Journal of Sound and Vibration 448 (2019) 247–258

Contents lists available at ScienceDirect

Journal of Sound and Vibration

journal homepage: www.elsevier.com/locate/jsvi

Special Issue: Acoustic Fatigue in Aero structures

On the validation of variable fidelity multi-physics simulations☆

Eann A. Patterson [a,*], Maurice P. Whelan [b]

[a] School of Engineering, University of Liverpool, Liverpool, UK
[b] European Commission, Joint Research Centre (JRC), Ispra, Italy

ARTICLE INFO

Article history:
Received 22 February 2018
Received in revised form 17 December 2018
Accepted 27 January 2019
Available online 1 February 2019
Handling Editor: Paolo Tiso

Keywords:
Complex system
Model validation
Model credibility
Epistemic values
Hermeneutic

ABSTRACT

The difficulties encountered in applying current normative approaches for validation to computational models of complex multi-physics engineering systems are identified and are associated with the untestable, and sometimes unprincipled, nature of these models. The behaviour of a structural panel on the surface of a hypersonic flight vehicle when subject to complex interactions between aerothermal, aeroelastic and material responses is employed as a key exemplar. A wide range of positions in the philosophy of science, that are applicable to validation, are discussed within the context of a schematic matrix, which allows models to be categorised according to whether they are testable and principled. In the absence of test data against which to assess the accuracy of predictions, it is proposed that a model's credibility should be established based on its epistemic values, theoretical ancestry and the credentials of the modelling techniques. This shift from an objectivist to a relativist approach requires the assignment of experts who acknowledge their biases while engaging intellectually and ethically with the model, the community of knowledge and stakeholders, in a hermeneutical approach.

1. Introduction

Hypersonic flight vehicles, fusion energy systems and spacecraft for inter-planetary exploration are engineering systems that are complex, involve multiple domains of physics and for which there is limited or no observational data to support the development of computational models that can be used to simulate system performance for the purposes of design, planning operations and maintenance, and predicting service life. The airframe of a potential hypersonic flight vehicle is employed as an exemplar to explore the issues associated with the validation of computational models of complex, multi-physics engineering systems.

A number of organisations are developing vehicles capable of hypersonic flight for a variety of applications including high-speed cruise for global travel and access to space for delivery and maintenance of satellites. Hypersonic flight corresponds to speeds around five times the speed of sound at which the vehicle is subject to an extreme, environment with combined fluid, thermal and structural excitations. Miller et al. [1] have identified the challenges to be resolved in predicting the structural response, lift and reliability prognosis through the complete life-cycle of the airframe of such a vehicle. These include the

☆ Second revision of manuscript for Special Issue on 1st International Aerospace Symposium on Acoustic Fatigue|September 11–12, 2017.

* Corresponding author.
E-mail addresses: eann.patterson@liverpool.ac.uk (E.A. Patterson), maurice.whelan@ec.europa.eu (M.P. Whelan).

https://doi.org/10.1016/j.jsv.2019.01.051

pressure excitations from turbulence in the boundary layer from transient deformation of the skin of the vehicle and from engine noise; high temperatures as a result of flow compression and viscous dissipation; the presence of geometric and material nonlinearities; the temperature dependence of material properties; and uncertainty in loads, material properties and boundary conditions. These factors have their origins in different domains of physics, which must be integrated in order to generate a successful design that has an appropriate compromise between structural integrity and weight for optimum range and load capacity. At a fundamental level, this involves the integration of computational fluid dynamics (CFD) and structural mechanics, in which the end-point is the structural design of the vehicle, but an essential earlier step is the CFD predictions of airflow, including turbulence causing broadband random excitation, that can be used in an iterative design process with load calculations. However, the thermal loading is also important and can cause out-of-plane displacement of the vehicle's skin panels that induces structural stresses and changes the flow regime. The interactions between these excitations are illustrated schematically in Fig. 1. These interactions result in a complex system whose behaviour cannot be reduced reliably into its constituent parts due to the dependency of one part on another that causes behaviour to emerge when the whole system is considered, which is not apparent when individual parts are considered. This renders experiments difficult to perform because it implies the need to conduct tests with prototype flight vehicles that are representative of the whole system. At the same time, simulations are also intrinsically difficult because of the coupled fluid-thermal-structural interaction. Elegant structural design, which meets performance requirements while making efficient use of material and energy, requires a high level of confidence in the computational simulations used in the design process. This confidence is usually gained through validation processes that establish the extent to which predictions from a model represent the real-world taking account of the intended use of the model [2,3]. This has motivated significant efforts to develop high-fidelity multi-physics models, often using reduced order approaches, for example by Tiso et al. [4]. However, this approach creates a conundrum, because the descriptor 'high-fidelity' implies that the predictions are an accurate representation of the real-world, which in turn suggests that a detailed quantitative comparison of the predictions with comprehensive measurements has been performed, for example using the process described by Sebastian et al. [5]; but this is unlikely, because the required comprehensive measurements are not readily available, as discussed above. Instead, it is more pragmatic to aspire to variable fidelity in a multi-physics simulation for which the level of fidelity is known, and its potential implications recognised and understood. This reduction in demand for fidelity allows a relaxation of the constraints of the quantitative, statistical comparison of measurements and predictions stipulated by validation guidelines. This paper discusses a framework within which this relaxation can be specified based on existing positions in the philosophy science. The intention is to provide a philosophical overview and direction for establishing conditions in which decision-makers are willing to utilise the predictions of models of complex engineering systems.

2. Approaches to model validation

For the past twenty years, validation processes in engineering modelling have been dominated by the definition provided in the AIAA and ASME Verification and Validation (V&V) guides [2,3], i.e. 'the process of determining the degree to which a model is an accurate representation of the real world from the perspective of the intended uses of the model'. The ancestry of this definition can be traced to early work by Fishman and Kiviat [6] and by Van Horn [7], who in the context of economic science, observed that a computational model is usually developed with objectives that reflect an intended use and so the predictions should be evaluated against these objectives. Sargent [8] added specificity by including the term 'intended uses of

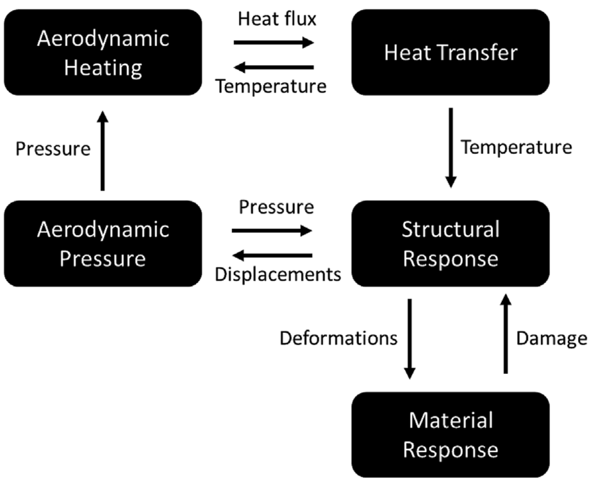

Fig. 1. Schematic representation of interactions between excitations of an airframe during hypersonic flight (based on Miller et al. [1]).

