Alternative Approaches to Validation

Emily Keefe, Stephen C Moore

Dstl/WP13453 V1.0
June, 2005

PCS
Dstl
Ively Road
Farnborough
GU14 0LX

Intellectual Property Department
Defence Science and Technology Laboratory
Porton Down, Salisbury, Wiltshire, SP4 0JQ.

Table of contents

INTRODUCTION

1. This working paper has been produced as part of the ongoing research being conducted by Dstl on behalf of the MoD's Research Acquisition Organisation (RAO) Human Capability Domain. The aim of this paper is to stimulate discussion on alternative methods of validating the human and social component of tools and techniques used to aid the understanding of human and social behaviour, for which the traditional realist approach to validation may not be feasible or appropriate.

2. The realist approach to validation is to determine the accuracy with which a model represents a real world situation by comparing model outputs with behaviour in that situation in reality. This approach is traditionally used for the validation of models, tools and techniques. However, difficulties are often encountered when trying to use the realist approach to validate human or socio-technical[1] models. Outputs of models concerned with human behaviour are often intended to facilitate understanding of a system rather than to definitively predict a future outcome. The nature of such outputs, combined with a lack of accessible data and difficulties in defining the cause of observed real world activities, make it difficult to validate against real events and thus adopt the realist approach.

PAPER STRUCTURE

3. This discussion paper introduces the concept of validation and the nature of the realist approach. It then outlines the key philosophical and theoretical problems with the realist approach to validation of models that address human and social behaviour. In the context of this paper, when a 'model' is referred to it is not necessarily a formal or computational model of a system as traditionally associated with the term. Instead, a 'model' could be something less formal, or more qualitative, such as a theory, process or technique for structuring thoughts or problems. Similarly, when 'outputs' are referred to they are not necessarily numerical or predictive. Instead, they might be offering greater insight into a problem area or an improved understanding of a situation. For ease of reading, models and outputs should be understood to be relatively broad terms in the context of this paper.

4. Alternative approaches to validation that may be used to address problems associated with the realist approach are then outlined. For each approach there is a summary of the concept, followed by a consideration of what aspects of the theory are useful and how it could be applied. Finally, a summary of how this fits into the extant Director General (Scrutiny and Analysis) (DG(S&A)) guidelines for Verification and Validation (V&V) within MoD is included.

5. To ease navigation around this document and to allow sections of particular relevance to be identified, the broad content and paragraphs numbers of sections of this report are as follows:

   a. Introduction to Validation - paragraph 8.

   b. Description of the realist approach - paragraphs 9 to 14.

   c. Philosophical and theoretical issues with the realist approach - paragraphs 15 to 27.

[1] Socio-technical systems comprise both human and non-human components (i.e. technical or physical) and are defined as 'the complex, interrelated, human and physical networks of a regime, its forces and their capabilities'.

d. Alternative approaches to validation - paragraphs 28 to 72.

e. A tabular summary of approaches to validation - paragraph 73.

f. Identifying subject matter experts to support validation - paragraphs 74 to 75.

g. Relationship between alternative approaches to validation with the extant DG (S&A) validation guidelines - paragraph 76.

h. Conclusions - paragraphs 78 to 83.

SCOPE OF THE WORK

6. The findings from this research into alternative approaches to validation have been presented as a working paper as the authors believe there are, at present, no firm solutions to the issue of validation of human science models and techniques. The theories that are introduced are of interest but it is not essential to understand them fully in order to use aspects of them for validation purposes. Indeed, the practical application is often quite simple. For this reason, Table 1 at the end of the document (paragraph 76) summarises the authors' view of when and how each approach to validation could be used. There is a large quantity of theoretical research and publications available and this paper is not exhaustive. Further input, alternative views, feedback and comments on this paper are actively sought by the authors.

7. This paper will form the basis of a second document, to be issued later this year, which will focus upon the practical application of ideas and theories discussed. An assessment of the utility of each alternative is anticipated to be included in this follow-on publication.

VALIDATION

CONTEXT

8. Validation is a means of increasing confidence in the accuracy of information derived from a model intended to support decision-making. Validation assesses the underlying assumptions of a model, the compromises and limitations of a model's processes, the associated information, and the model's management. These assessments unavoidably incur costs in terms of time, money and other resources. This raises the question of how to assess when the validation undertaken is sufficient for the circumstances under which the model is being used. There needs to be consideration of the effort required for validation and the potential impact of any decision-making supported by the model. As the impact of a decision increases, so does the need for the decision-maker to be aware of the validity and utility of any outputs from a model that is used to support their decision-making.

