or replaced, which can form a significant part of the design project. The requirements indicate the problem definition.

Functionality, among other things, regards the extent to which the artefact will satisfy the requirements, and the degree of certainty that this indeed will be the case. Pre-assumptions that are negotiated with the commissioner (preferably prior to the project) are acceptable. Such pre-assumptions will be assumed to hold at the time of assessment of the designed artefact. In the case of brown field design, the crux of the design can be in the localisation of the part to be re-designed.

The above leads to the following two indicators:

(1) S*atisfaction*, i.e., the extent to which the artefact satisfies the requirements. These requirements need to be formulated such that mere inspection of the artefact shows their fulfilment. So, a requirement like ‘it needs to be as light as possible’ is considered as an optimality criterion which is regarded as inacceptable.

(2) *Proof*, i.e., the degree of certainty that the artefact indeed fulfils the requirements and optimality criteria. Here two elements play a role: the extent to which verification took place, and the degree of rigor of the verification – being either *informally, empirically* (i.e., *using statistics* or *simulation*)*,* or *formally*.(i.e., a mathematical or logical proof).

When giving the proof, pre-assumptions can be made about availability or manufacturability. The validity of these pre-assumptions is assessed under the criterion ‘possibility of realisation’.

Other relevant indicators for functionality are the ease of use of the artefact. This refers to the external interfaces. The ease of use is characterized by the required effort (time) to learn to operate the system: the less the better. An artefact, in principle, has two types of users: operators and administrators. Each may have their own interface.

(3) *Ease of use* regards the utility of the artefact as defined by the primary functional specifications.

(4) *Ease of administration* regards configuring and maintaining the system, as determined in the secondary functional specifications.

2. Impact

Impact comes with three dimensions: the amount of damage resulting from malfunctioning of the artefact (to be called *dependability*); *economical value* and *societal value* . Dependability is characterized by three indicators:

(1) The *damage per incident*, where an incident is the failure of the artefact. This regards direct damage rather than follow-up damage (although this distinction may be difficult to make).

(2) The *chance of an incident* per year.

(3) The *sphere of impact* regards the number of people that depend on the artefact.

(4) The *economical value* is determined in comparison with existing solutions. The compared quantities include *expected cash value* (net present value, r_1) of the profit of the artefact in the present situation, and the expected cash value of the total *cost of ownership* in the present situation, c_1. All profits and costs of ownership are calculated over the total expected lifetime of the artefact.

(5) The *sustainability* regards the extent to which the ideal of a sustainable society was taken into account in the design. This regards (re-) use of materials and energy in production, use and disassembly of the artefact. Also the impact of the artefact on a sustainable environment is taken into account.

(6) The *impact on wellbeing* regards the effect of the artefact on people: does it bring mere material comfort, or can it improve the quality, or extend the length of human life.

3. Possibility of realisation

This involves two indicators: the *technical* and *economical* possibility of realisation.

(1) The *technical possibility of realisation* regards the certainty that the artefact can actually be realized. Absolute certainty exists if a 0-series has been manufactured and tested in practice. On the other extreme end of the scale, there are only models and unsubstantiated assumptions.

(2) The *economical possibility of realisation* can only follow from an economical analysis. In the case of an industrial product, an estimate of market prospects is necessary. For a unique artefact, we need the estimated profits. Reasons for failing realisation can be: an excessively high up front investment, or lacking sources of financing. The other extreme end of the scale is a commissioner-approved business case.

4. Inventiveness

Two indicators apply to this aspect:

(1) A *time indicator* is the amount of time elapsed since the most recent technological break-through necessary for the present artefact: the longer ago, the more inventive.

(2) A *surprise indicator* represents the extent to which various groups will be surprised about the findings in the design.

5. Complexity

Most artefacts comprise of *basic components* that are either available, or that can realised straightforwardly (i.e., using the standard methods of the discipline). Basic components are not decomposed any further. Next to basic components, we identify compound components., built up from basic components or other, more elementary compound components. Each component has one or more *interfaces*. An interface is a relation between two components, or the component and the exterior of the artefact. An interface, in various disciplines, can take various forms. It may represent the exchange of information, energy or momentum or some physical coupling.