the model'. While Balci and Sargent [9] framed the validation process in terms of the rejection of a null hypothesis: 'the model is valid for the acceptable range of accuracy under the experimental frame'. The concept of validation emerged as the comparison of predictions from the model with measurements from the real-world during the 1980s [10,11] leading to its incorporation in the AIAA and ASME guides [2,3]. Philosophically, this approach has been seen as Popperian with models being tested by attempting to falsify them within the context of the intended purpose, and when falsification is not achieved then models are regarded as valid or acceptable for the intended purpose and for the moment [12,13] but it does not imply they are completely correct or true representations. However, this focus on a single philosophical approach to the validation of computational models can become a constraint that limits the exploration of new or enlarged design domains due to a lack of measurement data to support the modelling or can be become a reason for limiting the validation process because of the challenges and costs involved. Kleindorfer et al. [14] have identified ten different positions in the philosophy of science that are applicable to the validation process; however, these have received little attention from the engineering community. Recently, some of these, including rationalism, empiricism and instrumentalism, have been proposed by Patterson and Whelan [15] as means of handling the difficulties encountered in the validation of computational biology models when there is a lack of measurement data. They used a 3×3 matrix diagram to schematically identify the appropriate validation approach for a model based on its epistemic foundation and the availability of data. This diagram was derived from an earlier 2×2 matrix [16] that was based on the epistemological division, identified by Tegmark [17], between physics in which a theory is testable and meta-physics where theories are not testable. This divide corresponds to the boundary between science and non-science identified by Popper [18]. There is a parallel epistemological division in engineering between computational models that predict the behaviour of systems that do not yet exist and those that retrodict or simulate systems that already exist and for which measurement data is available, or can be obtained. For the latter, the approaches described in the validation guides are usually appropriate because measurement data from the real-world is available to establish the accuracy of the simulation; however, this is not the case for the former because no data is available to demonstrate predictive accuracy. Hence, this suggests that at least two different approaches to validation should be available to modellers of engineering systems. The schematic in Fig. 2 blends the previous matrix diagrams [15,16] with the summary table provided by Kleindorfer et al. [14] to illustrate the potential relevance of a number of validation approaches described by their corresponding position in the

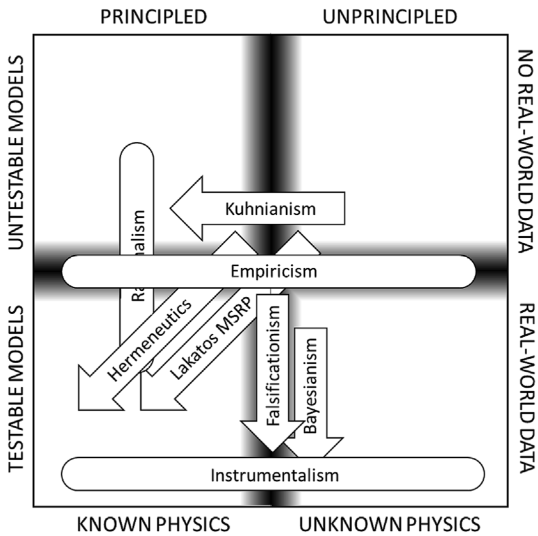

Fig. 2. Schematic diagram illustrating the epistemological divisions between principled models based on known physics and unprincipled models based on unknown physics, as well as between testable and untestable models with fuzzy boundaries forming the 2×2 matrix (based on Patterson [16]); also shown are various positions in the philosophy of science from Kleindorfer et al. [14] illustrating their potential applicability as approaches to model validation, namely:

Bayesian: probabilistic evaluation of predictive power updated with new observations;
Empiricism: demonstration of assumptions and predictive capability by comparison to observations;
Falsification: model success/survival in increasingly severe tests of predictive capability;
Instrumentalism: demonstration of predictive capability ignoring assumptions and theory;
Kuhnianism: expert determination of consistency with current paradigm;
Lakatos MSRP: model refinement using new understanding and data from research programs;
Rationalism: exposition of logical deduction of model from undisputed foundation.

philosophy of science. The horizontal axis relates to the knowledge-base underpinning the model: models on the left are built on principles that are well-understood and accepted, Barlas and Carpenter [19] use the term 'causal' or 'theory-like' to describe this type of model; whereas those on the right are unprincipled, unsupported by widely accepted scientific knowledge and often built on statistical correlations, Barlas and Carpenter described these as 'noncausal'.

These dichotomous views of models as testable/untestable and principled/unprincipled can be problematic for a complex system where some level of real-world data might be available and similarly some parts of the model might be based on principles while other parts might not. Hence, the boundaries in Fig. 2 are often fuzzy and ill-defined as illustrated in the diagram, and the precise position of a model in the domain of the diagram can be difficult to identify. However, the diagram is intended to provide a framework for reviewing the status of a computational model with regard to its likely credibility as well as the appropriate routes for validation; and in this context the fuzziness should not be an obstruction. Model credibility is the willingness of stakeholders to use the predictions of models to support decisions. In engineering, where models are used to inform decisions that have consequences for individuals and organisations, the preference is to operate in the bottom left corner of the diagram in Fig. 2, i.e. using testable models built on known physics, and in these circumstances most of the validation approaches identified by Kleindorfer et al. [14] and shown in Fig. 2 can be utilised; however, the engineering guidelines describe a methodology that is essentially Popperian because it is based on falsification [12], i.e. the model is subject to a series of severe tests, in which its predictions are evaluated against measurements made in the real-world, and if found inadequate the model is rejected or returned to the development phase for revision. 'Falsification' is shown in the bottom half of the diagram at the boundary between principled and unprincipled model because the approach is not dependent on the model being based on known physics and any testable model can be subject to a series of severe tests in an attempt to falsify it. In the process of attempting to falsify a model, more real-world data will be generated causing the model to move downwards in the diagram, as shown by the downward pointing arrow. If this testing is performed within a probabilistic framework then Bayesian statistics can be applied, so that increasing empirical data allows updating of the probability about the null hypothesis, as represented by the corresponding parallel downward arrow. Two other approaches to validation are applicable to testable models: empiricism and instrumentalism. Empiricism uses inductive generalisation, i.e. conclusions based on observational data, and only permits assumptions that can be independently and empirically confirmed to be included in a model. In Fig. 2, empiricism is shown along the fuzzy boundary between testable and untestable models because of the potential to use it to test the assumptions upon which an untestable model is constructed; and, hence provide a degree of validation that would not be possible by falsification or Bayesian approaches applied to the model itself. Empiricism can be deployed whether or not the underlying physics is understood. Patterson and Whelan [15] proposed combining empiricism with rationalism following the ideas proposed by Kant. Rationalism uses a series of logical deductions to build a theory from a set of undisputed truths, which implies that the principles underpinning a model are known and accepted but the model itself need not necessarily be testable, so that it is shown in Fig. 2 occupying the left side of the diagram. In practice, rationalism is the approach taken to building many models used in simulation [20] often by employing reductionist principles to resolve a system into its constituent parts for which the principles are known [14]. Hence, rational empiricism involves constructing the model using a series of logical deductions based on assumptions that have been tested by experiment; and the corresponding validation approach involves identifying the observational or experimental evidence supporting the relevant deductions and assumptions. Rationalism and empiricism are the first two steps in the historical three-step approach to validation proposed by Naylor & Finger [21]. The third step is instrumentalism or, in the context of economic modelling, positive economics. Instrumentalism demotes theories to the status of convenient instruments used to arrange observations, akin to a filing system, and focusses on the predictive success of the model. Instrumentalism is shown along the bottom of the diagram in Fig. 2 because it requires real-world data to demonstrate predictive success and, in common with empiricism, is agnostic about the knowledge-base of the model.