REALIST VALIDATION

9. Realist validation seeks to compare the predictive outputs of a model with experiments and trials, or real world events. The realist approach assumes that there is a truth in existence, and that research, modelling and analysis are attempts to discover this truth. Reality is considered to exist independently of the people trying to discover it. Whilst results obtained from experiments are dependent upon the means with which they are sought, and one approach may

yield different results to another, a realist may believe that there is only one way in which a problem can be formulated and solved, and that this approach is part of reality.[2]

THE REALIST APPROACH

INTRODUCTION

10. The realist approach to validation is one of the most convincing and universally accepted. A model validated this way generally attains credibility in the scientific community. This approach is highly appropriate for validating hard science methods and models due to the generally observable and predictable character of the natural sciences. However, there are several important problems that make the realist approach inappropriate for validating models and methods concerned with human behaviour.

11. The following section introduces key issues that should be considered when validating. These issues primarily concern inherent assumptions and origins of knowledge upon which models of human or social behaviour may be based. This problem is known as 'epistemology' and questions 'how we know what we know'. This is particularly important because, unlike theories of the natural sciences, theories of human or social behaviour are not independent from the observer and are inherently open to personal interpretation and bias. It is therefore necessary to be aware of the underlying assumptions of any theory or knowledge that is drawn upon by the tools and techniques that support the improved understanding of human and social behaviour.

LOGICAL EMPIRICISM

12. Logical empiricism is the basis of realism in the social sciences and argues that social behaviour can be objectively observed and explained using the same logical-mathematical language as the natural sciences. This doctrine suggests that only directly observable objects are amenable to scientific investigation and that scientific truth can only be reached through the objective observation of reality, which is considered to be external to the investigator. These observations must then be expressed using logical-mathematical language in order to receive the status of scientific knowledge. Within the context of logical empiricism, a model will only be considered valid if it adheres to these prescriptions.[3]

13. A problem with logical empiricism is the process of induction, in which knowledge is established through objective observation and experimentation.[3] Knowledge acquired through induction is considered to be an accurate translation of reality. However, logical empiricism cannot be completely objective and free from value judgements as, when studying human or social behaviour, there will always be a social, relational and constructed dimension to the investigator's perception of an object. The significance of this is that validation should appreciate that knowledge within a model may be affected by the perceptions and interpretations of the investigator.

14. Furthermore, logical-mathematical language assigns terms for objects observed by an investigator, but these terms are not actually determined by an external 'reality' and only have

[2] ROY, B., 'Decision science or decision-aid science.' *European Journal of Operational Research*, 1993, 66, pp., 184-203.
[3] DERY, R., LANDRY, M., and BANVILLE, C., 'Revisiting the issue of model validation in OR: An epistemological view'. *European Journal of Operational Research*, 1993, 66, pp., 168-183.

meaning and significance within the linguistic system to which they belong. Observation terms only receive their full meaning when they are related to a network of socially and culturally understood concepts.[3] For example, we can only understand the term 'criminal' if we appreciate the meaning we assign to this word within our own culture. This problem leads to similar issues as induction, in that the validation process must appreciate that language used within a model cannot be truly objective and that terms assigned to components necessarily have culturally or socially specific meanings. These terms may be inappropriate for different cultures or populations which the investigator is not part of and may inherently make value judgements about.

ADDRESSING CONCERNS WITH THE REALIST APPROACH THROUGH THE PHILOSOPHICAL AND THEORETICAL PERSPECTIVES

15. A common way to address concerns with a particular approach is to consider it from alternative philosophical and theoretical perspectives. Different perspectives in the social sciences each have their own strengths, weaknesses and abilities to raise concerns with other perspectives. In the context of this paper, they provide additional means and methods to employ when validating the tools and techniques that may be used to support the improved understanding of human and social behaviour.

THE HISTORICAL PERSPECTIVE

16. The historical perspective challenges declarations of absolute knowledge by proposing that truth is relative and varies between different people and times in history. This form of relativism argues that truth and value are relative to an observer or group of observers and that human judgement is conditioned by personal bias. This stands in direct contrast to the realist approach.