Complexity, to a large extent, is determined by the number of components and their mutual interfaces. Also, interfaces can have varying complexity.

Complexity is a complicated aspect. On the one hand, a design should be as simple as possible; on the other hand, we appreciate the design of an inevitably complex artefact more than the design of a simple artefact. For assessment it is necessary to quantify complexity. To this aim, we propose two indicators: *structural* complexity and *functional* complexity.

(1) The *structural complexity* regards the number of components and the number of interfaces between them. Complexity of basic components is implicitly represented by their number of interfaces.

(2) The *functional complexity* regards the interfaces of the artefact with its environment: the more, the higher the complexity.

Both structural and functional complexity are often inherent to the assignment. Unnecessary complexity prompts for lower appreciation; unavoidable high complexity will lead to higher appreciation.

Next to these there is a third indicator to valuate how the designer attempted to reduce complexity. This typically is achieved by clever hierarchical decomposition: which compound components are chosen for the design. We call it

(3) *Reduction of complexity*: it regards the hierarchical decomposition in the design, that is: the subdivision into compound components leading to a more transparent and more comprehensible design. The more complexity can be reduced, the better.

When dealing with re-design (e.g., optimisation) of an existing artefact, complexity regards the relevant part of the artefact that is involved in the re-design.

6. Elegance

Elegance is determined by inspecting the *architecture of the artefact.* Architecture should be transparent and comprehensible. We distinguish *features of beauty.* These can differ significantly over the disciplines.

Some examples are: using symmetry, homogenous and logical distribution of functionalities, substantial re-use of components, visual design of the artefact and its interface, and structural simplicity. Each of these features should be justified.

This justification should be indicated by the designer in the documentation. Assessors are free to accept or reject the validity of an argument. Alternatively, assessors can give arguments to discuss features of beauty.

We distinguish two indicators:

(1) The *number* of identified features of beauty.
(2) The *quality* of the arguments for these features.

So it is possible that a designer mentions many features of beauty, but gives weak justification; alternatively, the designer doesn't mention features of beauty, but these are identified by the assessor. In both cases, the first indicator could receive a high value whereas the second indicator receives a low value.

7. Genericity

Genericity regards the extent to which the designed artefact can be re-used, and the extent to which a *best practice* has been developed that can be applied in different situations. There are two indicators:
(1) *Artefact re-usability;*
(2) *Best-practice re-usability.*

8. Methodical approach

This regards the use of methods and techniques for *modelling* and *analysing*. There are two indicators:
(1) Methodical modelling
(2) Methodical analysis

9. Documentation

This both regards the documentation in the form of various deliverables, and a presentation in the form of a talk, a movie or a demonstrator. Four indicators are used:

(1) *Systematic organisation*
(2) *Completeness*
(3) *Correctness*
(4) *Accessibility*

Appendix: example scales for indicators

Here we give, for each indicator, a 5-point scale. These scales, on the one hand, are proof-of-concept, that is: they are an argument for the possibility of constructing scales for our indicators. On the other hand, scales are prototypes that, when necessary, can be modified by assessors to meet the specific features of a project at hand. Although the scales presented here have been applied successfully to some 10 different designs, wo foresee that ‘one size fits all’ might be too much to hope for at present.

1. Functionality

(1) *Satisfaction:*

1. does not satisfy at all
2. satisfies poorly
3. satisfies marginally
4. satisfies well
5. satisfies completely

(2) *Proof:*

1. has not been verified
2. has been partially verified using informal methods only
3. has been fully verified, but part of it using informal methods and part of it using empirical methods
4. has been fully verified, part of it using empirical methods and part of it using formal methods
5. has been fully verified using formal methods

(3) *Ease of use:* the amount of time needed to learn how to use the system:

1. ≥ 1 month
2. 1 week to 1 month
3. 1 day to 1 week
4. 1 hour to 1 day
5. < 1 hour

(4) *Ease of administration*: the amount of time needed to learn how to administrate the system:

1. ≥ 1 month
2. 1 week to 1 month
3. 1 day to 1 week
4. 1 hour to 1 day
5. < 1 hour

2. Impact

(1) *Damage per incident:*

1. app. € 1
2. app. € 100
3. app. € 10.000
4. app. € 1.000.000
5. app. € 100.000.000 or more

(2) *Chance of incident per year*

1. app. 10^{-8}
2. app. 10^{-6}
3. app. 10^{-4}
4. app. 10^{-2}
5. app. 10^{-1}

(3) The *sphere of impact,* measured in terms of number of affected people:

1. app. 10
2. app. 100
3. app. 10.000
4. app. 1.000.000
5. app. 100.000.000 or more.