The top right box in Fig. 2 is relatively empty of potential validation approaches because models belonging in this region are unprincipled and untestable leading to radical uncertainty. Any validation approach for such models should aim to move them out of this region. Two approaches might be considered useful in this context: Kuhnianism and the Lakatos MSRP (Methodology of Scientific Research Programs). Kuhn regarded validation as 'a complex process with social, psychological and historical dimensions' [14] in which an expert is called upon to determine whether the model fits the current paradigm using criteria such as consistency and fruitfulness. In this context, an expert is someone adept at recognising the contemporary paradigm and the process of validation involves identifying how the model confirms and extends the reigning paradigm, which might involve extending our knowledge of the underlying physics; hence the Kuhnian approach to validation is shown as a horizontal arrow in Fig. 2, representing a shift from unknown to known physics as the model is accepted into the existing paradigm. Lakatos suggested that a research program is successful when it generates more empirical information, which would clearly represent a downwards movement in Fig. 2; however, here it is suggested that a successful research program should also add to knowledge which would provide a leftward component to the arrow. In Fig. 2, the downward and leftward components of the Lakatos MSRP are equal but this need not be the case. Kleindorfer et al. summarise the Lakatos MSRP approach to validation as an 'increase (in) empirical and theoretical content without ad hoc adjustments' [14]. In other words, the Lakatos MSRP approach involves enhancing the rational and empirical basis for a model through a planned research programme rather than reactively as a consequence of falsifications. This might be particularly relevant to extending the scope of an existing model beyond the bounds of its original conception and validation. In practice, the process often occurs as part of an interaction between researchers, modelers and other stakeholders as part of a process of social epistemology [22,23], which involves the development of a shared knowledge and understanding that generates, what Welbourne called, a community of knowledge [24]. The interactions tend to lead to an enlargement and consolidation of the community and likely acceptance of the model by the members of the community. The concept that the truth, or an appreciation of it, evolves from an interaction between an individual and the situation of interest is known as hermeneutics [20] and is identified as an approach to validation by Kleindorfer et al. [14] but is not included Fig. 2 because it could be part of any of the processes highlighted. Engineering models tend to be used to support decision-making and the decision-maker is rarely the model developer. Hence, the developer needs to provide information that establishes the credibility of the model predictions in the perception of the decision-maker and this is often achieved through a process of interaction in which both the modeller and decision-maker bring their expertise and experience to bear on the issue. During the process, the value of the model predictions becomes apparent in the minds of the modeller and decision-maker, which could be described as a hermeneutic approach regardless of the validation approach or approaches employed.

Perhaps at this juncture, it is useful to consider the possible location in the schematic diagram of a small selection of recently-published models in the field of sound and vibration prior to a more detailed examination of the position of a multi-physics model of the structure of a hypersonic vehicle that has motivated this study. In Fig. 3, the positions in the philosophy of science have been removed from the schematic diagram and been replaced by circles indicating the likely positions of five recently-published models [25–29] that were chosen for their contemporaneousness and to provide a distribution across the diagram. Germanpré et al. [25] studied the contributions of longitudinal track unevenness and track stiffness variation to railway-induced vibration based on measurements from a Swedish measurement vehicle, IMV100. They compared the predictions from their finite element-boundary element model with measurements of vibration that occurred during the passage of an X31 Öresund passenger train. Their model was based on known physics and they were able to test their model using data from the real-world obtained directly from the system of interest, i.e. a passenger train on a stretch of track at Furet, Sweden; hence, this model lies in the bottom left corner of the schematic diagram in Fig. 3. Bajrić and Høgsberg [26] also obtained measurements related to structural vibration; however, they used a model-scale five-story building rather than a real-world system which moves their model upwards in the diagram in Fig. 3 because measurement data from the actual

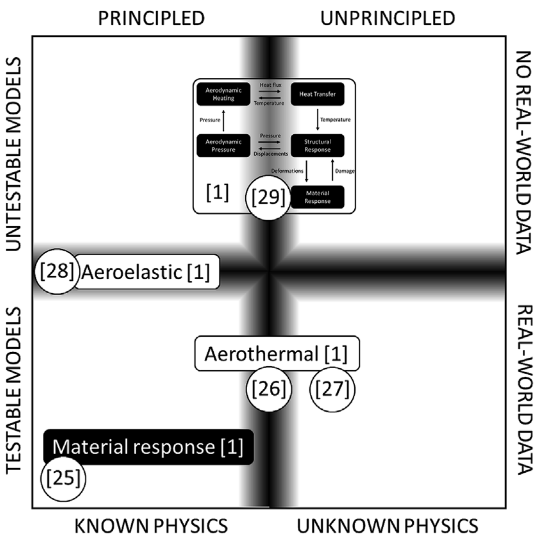

Fig. 3. Schematic diagram combining elements of Fig. 1 with Fig. 2 to show the state of underpinning knowledge and real-world data for the constituent parts of the aerothermoelastic system model and the system model for airframe panels in a hypersonic flight vehicle together with the possible locations of recently published models, shown in circles:

[25] Non-periodic track model to railway-induced vibration compared to measurements from Swedish measurement vehicle by Germanpré et al. [25];
[26] Novel expression of non-classically damped structure compared to numerical simulations and experimental measurements for five-storey building by Bajrić and Høgsberg [26];
[27] Analytical model of rocket combustion chamber compared to data from firing tests of laboratory-scale rocket by Carmicino and Di Martino [27];
[28] LES model of vortex structures in flow field of pantographs compared to theory and wind-tunnel tests by Xiao-Ming et al. [28];
[29] Dynamic substructuring method used to establish equations of motion for large space structures and compared to other numerical models by Hu et al. [29].