17. The historical perspective suggests that scientific knowledge is not objective and is influenced by presuppositions and research strategies. Knowledge production is directed and guides the construction of our perception of reality. A central concept of this perspective is that of 'paradigms'. Paradigms are the social and historical contexts in which research is conducted and they are built upon intellectual axioms that guide scientific research, reasoning and interpretation. Scientific research is guided by social mechanisms within a hierarchic power structure. Young scientists serve apprenticeships and acquire knowledge from their seniors, which reinforces and perpetuates a particular paradigm. Competition for research funds may also result in scientific research being directed to satisfy certain demands and the further perpetuation of a particular paradigm.[4]

18. The key implication of these issues is that any validation must appreciate the paradigm in which the knowledge or hypothesis, upon which a model is based, was acquired. Scientific knowledge is not just determined by relations with objects but by social relations as well and this leads to the questioning of the validity of the knowledge and hypotheses upon which models and methods are based.

[4] HOLLIS, M., *The Philosophy of Social Science: An Introduction*, Cambridge University Press, 2002, pp. 84-86.

THE SOCIOLOGICAL PERSPECTIVE

19. The sociological perspective focuses on how norms and universal criteria of what is considered 'scientific' exist due to scientists operating on consensus. Consensus results when scientists share similar interests and produce knowledge that contributes to a greater body of research. Unanimous acceptance of research as scientific in a specific area affords the consensus of the wider scientific community. This reciprocal recognition within a scientific community is considered to be the basis of the social organisation of science.[3]

20. A programme colloquially known as the 'strong programme' emerged from the sociological perspective in reaction to the tendency to take for granted the absolute nature of scientific knowledge due to consensus within a scientific community. The strong programme acknowledges that scientific research processes are diverse in different times and places, and that this is dependent on the social and historical contexts that structure scientific research. This is similar to the concept of paradigms, except it also argues that scientific knowledge is only one of many social systems of beliefs. Other types of knowledge, such as religion or politics, should not be dismissed as irrational or false because they do not meet a wider scientific consensus.[3]

21. The arguments of the strong programme have two key implications for validation. It is implied that for models to be valid they must be based on theories or knowledge that are socially credible and fall within a general consensus of legitimacy. To be able to persuade, a model must have legitimacy within the community for which it will produce recommendations. There is also the need for validation to appreciate what theories or knowledge a model is based upon and to question why these beliefs are more appropriate than any alternatives. Competing (but potentially more appropriate) beliefs may have been dismissed at an earlier stage because they were not considered to be 'scientific' by general consensus at the time.

CRITICAL THEORY

22. Within the critical approach it is believed that a social scientist cannot be independent from their subject matter as they are part of the society that they study. Critical theories propose that there is a close connection between knowledge and power because theory actually affects the way we think about the world that we live in. The realist approach encourages a particular view of the world by making certain power structures and social relations appear to be the norm and this affects the distribution of power in society.

23. The key implication of the critical approach is the concept that theories are always *for* someone and *for* some purpose. The realist approach is purportedly objective, however, it has underlying views and assumptions that reinforce a particular viewpoint. It is therefore important for validation to appreciate what viewpoints or patterns of institutions and relationships are reinforced and legitimised by the theories and knowledge upon which a model is based.[5]

[5] SMITH, S., 'Reflectivist and constructivist approaches to international theory' in BAYLISS, J., and SMITH, S., (eds.), *The Golbalisation of World Politics: An Introduction to International Relations* (Oxford University Press, Oxford, 2001, pp., 224-49.

POST-MODERNISM

24. Post-modernism is a highly complex approach to analysing theories of human and social behaviour and has often been perceived as a relatively extreme theoretical development. However, the post-modernist movement raises several important issues that are particularly relevant to the validity of the realist approach and logical empiricism. One of the key goals of post-modernism is to deconstruct and distrust any approach that claims to have direct access to 'the truth' or that accurately represents 'reality'. Post-modernism refers to these approaches as meta-narratives, which means they create an image of the world supposedly from an external and objective perspective, and it has two central concepts that analyse and critique these meta-narratives. These two concepts are power-knowledge relationships and textual strategies.[5]