(4) The *economical value*: $q=((r_2-r_1)-(c_2-c_1)) / (|r_2-r_1|+|c_2-c_1|)$; r_1 = profit in present situation; r_2 = profit in new situation; c_1 = cost of ownership in present situation; c_2 = cost of ownership in new situation;

1. $-1 \leq q < -0{,}6$
2. $-0{,}6 \leq q < -0{,}2$
3. $-0{,}2 \leq q < 0{,}2$
4. $0{,}2 \leq q < 0{,}6$
5. $0{,}6 \leq q \leq 1$

(5) Contribution to *sustainability:* we define 4 realms where an artefact can contribute to sustainability: production, use, disassembly, and the context of the product. The scale is:

1. contribution to none of the above
2. contribution to one of the above
3. contribution to two of the above
4. contribution to three of the above
5. contribution to all four of the above

(6) The *impact on well being:*

1. no impact
2. material comfort
3. social comfort
4. illness or health
5. life or death

3. Possibility of realisation

(1) The *technical possibility of realisation*:

1. unknown whether the artefact can be realised
2. realization has been made plausible with informal arguments
3. realization has been shown to be feasible with model studies
4. a prototype has been built
5. a 0-series has been manufactured and realized in practice

(2) The *economical possibility of realisation*:

1. accurate estimates are lacking
2. there are accurate estimates for the total cost of ownership
3. moreover, there are accurate estimates of the profit, and the profit extends the cost of ownership
4. moreover, there is a business plan and a plan for financing (that may not have been approved)
5. moreover, there is a business case that is approved by the commissioner

4. Inventiveness

(1) The *time indicator*: the present artefact could have been realized …

1. <1 year (it was not possible before that time)
2. app. 1 year
3. app. 10 years
4. app. 50 years
5. it has always been possible

(2) The *surprise indicator*. This design surprises …

1. nobody
2. interested laymen
3. peers (fellow trainees)
4. experienced designers in practice
5. supervisors (from academia or industrial laboratories)

5. Complexity

The method to determine the three complexity indicators, to be presented here, is mainly intended for artefacts that can be naturally subdivided in components. Unfortunately, application of this method can be somewhat elaborate. Interfaces are the essential ingredient for determining the complexity indicators. The interfaces are weighted with their number of chosen (i.e., designed) features. In other words: the number of choice variables for the interfaces. To quantify this, we count the number of motivated choices, that should be evident from the documentation. Every documented chosen feature contributes 1 to the weight of the interface considered.

An interface can have weight 0 (in case no single specific feature of the interface is mentioned), and multiple interfaces may exist between any two components. There is some freedom left: an interface can be split, but then for each of the parts, the number of features will be less than the number of features in the undivided interface.

This brings about two perspectives to a system: the network perspective (Fig. 1) which only contains the basic components and their interfaces, and the hierarchical perspective (Fig. 2) which also contains the compound components.

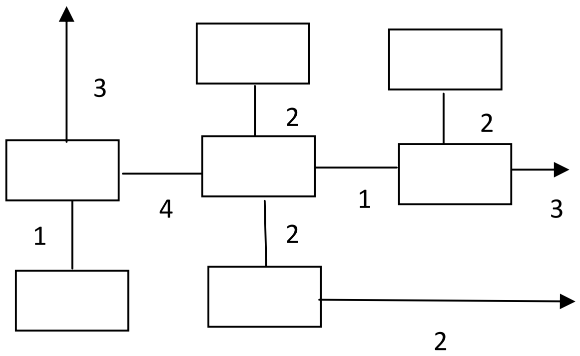


Fig 1 Network perspective

The structural complexity is defined as the double sum over all basic components, and per basic component over all its interfaces, of the number of documented chosen features. Formally: for a basic component *i* , I_i is the total weight of *i*'s interfaces with other basic components, and E_i is the total weight of the interfaces with the artefact's environment for that basic component. The structural complexity S is defined by:

$$S=\sum_{i:1..N} (I_i + E_i)$$

Here, *N* is the total number of basic components. In Fig.1, we find *S*=32.