object of interest, i.e. the real five-story building is not available. In addition, the complex nature of the engineering structure means that it is difficult to predict the modal shapes and damping matrix for the dynamic structure because a full understanding of the physics is not yet available, in particular of the interactions between the parts of the structure; and, this moves the model to the right side of the diagram in Fig. 3. Similarly, Carmicino & Di Martino [27] predicted the longitudinal acoustic modes of combustion chambers in rockets and compared their results with laboratory-scale firing tests, which places their model in the middle of the testable model region. They used an analytical approach based on a one-dimensional description of waves which worked well until vortex shedding started to occur. Hence, the model can be placed at the boundary between known and unknown physics in Fig. 3, because the physics appears to be well-founded on known physics for a limited range of operation beyond which the physics is less well-known and not incorporated into the model. Xiao-Ming et al. [28] constructed an LES model of the vortex structures in the flow field associated with pantographs, which was based on well-understood physics of fluid flow, placing the model close to the left edge of the schematic diagram in Fig. 3. However, obtaining real-world data for this scenario is almost impossible and only wind tunnel tests on a scale-model were viable, which yielded limited data, and none directly related to the predicted radiated noise from the pantograph. Hence, the model could be considered to be almost untestable and placed at the boundary between testable and untestable in Fig. 3. Finally, Hu et [29] modelled a decentralized simple adaptive control of vibration in large space structures on which they were unable to conduct any physical tests and they identify a number of limitations in their model that require further research. Hence, taking these two factors together places their model in the untestable model category and at the boundary between known and unknown physics. Since, there are some measurements available from physical tests for the first four of these models [25–28], then it would be possible to apply the validation procedures outlined in the ASME [2] and AIAA [3] guides, i.e. to attempt to falsify their predictions in increasingly arduous comparisons with physical tests. However, in the reported studies, either a form of empiricism was practised in which data from experiments were used to confirm the underlying assumptions or theory, e.g. in the model of the pantograph [28] for which aerodynamic resistance and pressure levels were compared rather than far-field radiated noise; or, a form of instrumentalism was practised, to directly demonstrate the predictive capability of the model, for example in the case described by Germanpré et al. [25] for the railway vibrations. However, the situation for the model of the large space structures, described by Hu et al. [29], is different because no data from physical tests are available and instead, they compare their predictions to those from other models in order to determine their consistency with the existing paradigm, which can be construed as a form of Kuhnianism. The situation is similar for the multi-physics model of an airframe in a hypersonic flight vehicle that provided part of the motivation for this study and is described in more detail in the next section.

3. Exemplar: multi-physics models of hypersonic vehicles

It is appropriate to consider the location of multi-physics models in the schematic in Fig. 3 so that suitable approaches to establishing their reliability and utility can be discussed. The behaviour of a structural panel on the surface of a hypersonic flight vehicle will be employed as an exemplar. This behaviour, which is commonly characterised as thermo-acoustic behaviour, involves complex interactions between aerothermal, aeroelastic and material responses as shown in Fig. 1. A typical system model used to predict the behaviour of such a structural system would be expected to exhibit emergent behaviour, i.e. behaviour that is not evident from that of its constituent parts. The usual approach to building a computational model of the complex system is to attempt to connect models of the aeroelastic, aerothermal and material response. For example, Miller et al. [1] in their aeroelastic model used von Karman plate theory to describe a panel's equation of motion and included (a) the thermal loading due to arbitrary, non-uniform, in-plane and through thickness temperature distributions; (b) the chord-wise variation of the elastic modulus and the thermal expansion coefficient; (c) randomly generated zero mean value acoustic disturbance forces; (d) rotary inertia and (e) mass-proportional structural damping. They simplified the problem by assuming that in-plane forces were constant, in-plane displacements were zero at immovable supports, mass-proportional structural damping was constant, and the damping coefficient was taken as 2% of the critical damping coefficient of the first mode of the unloaded panel at the initial temperature. They also simplified the acoustic disturbance force to be spatially uniform and temporally random about a zero mean. This aeroelastic model of a structural panel is well-founded on accepted knowledge of the dynamics of panels; however, the complicated boundary and loading conditions make experimental tests challenging, so that it is reasonable to place this isolated part of the system model on the left of the diagram in Fig. 3, i.e. in the known physics regime, but in the fuzzy zone between testable and untestable models.

In their aerothermal model, Miller et al. [1] accounted for the thermal equilibrium and heat transfer between the boundary layer and the thermal structure. Two-dimensional heat transfer equations were used to describe the chord-wise and through-thickness conduction paths. Eckert's reference enthalpy method was employed to model convective heat transfer, which is an entirely empirical approach, such that the aerothermal model can be placed in the bottom half of Fig. 3 because it is testable, but in the fuzzy zone between principled and unprincipled models because the exact process of hypersonic convective heat transfer is not fully understood.

Miller et al. [1] did not model the material response; however, the damage mechanics for high-temperature alloys has been reliably modelled by others [30] including, recently, a probabilistic approach [31] that takes account of the statistical nature of the material's microstructure, the complex three-dimensional stress/strain field and the grain boundary energies. This modelling is supported by extensive empirical data at the material or coupon scale, for example Li et al. [32] and Bache et al. [33]. Hence, the material response model can be placed in the bottom left corner of Fig. 3 because there is a strong knowledge foundation for the models and ample real-world data to validate them. However, this real-world data does not include tests on large-scale structures; hence, when structures rather than coupons are the focus of the modelling then the material response model should be moved upwards in Fig. 3 to represent the decrease in availability of real-world data. When the material response model is integrated into the system model, then the viability of testing is very substantially reduced, effectively to zero and the system model should be moved into the untestable region. At the same time, it needs to be acknowledged that appropriate methodologies for connecting together the constituent parts of the system model are largely unknown, which places the system model in the fuzzy region between principled and unprincipled models, as shown in Fig. 3.

The exact position of the system model in Fig. 3 is unimportant. Most models of complex multi-physics systems involve several layers of discretization that lead to a series of discrete sub-models; and the connections between these sub-models are often 'kluge' that cobble pieces of code together and add substantially to the uncertainty in the predictions. These kluges are specially designed algorithms that connect otherwise incompatible sub-models [34], which are not based on scientific knowledge but on programmers' experience of what works, in other words they are unprincipled and hence pull a system to the right, at least to the fuzzy zone between principled and unprincipled models. Similarly, a complex system is unlikely to be deterministic and will generate a set of solutions that can be described by a phase portrait or attractor; so that, real-world measurements will only represent a tiny fraction of the possible locations in the phase portrait and hence, a model of a complex system is likely, at best, to lie on the fuzzy zone between testable and untestable models in Fig. 3.