25. The concept of power-knowledge relationships proposes that knowledge is not immune from power as it relies on and reinforces existing power relationships. Knowledge is not external to a social context but is actually part of it, and a key question posed by post-modernism is how can history have a truth if truth has a history? Post-modernists approach the study of history using an approach known as genealogy, which emphasises that there is no truth, only regimes of truth. These regimes reflect the ways that power and knowledge have developed through history in a mutually sustaining relationship. Therefore, as with paradigms, statements about the world are only 'true' within a specific social and historical context. The key implication of this concept is that validation should be aware that the 'truth', upon which a model is based, is only one regime of 'truth' amongst many possible alternatives. Validation should question why this particular regime of truth was selected for a model and which power relationship this regime of truth may serve to reinforce.[5]

26. The concept of textual strategies proposes that the social world has been constructed textually. The world is constituted like text in the sense that interpreting the world reflects the concepts and structures in our language. There are two main ways of exposing these textual strategies, which are deconstruction and double-reading. Deconstruction is a process of showing that seemingly natural terms and concepts are actually artificial linguistic constructs, and that these are arranged hierarchically in opposition where one term is always privileged over the other (i.e. rich/poor, good/bad, right/wrong). Double-reading is a way of revealing how these seemingly natural concepts operate by subjecting a text to two readings. The first reading would be a repetition of the dominant interpretation to demonstrate how it achieves coherence but the second reading would highlight the internal tensions that result from the use of seemingly natural concepts. The aim is not to reach a 'correct' interpretation of the text but to reveal how knowledge can be interpreted in different ways. The key implication of this concept is that validation should be aware of how the knowledge, upon which a model is based, has been constructed by an investigator and how this knowledge could have been interpreted differently.[5]

SUMMARY AND IMPLICATIONS

27. The philosophical and theoretical perspectives introduced above raised a number of problems with the realist approach and the possibility of defining a social "reality". Table 1 summarises these problems and the implications of these perspectives for the conduct of a validation process that is more appropriate for the tools, techniques and methods for understanding and predicting human and social behaviour.

Table 1: Problems with the realist approach and implications for more appropriate validation process

Philosophical and Theoretical Perspectives.	**Problems with the realist approach raised by these perspectives.**	**Implied validation checks required for a more appropriate validation process.**
The Interpretative Perspective	Society and human behaviour cannot be observed and explained entirely objectively like the natural sciences. It can only be interpreted subjectively by the investigator. Models will therefore necessarily be a subjective interpretation of the subjects and behaviour being modelled.	Validation should query the source data that a model is based upon and whether alternative data, collected using a different approach, would be more appropriate. It should also query whether an investigator's interpretation of society and human behaviour appropriately reflects an interpretation that could be given by the subjects themselves.
The Historical Perspective	Knowledge of society and human behaviour cannot be absolute. The underlying theories upon which models may be based do not necessarily explain an indisputable and unchanging social "reality". The views and assumptions in these theories may only reflect the focus of the dominant research programmes at that time. These are known as paradigms.	Validation should identify the scientific paradigm underlying the theories used to support the tasks of understanding and predicting human and social behaviour. It should assess what particular view or methodology this paradigm encourages and whether it is appropriate for these tasks, in comparison to other alternatives.
The Sociological Perspective	Our knowledge and theories of "reality" are not indisputable and unchanging. They are relative to time and context. They reflect change over in time in line with new research discoveries and changes in the wider consensus of what is considered to be "true".	Validation should identify whether the theories and knowledge, upon which models may be based, are credible with peers at that time. It should also periodically re-evaluate over time to account for changes and developments in theories, knowledge and the wider consensus of peers.
Critical Theory	Theories, and their inherent elements, assert a certain viewpoint. This reinforces the perception that these elements are necessarily true and indisputable. These theories may therefore legitimise an inaccurate or false explanation or interpretation of the society or human behaviour that is being modelled.	Validation should identify the key elements of the theories, upon which models may be based, and query why these elements are included or prioritised at the expense of others. It should also query what viewpoint and perception of reality is reinforced and legitimised by adopting a certain theory and its inclusive elements.

Philosophical and Theoretical Perspectives.	Problems with the realist approach raised by these perspectives.	Implied validation checks required for a more appropriate validation process.
Post-Modernism	The "truth", upon which a model may be based, is actually only one interpretation of truth amongst many possible alternatives. We constantly interpret and reconstruct "truth". Our knowledge of society and human behaviour is therefore an interpretation, which is reconstructed in different ways by the researcher, historian, theorist etc.	Validation should reflect on how a modeller has interpreted and incorporated theories, knowledge or historical explanations into a model. It should also identify how different a model or assessment might be if the original source knowledge actually offered a different interpretation of its subject matter.