(Notice that internal interfaces are counted twice, external interfaces are counted once. Also, in Fig. 1, there happens to be at most a single interface between any two components. This doesn't have to be the case in general.)

The functional complexity F is defined, counting the exterior interfaces only:

$$F=\sum_{i:1..N} E_i$$

In Fig.1, we find F= 8.

The *reduction of complexity* is determined by the hierarchy in the design, that is: the way how compound components have been chosen. In Fig. 2 we depict the same artefact as in Fig. 1, where compound components A, B, C and D have been added. A, B, C are at the same level; D is one level higher. Interfaces with the environment have been extended from the basic components to the highest component.

The hierarchical structure aims to improve the overview and the clarity of the system. Hierarchies can be depicted as tree structures, as depicted in the right part of Fig. 2.

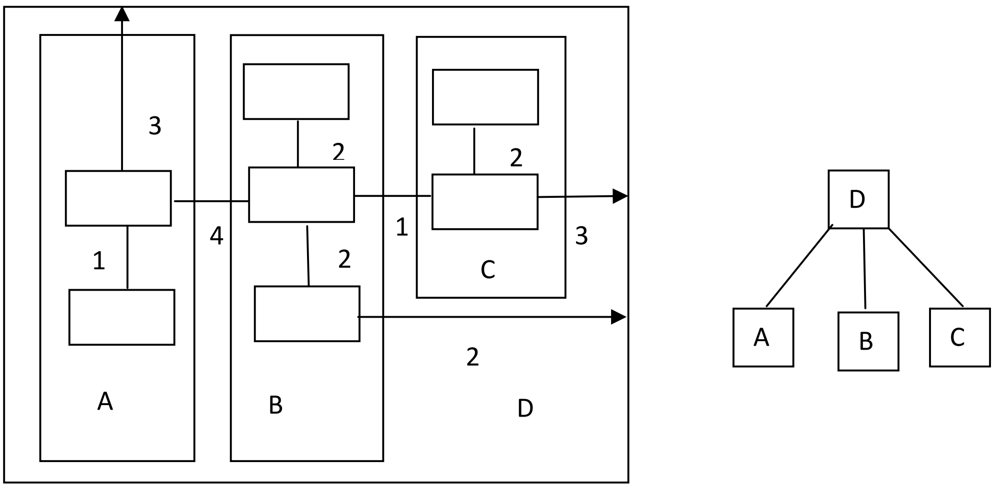


Fig. 2 Hierarchical perspective

The reduction of complexity for a compound component *i* is defined as $I_i / (I_i+E_i)$ where I_i is the sum of all weighted interfaces *within* the compound component, and E_i equals the sum of weighted interfaces projecting from the compound component outward – not necessarily to the exterior of the artefact. Only interfaces with components at the highest level within that component are counted: interfaces within subcomponents are left out.

The rationale is, that a compound component is 'better' if it hides a larger part of the interfaces. In software engineering, *cohesion* and *coupling* are terms often used to express this intuition. So I is a measure for cohesion, and E is a measure for coupling.

For an artefact, its reduction of complexity R is defined as:

$$R=(1/M).\sum_{i:1..M} I_i/(I_i+E_i)$$

Where M counts the total number of compound components. Notice that E of the topmost compound component in the hierarchy equals the functional complexity.

In particular, we find that a compound component without any internal interfaces has reduction of complexity equal to 0 – as it should be. In the example in Fig. 2, the reduction of complexity for components A, B, C and D is 1/8, 4/9, 1/3, and 5/13. For R we find therefore a value of 0,32.