4. Credibility of complex system models

In section 2 a plethora of validation approaches for models and simulations drawn from a wide range of disciplines, including both physical and social sciences, were outlined. The approaches are not as distinct as portrayed in Fig. 2 and section 2. There are many variations on, and overlaps between, approaches that are not appropriate to discuss in this context; however, it is clear that there are significant choices available for the validation of engineering models beyond the empirical, and largely Popperian-based approaches that have become dominant in engineering. At its most basic level, 'validation is the process of determining whether a model reflects reality' [35]; the additional specificity provided by a statement about 'its intended purpose/use' limits the domain over which falsification attempts need to be made and constrains the costs of validation. However, the requirement to attempt falsification requires the acquisition of empirical data, which constrains the validation process to testable models and limits the acceptable modelling domain to well-established engineering. A broader interpretation of determining whether a model reflects reality is required to allow exploration of the domain represented by the top half of the diagrams in Figs. 2 and 3 where, for an entire complex system, the system model is untestable and the physics might be unknown. Previously, models in the top right box were seen as solely heuristic, i.e. constructed purely for the purposes of discovery and learning; however, the increasing cost and complexity of engineering systems, such as large space structures, hypersonic flight vehicles and fusion reactors, may place their models in this box because reduced-scale prototypes for generating validation data are not viable and emergent behaviour is likely so that the physics may be unknown, as discussed in the previous section. Hence, it is appropriate to consider alternative validation approaches as proposed by Kleindorfer et al. [14], who sought to free practitioners to pursue a varied set of approaches and to convert the validation problem into an ethical one in which the practitioner must responsibly and professionally argue for the worth of their model. This converts the task from one of solely validation to one focussed on establishing model credibility in the perception of other stakeholders, including decision-makers. However, establishing model credibility, i.e. a willingness of others to use predictions to inform decisions [36], is not straightforward in the absence of empirical data and possibly of well-defined knowledge of the physics. These circumstances are encountered elsewhere, for instance, in climate modelling when reliable detailed predictions of long-term events are sought with only sparse observational data available in the short-term. In this context, Biddle and Winsberg [37] suggested that models with epistemic values are more likely to be appropriate. This is an extension of the concept proposed by McMullin [38] that there is a set of characteristics or values, which are implicit in modern science and which enhance 'the truth-like character of science'. Epistemic properties include simplicity, consistency and explanatory power as well as predictive accuracy; although the latter is not relevant for untestable models. Credibility is more likely to be achieved when a model and its underpinning concepts are understandable by stakeholders as well as modellers and so there is a direct link between simplicity and credibility. However, there is a more fundamental argument in favour of simplicity because it is inefficient to replace a simplifying assumption, termed as a tractability assumption by Kuorikoski et al. [39], by a more complicated one unless it results in a more reliable or realistic prediction [40]; although the same does not hold for substantial assumptions that define the causal core of the model. The process of investigating the sensitivity of the predictions from a model to its tractability and substantial assumptions has been described as derivational robustness and sufficient robustness analysis respectively by Raerinne [41]. The credibility of a computational model is dependent on its insensitivity to tractability assumptions and its sensitivity to substantial or causal assumptions. The latter are also related to its internal consistency or coherence which requires that there should be no logical contradictions in the model. Whereas external consistency requires that a model predicts similar effects and behaviour as predicted by alternative models and as expected from a priori knowledge. Explanatory power effectively combines two values of scientific theories identified by McMullin [38], namely unifying power and fertility. He defined unifying power as 'the ability to bring together hitherto disparate areas of inquiry' and fertility as the ability to make predictions beyond the situation on which the model was based originally. The former is perhaps less important when modelling a complex engineering system; but clearly a model that can elucidate or describe observed or known phenomena is likely to be more credible and Bailey has highlighted that the plausibility of models is enhanced when they are able represent situations beyond those used to guide their construction [42], i.e. they possess fertility.

It can be unviable to demonstrate epistemic values in the absence of real-world observations. In the latter circumstances, the theoretical ancestry of a model must make a large contribution to establishing its credibility as well as the established credentials of the model-building techniques, with the latter leading to 'self-vindicating' models [34]. In this context, it is useful to sub-divide the process of generating predictions from models into two stages: (i) laying out or selecting the theoretical concepts or principles if the model is to be principled; and (ii) applying these concepts to the real-world problem of interest. From the perspective of principled models, Hacking [43] called these two processes 'speculation' and 'calculation' respectively, though Kuhn called the latter 'theory articulation', which is perhaps more descriptive of the process involved [45]. Thus, the authority of the theoretical concepts, selected to construct the model, form the theoretical ancestry of the model; while the techniques used to articulate the theory in the model are critical to the reliability of the predictions. Winsberg suggests the construction of a model draws not just from the selected theoretical concepts but also from the modellers intuition and speculative acquaintance with the real-world problem [34]. There is an extensive list of activities, practices and assumptions that are incorporated into a model as part of its construction and their track record forms the modelling credentials. This list includes: assumptions about which parameters can be neglected, rules of thumb for surmounting computational difficulties, and choices about differencing schemes and exploitation of symmetry, for example [34]. The past performance of this modelling praxis provides some level of confidence in models built using the same praxis and when this past performance includes clearly recognised successes then their reputation as reliable praxis grows and enables the creation of self-vindicating models.

These concerns about epistemic values, theoretical ancestry and model-building credentials are somewhat vague and likely to yield only qualitative information in support of a model's credibility. This implies some subjective judgment of value will be required, which Rudner [44] and Kuhn [45] have argued is an essential part of science with choices between theories being influenced by a combination of objective and subjective factors, including idiosyncratic influences. This is difficult to implement in an engineering community where the prevailing philosophy is objective reductionism based on the work of Newton and Descartes. Hence, an attempt has been made in Table 1 to identify the factors that influence the credibility of a model that is untestable. In the same table, strategies for assembling evidence that a model possesses these factors are outlined, based on the foregoing discussion and the work of Naylor & Finger [21], McMullin [38], and Winsburg [34]. The table draws on earlier work in chemical risk assessment [46] where a combination of data-based and knowledge-based evaluation of model credibility can be undertaken. In the current context, data-based credibility factors will not be applicable for an untestable model but are left in place because of their potential relevance to component or sub-models. Additional factors for theoretical ancestry and model-building credentials have been included to support the evaluation of self-vindicating models. It is proposed that the credibility of a complex model, which is untestable, could be established through a process of interaction between the modeller and decision-maker using the list of credibility factors as an agenda to guide a series of discussions and, ultimately, as the basis for structuring evidence provided by the modeller to the decision-maker in support of establishing the model's credibility in the perception of the decision-maker.