ALTERNATIVE APPROACHES TO VALIDATION

INTRODUCTION

28. The realist approach claims to be able to identify the 'truth' or 'reality' of the social world in the same way as is possible for the natural world. However, the previous section has highlighted several key issues with the way in which the realist approach acquires and creates knowledge that makes it inappropriate for validating models of human and social behaviour. The approach is not suitable for models / methods designed to facilitate understanding of a socio-technical system rather than definitively predict a future outcome as there are no 'real' events to validate against. In light of these problems the authors suggests that the realist approach is inappropriate for the validating the human and social behaviour aspects of tools and techniques used to support the improved understanding of such behaviour. The following section provides a series of alternative approaches to validation that could be utilised. These approaches are assessed on their compatibility with epistemological issues already raised and their appropriateness in supporting validation of models and techniques that address human or social technical systems.

THE INTERPRETATIVE APPROACH TO VALIDATION

29. The interpretative approach in social science is the principal alternative to realism. The key proposition is that we cannot explain the social world objectively because we understand and interpret it subjectively. We must appreciate how we acquire knowledge and how we attach subjective meaning to human and social behaviour. In this approach, a distinction is made between explaining (*erklären*) and understanding (*verstehen*) human behaviour. '*Verstehen*' has evolved into a holistic concept that requires an investigator to understand the underlying values, morals and ethics of societies and specific patterns of behaviour. These are called social norms and *verstehen* argues that they are not objectively observable facts in the same way as objects in the natural sciences are.[6]

30. Within the verstehen concept it is argued that human actions have meanings specific to those involved in a situation. An investigator should communicate with their subjects in order to understand their behaviour and social processes, and to help avoid imposing bias and personal perception on the interpretation of actions observed. Within realism, subjects are viewed

[6] HOLLIS, M., *The Philosophy of Social Science: An Introduction,* Cambridge University Press 1994, p. 147.

externally and treated as objects. If a model is tested and errors observed, then it is adjusted to match 'real world data' before testing again. In contrast, the interpretative approach stresses the possibility of an investigator misunderstanding any given action. Interpretative validation seeks to resolve misunderstandings encountered during testing by consulting with participants in the system that is being analysed and modelled.[7]

31. There are strong grounds for believing that an interpretative approach is more appropriate for validating models of human behaviour than the realist approach. Epistemological problems are addressed through the way in which knowledge of the social world is acquired. The approach does not try to prove that a social 'reality' can be knowable in a direct way without interpretation. Instead it is concerned with accurately interpreting human and social behaviour from the perspective of the subjects involved. Measuring and recording of this behaviour should refer to rules, norms and principles of the society in which it took place. Through this process, validation should be able to help ensure that models and methods used to support the improved understanding of human and social behaviour are as accurate and appropriate as possible.

32. This interpretative approach to validation should ideally involve discussion with subjects from the system being modelled, yet this is not always possible or appropriate. Under such circumstances, validation could instead involve consultation with appropriate and suitably informed Subject Matter Experts (SMEs).[8] SMEs may be a viable alternative for validating the interpretation of a particular pattern of behaviour being modelled.

VALIDATION BY CHECKING THE CONSTITUENT ELEMENTS

33. Models may be based upon a particular phenomenon, supporting data, and transformations undertaken according to certain theories or rules. This approach concentrates on confirming that these constituent elements of a model are valid, rather than comparing the outputs of a model with reality. Checking each of the constituent elements may be used as an approximation to overall validation and may be sufficient to give confidence in the entire model.[9] In some ways this process is similar to computer science verification, in which each section of code is tested before being added to the model and the whole model tested with the new code in place.

34. Checking of constituent elements is similar to the interpretative approach and has significant relevance for the validation of models of human behaviour. Models are constructed from different pieces of information and each constituent element may fall into a different area of expertise. The validation of these elements is a check on whether the source phenomenon, supporting material and internal processes are trustworthy and this may require consultation with relevant SMEs from different areas of expertise. The SMEs, once identified, can assess whether interpretations of subject matter, supporting material and data sources used within the model are reasonable. These activities could be undertaken on a rolling basis throughout modelling. The main difficulty with this approach may be the identification of, and access to, appropriate SMEs.