The scales for these indicators are defined as follows:

(1) The structural complexity:

1. $0 \leq S < 10$
2. $10 \leq S < 50$
3. $50 \leq S < 100$
4. $100 \leq S < 250$
5. $S \geq 250$

(2) The functional complexity:

1. $0 \leq F < 5$
2. $5 \leq F < 10$
3. $10 \leq F < 20$
4. $20 \leq F < 100$
5. $F \geq 100$

To illustrate the latter values, we give some estimated scores: a new polymer: 1, a glucose meter for diabetics: 2, a cash dispenser: 3, a TOMTOM: 4, a CT scanner: 5.

(3) The reduction of complexity*:*

1. $0 \leq R < 0{,}2$
2. $0{,}2 \leq R < 0{,}4$
3. $0{,}4 \leq R < 0{,}6$
4. $0{,}6 \leq R < 0{,}8$
5. $0{,}8 \leq R \leq 1$

6. Elegance

(1) *Number of identified features of beauty* (K):

1. $K = 0$
2. $1 \leq K < 5$
3. $5 \leq K < 10$
4. $10 \leq K < 20$
5. $K \geq 20$

(2) *Quality* of the argumentation:

1. very poor
2. poor
3. marginal
4. good
5. very good

7. Genericity

Genericity refers to the degree of re-usability of an artefact, and whether a best practice has been developed that can be applied in other situations. A best practice can be either a method or a technology. The first indicator is called *artefact re-usability*. The second one is called *best practice re-usability.*

For artefact re-usability we propose the scale:

1. not re-usable
2. re-usable in other products belonging to the same product family
3. re-usable in other products fulfilling the same internal function
4. re-usable in other products used for the same purpose
5. re-usable in other products, including products that serve another purpose

For best practice re-usability we propose the scale:

1. not re-usable
2. re-usable in the same context of operation, and the same regime
3. re-usable in the same context of operation, and in a different or a broader regime
4. re-usable in a different context of operation, but within the same discipline
5. re-usable in a different or broader discipline

The term ‘context of operation’ is clarified with some examples: a software environment, a mechanical device, a type of chemical reactions are *contexts of operation*. The term ‘regime’ refers to the range of values for relevant quantities in the environment, such as pressure, temperature, time scale or length scale. ‘North pole’ or ‘equator’ are two examples of temperature regimes. The term ‘discipline’ refers to academic fields such as physics, chemistry and biology.

8. Methodical approach

This amounts to two indicators for the use of methods and techniques: one for modelling and one for analysing. We use the same scale for both:

1. no method was used
2. a method was used that is not state-of-the-art, and it was not used correctly
3. a method was used that is not state-of-the-art, but it was used correctly
4. a method was used that is state of the art, but it was not used everywhere, or it was not used fully correctly
5. a state-of-the art method was used correctly, everywhere when applicable

9. Documentation

This regards both the documentation containing of the various deliverables, and the presentation in the form of, e.g., a talk, a movie or a demonstrator. We use four indicators: the systematic organisation, the completeness, the correctness, and the accessibility. All four are measured against the same scale::

1. very poor
2. poor
3. marginal
4. good
5. very good.

References:

[1] Measuring Excellence in Engineering Research, The Royal Academy of Engineering, www.raeng.org.uk (January 2000) ;

[2] Regulations of the Dutch Certification Committee for Technological Design Programs (CCTO), effective as of February 1st 1993;

[3] Handbook Stan Ackermans Institute, August 1999;

[4] M.J.W. Schouten Kwaliteitsbeoordeling ontwerp- en constructiedisciplines, samenvatting, mei 2009;

[5] H. Simon; The science of the artificial; MIT press 1969, 3e edition 1996;

[6] A. Hevner, S. March, J. Park en S. Ram; Design Science in Information Systems Research; MIS Quarterly vol 28 nr 1;

[7] P.A. Kroes (ass.ed.); Philosophy of Engineering Design, part III of Handbook Philosophy of Technology and Engineering Sciences, edited by A.W.M. Meijers, Elsevier 2009, pp 405-631;

[8] D. Schön The reflective practitioner: how professionals think in action.Temple Smith (London) 1983.

3TU.School for Technological Design, Stan Ackermans Institute offers eleven two-year postgraduate technological designer programmes. This institute is a joint initiative of the three technological universities of the Netherlands: Delft University of Technology, Eindhoven University of Technology and University of Twente. For more information please visit: www.3tu.nl/sai.