In the context of the multi-physics model of a hypersonic flight vehicle shown schematically in Figs. 1 and 3, and described in the previous section, the models of material response and aerothermal behaviour can be validated following the procedures described in the engineering standards [2,3] because they are testable. However, as discussed in section 3, the aeroelastic model of panel behaviour lies in the fuzzy zone between testable and untestable models in Fig. 3 and thus consideration of the data-based credibility factors in Table 1 would be appropriate. The system model of the airframe panels is complex with emergent behaviour and untestable, so that it would appropriate to gather evidence of its possession of the credibility factors associated with epistemic and self-vindicating properties listed in Table 1.

5. Discussion

The normative approach to validation is biased towards a Boolean decision on a model being valid/invalid or acceptable/unacceptable based on comparing predictions and observations [47]. The ASME V&V guide [2] states explicitly that the truth of a prediction 'cannot be proven in sense of deduction logic' and 'should be tested for trustworthiness by the accumulation of evidence', while at the same time it recommends the reduction of a system into its component parts for a 'bottom-up approach to validation and verification'. These recommendations are useful for models that lie in the bottom half of the diagrams in Figs. 2 and 3; however, they impose substantial constraints on model development and deployment for complex multi-physics systems whose models are untestable at some level and this potentially inhibits advances in the design and production of these engineering systems.

In the previous section, it has been argued that establishing credibility, rather than undertaking a process of validation, should be pursued for untestable models of complex engineering systems, located in the top half of the diagrams in Figs. 2 and 3. The factors that influence the credibility of such computational models have been identified as the epistemic values of simplicity, internal and external consistency and explanatory power, the theoretical ancestry of the model, and the credentials of the model-building techniques. Credibility is the willingness of others to use model predictions to inform their

Table 1
Credibility factors and strategies for an untestable model [based on Patterson et al. [46].

Credibility Factor	Potential strategies to demonstrate possession of factor	
Assumption confirmation	≻ Identify assumptions underpinning model & their limitations ≻ Collate observational evidence to justify each assumption ≻ Identify [characterise] uncertainty associated with assumptions	Data-based factors, which might be relevant for sub-models; based on Naylor & Finger [21].
Qualitative concordance	≻ Assess extent to which predicted behavioural trends [effects] match observed behaviour ≻ Assess uncertainty for behavioural trends	
Quantitative concordance	≻ Quantify the extent to which predictions represent observational data for the intended use of the model ≻ Quantify uncertainty in the quantitative predictions	
Explanatory power	≻ Explain observed phenomena and behaviour [effects] using predictions ≻ Using model explain situations & effects other than those on which approach is based [from Bailey [42]] ≻ Identify limitations in explanatory power	Epistemic properties; based on McMullin [38]
Internal coherence	≻ Demonstrate model predicts already known result[a1] [e.g. training dataset] (calibration) ≻ Demonstrate perturbation of input parameters produces expected result[a2] (sensitivity) ≻ Demonstrate predicted behaviour [effects] disappears in appropriate circumstances[a3] (disappearance) ≻ Demonstrate predictions unchanged by elimination of all plausible sources error[a4] (artefact removal)	
External consistency	≻ Predict similar effect/behaviour with an alternative model[s][a5] (external comparison) ≻ Assess reproducibility of model in different conditions ≻ Quantify uncertainty in reproducibility and alternative prediction comparisons	
Simplicity	≻ Demonstrate appropriate degree of complexity by removal of each core assumption producing a significant change in prediction ≻ Build a strong narrative with appropriate level of detail that is both precise and concise ≻ Express uncertainty about level of complexity and influence of known unknowns and unknown unknowns	
Theoretical ancestry	≻ Identify theoretical concepts underpinning the model & their limitations ≻ Collate historical evidence for relevance and reliability of theoretical concepts ≻ Identify [characterise] uncertainty associated with theoretical concepts	Properties of self-vindicating models; based on Winsberg [34]
Model-building credentials	≻ Identify components of model praxis, e.g., neglected parameters, 'rules of thumb', choice of differencing schemes, exploitation of symmetry. ≻ Collate historical evidence for reliability of each component of modelling praxis and for their integrated use. ≻ Identify [characterise] uncertainty associated with theoretical concepts	

[a] Based on five strategies commonly used by experimentalists to ascertain the dependability of their data, according to Franklin [53].

decisions and hence belongs to the decision-maker rather than the modeller. This implies the need for modellers to convey information to decision-makers that demonstrates the extent to which a model possesses these factors, which Yang et al. [48] has proposed can be achieved as part of an on-going process employing a broad range of tools, including focus groups and questionnaires. This is likely to be necessary because the initial distribution of opinions on the credibility of a model will likely be a normal distribution, particularly in a large group of stakeholders. Hence, a series of interactions will be required to allow individuals to interact with the model in a hermeneutic process that permits a broad range of approaches to establishing credibility to be considered. This will often involve social interaction between modellers and stakeholders, such that the process of model building, development, validation and establishing credibility can be viewed as a socio-technical activity that could be enhanced by social media tools, which have been shown to increase productivity and reduce costs and lead times [49]. This will enhance connectivity with the community of knowledge and support epistemic dependence which is the inclination to depend on the understanding of others and this is greater when individuals have apparent access to experts who already have that understanding [50].

The proposed change in focus, from validation to establishing credibility, might be troublesome for some stakeholders because it involves crossing the divide between objectivism and relativism. Objectivism is based on the idea that there is a unique, ultimate foundation for things, including models – it is also referred to as foundationalism. Relativism is the doctrine that knowledge and truth are not absolute – this is an anathema to many engineers who are trained to seek the absolute answer to a problem using the reductionist principles of Descartes and Newton. A reductionist approach involves measuring and quantifying behaviour; whereas the contrary approach, sometimes known as a systems view, places the emphasis on the mapping of relationships and behavioural patterns because an exact knowledge of the relationships and behaviour of a complex system will usually be unattainable [51], which renders an objective approach to validation problematic or unviable.

For a system model that lies in the top half of Fig. 3, an alternative and potentially viable approach to validation would be to engage in a programme of scientific research to translate the model diagonally down and leftwards using the Lakatos MSRP approach. This would involve planning, conducting and interpreting scientific research on the system of interest so that the body of empirical data is enhanced, and at the same time, the scientific knowledge of the system is increased which means that information about the system gained through systematic study is organised by principles that are accepted by the stakeholders. This is the process of engineering research with which many are familiar and that solves the validation problem by translating the model towards the bottom left corner of the diagram in Fig. 2. However, it is often time-consuming and usually costly, sometimes to a prohibitive level as in the case of large space structures, hypersonic flight and fusion reactors. Hence, the process described in the previous section provides some guidance for establishing the usefulness of a system model that remains untestable and possibly unprincipled.