[7] FROST, M., *Ethics in International Relations*, Cambridge University Press, 1996, pp. 23-28.
[8] Paras 74 to 75 provide greater detail on the role of SMEs in validation.
[9] MISER, H., J., 'A foundational concept of science appropriate for validation in operational research.' *European Journal of Operational Research*, 1993, 66, pp., 204-215.

CONSTRUCTIVIST VALIDATION.

35. Constructivist validation involves looking for 'keys' that are considered to be suitable for developing and evolving beliefs. These keys are different concepts, models or procedures that can help an investigator to understand, organise or develop a situation. The aim of constructivism is not to discover an existing truth, but to construct a set of these keys that are appropriate to the subjects' objectives and values. The selection and development of keys should be related to working hypotheses, which should be suitable for the subject matter and should result in recommendations for the decision-maker. The overall goal within constructivism is to add new insights.[2]

36. There are two central aspects to constructivist validation; whether outputs or recommendations are accepted and whether they are productive. Keys must provide useful outputs that serve the purpose for which they were developed in order to be valid. If understanding and recommendations add value, within the context for which they were developed, then they have served their purpose and are valid for it.

37. Constructivism is not significantly undermined by the epistemological issues raised and may consequently be more appropriate than the realist approach for validating models of human or social-technical systems. Constructivism often involves conceptualising criteria and presenting these as a description of reality. Whilst the realist approach argues that these are rooted in an objective reality, the constructivist approach acknowledges that these are not 'true values' but approximations intended to help us deal with complex situations. These indices and criteria are not overtly deterministic and are only considered to be a basis for reasoning and comparison.

38. Recommendations are dependent upon keys used to understand a situation. As keys may be social constructs, these recommendations are only valid for the situation within which they were made and it is not possible to generalise for a multitude of different scenarios. Also, because outputs are wholly dependent upon keys, which are in turn wholly dependent upon the context within which they were developed, the acceptance of outputs infers acceptance of the keys. Therefore, the constructivist approach to validation depends on whether there is a community of people interested in the keys and accepting of the outputs that they produce.

AXIOMATIC VAILDATION

39. Axioms are principles, rules or frames of reference for working hypotheses. If these principles are accepted then a model for representing a situation should be created in accordance with them. Axiomatic validation tests whether axioms have been applied correctly and appropriately. This involves confirming that a model conforms to underpinning hypotheses, evaluating whether the model is relevant (i.e. purposeful within a specific context) and ensuring that it has been competently built. This process of validating should involve consultation with relevant SMEs. For example, if theories of behaviour have been used then it would be appropriate to seek input from psychologists. Similarly, if theories specific to a country of interest have been used, then an SME on this country should be consulted.[2]

40. This approach to validation is similar to the interpretative approach as it involves consultation with relevant experts, but differs as SME consultation seeks to confirm that the model accurately conforms to the axioms, not that the axioms are correct in the first place. The main weakness with this approach is that it leads us to accept the axioms and to look for evidence to

support the outputs of the model in order to justify the underlying hypothesis. It does not lead us to question whether the underpinning hypothesis itself is valid. An axiom is not necessarily the correct, or only, formulation of rationality that it may appear to be. To validate results though the 'truth' suggested by axioms is not scientific validation and leads to the problem of paradigms discussed previously; accepting a particular hypothesis as a starting point limits the resulting model and reinforces the paradigm from which the hypothesis originated. There is value therefore in challenging properties which may seem to be natural or 'right'. Further, accepting individual axioms does not mean that the collective group should also be accepted.[2]

41. Axiomatic validation is most appropriate for models that are built on hypotheses which are relatively uncontested and without any significant debate regarding their assumptions. This may not always be the case for hypotheses that address the human and social behaviour where underlying assumptions may be debatable. This approach is also not appropriate for models which do not have a specific underpinning theory and therefore no axiomatic framework to validate against.