The untestability of models mandates a shift from an objectivist approach strongly based in empiricism to a relativist approach. This shift implies a greater need to make value judgments, which in turn increases the need to be cognizant of bias which has been identified as arising in any, or all, of the following four phases: characterisation of model requirements; data collection; preliminary design; and final design evaluation [52]. Feinstein and Cannon [20] argue that stakeholders should not attempt to achieve 'a mythical, value-free objectivist environment' but to understand their prejudices and evaluate information on its merits. This evaluation could involve estimating uncertainties. These are likely to be multi-faceted for untestable and unprincipled models, including: parameter uncertainty because the untestability makes it unviable to establish the values of the input parameters of the model; parametric variability arising for instance from manufacturing of the engineering system; structural uncertainty when the underlying physics is unknown; algorithmic uncertainty from the discretization process; experimental or data uncertainty from any measurements that are possible; and interpolation uncertainty from the lack of available data. The term 'interpolation uncertainty' does not really cover the dominating uncertainty that maybe present with untestable models, which is associated with an inability to quantify the real-world behaviour of the complex system, and hence could be termed 'reality uncertainty'. The quantification of uncertainty in complex systems is a current topic of active research for both mathematicians[1] and engineers tackling complex aerospace systems[2]; and hence is beyond the scope of this study.

The approach to establishing the credibility of models of complex engineering systems is based heavily on pragmatism, and in this sense, it follows the approach of 'operators' who, according to Forrester [reported by Barlas and Carpenter [19]], see validity as 'relative usefulness' as opposed to 'observers' and the literature which considers it as a 'formal logic concept rather than a pragmatic issue'. It acknowledges that some models are better than others as a result of possessing a greater authenticity while not necessarily being high fidelity representations of the real-world. And, it recognises that it may not be viable or possible to build a computational model that is entirely representative of a complex multi-physics engineering system and its interactions with its surroundings.

6. Conclusions

The normative approaches to validation of engineering models are largely reductionist and empirical with an emphasis on demonstrating that a model cannot be falsified within the scope of its intended use. This imposes severe constraints on the use of computational models of complex multi-physics engineering systems for which empirical data is unavailable due to a lack of appropriate resources, time or capability; thus, rendering the models untestable. In addition, such models might also be unprincipled in the sense that they are built on unknown physics, which makes the use of deductive rationalist approaches to validation impractical. After reviewing the wide range of philosophical positions available to support model validation using a matrix categorisation, a relativist approach to establishing the credibility of such models is described and is based on the epistemic values of the model, its theoretical ancestry and the credentials of the model-building techniques. The focus of describing the usefulness of an untestable model has been moved from validation to credibility, in recognition that the physical realm of interest cannot be measured and hence predictions cannot be compared to it, and that the willingness of a decision-maker to use the predictions of a model defines its usefulness. The value-based approach requires experts to make decisions, acknowledging their biases and bringing their expertise to bear on the specifics of the situation, which approximates to Kuhnianism. The importance of social interaction between experts and the interaction between the individual and the model are highlighted. Finally, a list of credibility factors is provided that could form an agenda for such interactions.

Acknowledgements

Discussions with Dr Ravi Chona from the Structural Sciences Center, U.S. Air Force Research Laboratory, Dayton, OH 45433, during the 1st International Aerospace Symposium on Acoustic Fatigue in September 2017 led to the preparation of this paper and are gratefully acknowledged. The effort by the University of Liverpool reported here was sponsored by the Air Force Office

[1] See for example current scientific programme of Isaac Newton Institute for Mathematical Sciences on Uncertainty quantification for complex systems: theory and methodologies, January 3rd, 2018 to June 29th, 2018: www.newton.ac.uk/event/unq (last accessed on January 10th, 2018).

[2] See 2017 Year in Review by AIAA under Applying Uncertainty Quantification to Complex Systems: https://aerospaceamerica.aiaa.org/year-in-review/applying-uncertainty-quantification-to-complex-systems/ (last accessed on January 10th, 2018).

of Scientific Research, European Office for Aerospace Research and Development under grant number FA9550-16-1-0091. The U.S. Government is authorized to reproduce and distribute reprints of Governmental purpose notwithstanding any copyright notation thereon. Lt. Col. Dave Garner (EOARD) and Dr Jaimie S Tiley (AFOSR) respectively were the program officers for this grant.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.jsv.2019.01.051.