VALIDATION BY FALSIFICATION

42. Falsification is the process through which a theory is offered to the scientific community to challenge and attempt to disprove it. An alternative approach to validation is suggested by the falsification approach, in which the difficulties of logical empiricism are resolved while preserving the goal of identifying social truths through formal scientific practices. This process is deemed to be central to the production of scientific knowledge and essential for a model to be valid. The scientific character of knowledge depends on the way it is produced and the internal coherence of a theory is a necessary condition of its validity.[3]

43. The main strength of this approach is that it can lead to a rigorous validation process, which can refute key aspects of an underlying theory and therefore invalidate key components of a model. According to the falsification approach a model can never be truly valid and a model's validity therefore lies in the tests to which it is subjected. However, falsification is only a process of validating underlying theories and does not confirm the outright validity of a model.[3]

44. Falsification concentrates on justifying or refuting theories and does not challenge the origin of these theories. This does not address the concern that theories and knowledge of the social world cannot be induced objectively from an external reality. This means that falsification may not be appropriate for validating contentious theories of human and social behaviour.[3]

INSTRUMENTAL VALIDATION

45. The overall concept of the instrumental approach is that the truth of knowledge is to be found in its practical character. Knowledge is only genuinely scientific if it is a useful instrument. Instrumental validation is often considered as a complement to the falsification approach as it is a way of explaining why certain theories are not rejected, despite their refutation, because they are successful in producing practical outputs. This approach argues that models are constructed for practical application in particular situations and are completely instrumental. Therefore, validation should not question whether a model is true, only if it useful.[3]

46. This approach is useful if it can be accepted that the purpose of validation is not to question the validity of the underlying assumptions and knowledge upon which a model is based. The

instrumental approach is intended to validate whether a model provides practical recommendations, yet the practicality of recommendations is dubious if the outputs are based upon inaccurate or inappropriate hypotheses. As the instrumentalist approach does not validate important underlying elements it fails to evaluate one of the most important aspects of any model used to support the improved understanding of human and social behaviour.

EXPERIMENTAL VALIDATION

47. A conventional experiment is a regulated, scientific procedure that results in findings that can be subjected to quantitative, mathematical or statistical validation techniques. Experimental validation focuses on the type and quality of outputs produced by experiments. It is also concerned with the techniques used to produce findings and the efficiency of procedures adopted during an experiment. The quality of findings are considered in terms of the sensitivity of the model to changes in input parameters and the level of insight gained, acceptability, applicability and usefulness in making a decision.[10]

48. This approach is similar to instrumentalist validation which considers whether knowledge provided is useful. In experimental validation it is necessary to determine who is going to be using model outputs and for what purpose. These individuals or groups would then need to be approached to provide an assessment of the scope and quality of the model's outputs. In reality, it may not be possible to identify or access a generic set of end users and an intermediate user may have to be used as an alternative. An assessment of the quality of solutions could be undertaken through the use of relevant interviews and or questionnaires.

49. The concern with this approach is that it concentrates on the utility of a model's outputs, without validating whether the underlying hypothesis is appropriate.

OPERATIONAL VALIDATION

50. The operational approach to validation is concerned with outputs produced by a model rather than the processes through which they were obtained. Operational validation considers the recommendations produced by a model, to support a decision, in terms of the following criteria:[9]

    a. Usability: Resource availability, training requirements etc.

    b. Utility: Level of understanding or insight gained, enhancement in communications etc.

    c. Timeliness: How quickly solutions can be obtained.

    d. Synergism: Whether current decision complements previous decisions.

    e. Cost: Implementation costs, equipment, training, data etc.

51. The main strength of this approach is that it encompasses many of the criteria against which a specific recommendation should intuitively be assessed. Validation requires the identification of an appropriate assessor for each criterion, who is questioned through a discussion, interview or questionnaire as appropriate. The appointment of these assessors is dependent upon the criteria, for example an assessment of the time required to produce results could be

[10] ORAL, M., and KETTANI, O., 'The facets of the modelling and validation process in operations research.' *European Journal of Operational Research*, 1993, 66, pp., 216-234.

undertaken by the modeller whereas the utility of a recommendation should be an end user assessment.

52. However, operational validation is less appropriate for models of human behaviour because the outputs are often qualitative, such as insight into a situation or behaviour of an individual. These can be difficult to quantify and evaluate through operational validation.

APTNESS VALIDATION

53. This approach considers the same components as operational validation (usability, utility, timeliness, synergism and cost) but does so from the perspective of potential model users. Aptness validation produces information to help potential users to accept or reject a model through consideration of the following:[9]

    a. Insight generating capacity.

    b. Descriptive realism.

    c. Mode reproduction ability: replicating the important modes of dynamic behaviour observed in the real system.

    d. Transparency.

    e. Relevance: whether the model addresses issues considered to be important by SMEs.

    f. Ease of enrichment: whether the model be modified.

    g. Fertility: whether it generates new ideas or approaches.

    h. Formal correspondence with data.

    i. Point prediction ability.