References

[1] B.A. Miller, J.J. McNamara, S.M. Spottswood, A.J. Culler, The impact of flow induced loads on snap-through behavior of acoustically excited, thermally buckled panels, J. Sound Vib. 330 (2011) 5736–5752.
[2] ASME V&V 10-2006, Guide for Verification & Validation in Computational Solid Mechanics, American Society of Mech. Engineers, New York, 2006.
[3] Computational Fluid Dynamics Committee on Standards, Guide for the Verification and Validation of Computational Fluid Dynamics Simulations, American Institute of Aeronautics and Astronautics, January 1998. AIAA G-077-1998.
[4] P. Tiso, A multiscale, parametric reduced order model for thermo-structural dynamics, in: 1st International Aerospace Symposium on Acoustic Fatigue, Glasgow, September 11-12, 2017.
[5] C.M. Sebastian, E. Hack, E.A. Patterson, An approach to the validation of computational solid mechanics models for strain analysis, J. Strain Anal. Eng. Des. 48 (2012) 36–47.
[6] G.S. Fishman, P.J. Kiviat, The statistics of discrete-event simulation, Simulation 10 (1968) 185–195.
[7] R.L. Van Horn, Validation of simulation results, Manag. Sci. 17 (1971) 247–258.
[8] R.G. Sargent, Validation of simulation models, in: Proc. Winter Simulation Conference (WSC'79), vol. 2, 1979, pp. 497–503.
[9] O. Balci, R.G. Sargent, A methodology for cost-risk analysis in the statistical validation of simulation models, Commun. ACM 24 (11) (1981) 190–197.
[10] R.E. Shannon, Tests for the verification and validation of computer simulation models, in: Proc. 13th Winter Simulation Conference, vol. 1, 1981, pp. 573–577.
[11] O. Balci, How to assess the acceptability and credibility of simulation results, in: Proc. IEEE Winter Simulation Conference, 1989, pp. 66–71.
[12] J.S. Carson, Convincing users of model's validity is challenging aspect of modeler's job, Ind. Eng. 18 (1) (1986) 74–78.
[13] E.E. Klein, P.J. Herskovitz, Philosophical foundations of computer simulation validation, Simulat. Gaming 36 (3) (2005) 303–329.
[14] G.B. Kleindorfer, L. O'Neill, R. Ganeshan, Validation in simulation: various positions in the philosophy of science, Manag. Sci. 44 (8) (1998) 1087–1099.
[15] E.A. Patterson, M.P. Whelan, A framework to establish credibility of computational models in biology, Prog. Biophys. Mol. Biol. 129 (2013) 13–19.
[16] E.A. Patterson, On the credibility of engineering models and meta-models, J. Strain Anal. 50 (4) (2015) 218–220.
[17] M. Tegmark, Our Mathematical Universe: My Quest for the Ultimate Nature of Reality, Allen Lane, London, 2014.
[18] K. Popper, The Logic of Scientific Discovery, Harper and Row, New York, 1959.
[19] Y. Barlas, S. Carpenter, Two philosophical roots for model validation: two paradigms, Syst. Dynam. Rev. 6 (2) (1990) 148–166.
[20] Feinstein Ah, H.M. Cannon, A hermeneutical approach to external validation of simulation models, Simulat. Gaming 34 (2) (2003) 186–197.
[21] T.H. Naylor, J.M. Finger, Verification of computer simulation models, Manag. Sci. 14 (2) (1967) B92–B101.
[22] S. Fuller, Social Epistemology, second ed., Indiana University Press, Bloomington, IN, 2002.
[23] A.I. Goldman, A guide to social epistemology, in: A.I. Goldman, D. Whitcomb (Eds.), Social Epistemology: Essential Readings, Oxford University Press, Oxford, 2011, pp. 11–38.
[24] M. Welbourne, The community of knowledge, Phil. Q. 31 (1981) 302–314.
[25] M. Germonpré, J.C.O. Nielson, G. Degrande, G. Lombaert, Contributions of longitudinal track unevenness and track stiffness variation to railway induced vibration, J. Sound Vib. 437 (2018) 292–307.
[26] A. Bajrić, J. Høgsberg, Identification of damping and complex modes in structural vibration, J. Sound Vib. 431 (2018) 367–389.
[27] C. Carmicino, G.D. Di Martino, Analytical modeling of the one-dimensional acoustic field in hybrid rocket combustion chamber, J. Sound Vib. 437 (2018) 180–197.
[28] T. Xiao-Ming, Y. Zhi-Gang, T. Xi-ming, W. Xiao-long, Z. Jie, Vortex structures and aeroacoustic performance of the flow field of the pantograph, J. Sound Vib. 432 (2018) 17–32.
[29] Q. Hu, L. Su, Y. Cao, J. Zhang, Decentralized simple adaptive control for large space structures, J. Sound Vib. 427 (2018) 95–119.
[30] R. Viswanathan, Damage Mechanism and Life Assessment of High Temperature Components, ASM International, Metals Park OH, 1989.
[31] S.R. Yeratapally, M.G. Glavicic, M. Hardy, M.D. Sangid, Microstructure based fatigue life prediction framework for polycrystalline nickel-based superalloys with emphasis on the role played by twin boundaries in crack initiation, Acta Mater. 107 (2016) 152–167.
[32] H.Y. Li, J.F. Sun, M.C. Hardy, H.E. Evans, S.J. Williams, T.J.A. Doel, P. Bowen, Effects of microstructure on high temperature dwell fatigue crack growth in a coarse grain PM nickel based superalloy, Acta Mater. 90 (2015) 355–369.
[33] M.R. Bache, J. O'Hanlon, D.J. Child, M.C. Hardy, High temperature fatigue behaviour in an advanced nickel based superalloy: the effects of oxidation and stress relaxation at notches, Theor. Appl. Fract. Mech. 84 (2016) 64–71.
[34] E. Winsberg, Science in the Age of Computer Simulation, The University of Chicago Press, Chicago, 2010.
[35] M.W. Rohrer, Seeing is believing: the importance of visualisation in manufacturing simulation, in: Proc. 2000 Winter Simulation Conf., 2000, pp. 1211–1216.
[36] L.W. Schruben, Establishing credibility of simulations, Simulation 34 (1980) 101–105.
[37] J. Biddle, E. Winsberg, Value judgements and the estimation of uncertainty in climate modelling, in: P.D. Magnus, J. Busch (Eds.), New Waves in Philosophy of Science, Palgrave MacMillan, Basingstoke, 2010.
[38] E. McMullin, Values in science, in: Proc. Biennial Meeting of Philosophy of Science Association, 1982, pp. 3–28, 1982(2).
[39] J. Kuorikoski, A. Lehtinen, C. Marchionni, Economic modelling as robustness analysis, Br. J. Philos. Sci. 61 (2010) 541–567.
[40] W. Wimsatt, Robustness, reliability and over-determination, in: M. Brewer, B. Collins (Eds.), Scientific Inquiry and Social Sciences, Jossey-Bass, San Francisco, 1981.
[41] J. Raerinne, Robustness and sensitivity of biological models, Philos. Stud. 166 (2013) 285–303.
[42] J.E. Bailey, Mathematical modelling and analysis in biochemical engineering: past accomplishments and future opportunities, Biotechnol. Prog. 14 (1998) 8–20.
[43] I. Hacking, Representing and Intervening: Introductory Topics in the Philosophy of Science, Free Press, New York, 1983.
[44] R. Rudner, The scientist qua scientist makes value-judgments, Philos. Sci. 20 (1953) 1–6.
[45] T.S. Kuhn, Objectivity, value judgement and theory choice, in: T.S. Kuhn (Ed.), The Essential Tension: Selected Studies in Scientific Tradition and Change, University of Chicago Press, Chicago, 1977.
[46] E.A. Patterson, A. Worth, M.P. Whelan, On the Scientific Credibility of Toxicological Methods Used in Chemical Risk Assessment, Royal Society Open Science, (in review).
[47] CEN, Validation of Computational Solid Mechanics Models, Brussels: Comite Europeen de Normalisation (CEN), CWA 16799(2014), 2014.

[48] N.Y. Yang, M.M. Kumaraswamy, H.J. Pam, G. Malush, Integrating qualitative and quantitative methodology to assess validity and credibility of models for bridge maintenance management system, J. Manag. Eng. 27 (3) (2011) 149–159.
[49] D. Wu, D.W. Rosen, J.H. Panchal, D. Schaefer, Understanding communication and collaboration in social product development though social network analysis, ASME J. Comput. Inf. Sci. Eng. 16 (2016), 011001.
[50] S.A. Sloman, N. Rabb, Your understanding is my understanding: evidence for a community of knowledge, Psychol. Sci. 27 (11) (2016) 1451–1460.
[51] F. Capri, P.L. Luisi, The Systems View of Life: a Unifying Vision, Cambridge University Press, Cambridge, 2014.
[52] S.R. Irvine, R.R. Levary, M.S. McCoy, The impact of judgment biases on the validation of simulation models, Simulat. Gaming 29 (2) (1998) 152–164.
[53] A. Franklin, The Neglect of Experiment (chapter 6.3), Cambridge University Press, Cambridge, 1986.