54. The main strength of this approach is that it does have limited scope for validating the underlying assumptions and hypotheses of models. The consultation with SMEs to confirm a model's relevance is an important aspect for any validation process that evaluates models of human and social behaviour.

55. The main weakness in this approach is that it necessitates a good working relationship between the modeller and the end users. This relationship may be similar to that described within operational validation, where the end user is first briefed on the model and outputs produced, followed by an interview, questionnaire or discussion to assess the criteria listed. If this working relationship is not positive or constructive then this validation process is unlikely to be successful.

BLACK-BOX VALIDATION

56. This form of validation is interested only with the input : output relationships of the model. It does not consider the internal processes that cause the relationships to exist. Black-box validation is concerned with the predictive power of a model, and considers whether the outputs from that model are distinguishable from reality.[11] It is this approach to validation which is the essence of the Turing Test (TT). Within the TT, an interrogator is connected via a

[11] PIDD, M. *Tools for Thinking*, John Wiley & Sons, Inc, 1996.

terminal to a person and a machine. By asking questions, they are required to identify which is which.[12]

57. A key problem with this approach is that, in black-box modelling, the internal processes of the model do *not* reflect reality. A black-box model is supposedly calibrated to ensure realism but this raises the key problem of who decided what is 'real'. The key requirement for a black-box is only that the input : output relationships are the same, and therefore the process through which this is achieved is not of interest. A black-box model is considered to be a relatively good predictor of behaviour, within routine situations, and sufficient to represent non-central and non-critical actors within a system, provided that the actors only act in a routine and understood way. However, this approach does not validate the underlying human cognitive and social processes within the model and this should be a central task for models of human and social behaviour.

58. If black-box modelling is to be used, it should be carefully considered whether the underlying processes are of importance. If they are not, then validation could be undertaken by comparing the actions of these actors within the model to the real world situation that they represent, using open-source literature or historical data. The model is valid as long as the behaviour of the actor in the model matches the behaviour of the actor in reality. If these underlying processes are considered important then it can be argued that, if the actors in the system are not understood and simply replicated, the system will not reflect a real world situation in unusual or unprecedented situations.

59. The utility of this approach is dependent on its fitness for purpose. The validation process may conclude that the use of this model is satisfactory for routine situations, where the black-box model may be a relatively good predictor. However, it is recommended that black-box modelling would be less satisfactory for the non-routine situations, where the underlying processes are not understood and the outputs of the model are less likely to be reliable or accurate. Furthermore, this technique may be considered satisfactory for decisions regarding peripheral members of a target system. However, if the underlying processes are not understood then the outputs may be less reliable and it is recommended that this form of modelling may be less appropriate for critical decisions regarding key members of the target system.

OPEN-BOX VALIDATION

60. Open-box validation provides greater fidelity than black-box validation as it is concerned with understanding and representing the internal relationships and workings of a system that is being modelled.[11] Open-box modelling considers 'why' and 'how' a decision is made, but not 'what' that decision may be. It is a way of representing a decision-making process so that it accurately represents a real world situation. For example, if considering modelling the factors that make a farmer decide to cultivate opium poppy in Afghanistan, the black-box approach may be that the decision is made 40% of the time if it is sunny or 20% of the time if wet. In an open-box model however, the decision to cultivate could be some more like the equation in Figure 1. Open-box validation evaluates the accuracy of each constituent factor within this decision-making model.

[12] TURING, A.M. *Computing Machinery and Intelligence*. Mind, 1950, 49, 433-460.

$$d_c = f(w) - f(e) + f(r) + f(s) + f(p) + \ldots$$

Where d_c is the likelihood that the farmer decides to cultivate opium poppy;
f(w) is some function describing the influence of the weather;
f(e) is some function describing the likelihood of crop eradication;
f(r) is some function describing religious influences;
f(s) is some function describing social influences;
f(p) is some function describing the expected profit from cultivating.

Figure 1: Equation for Decision to Cultivate Opium Poppy

61. The main strength of this approach is that it does attempt to compare the internal workings of a model to the real world system that it has interpreted. Pragmatically, this may involve ensuring that the relationships programmed into the model make sense and seem plausible to SMEs. This shares the advantages of the interpretative and constructivist approaches in that it does not necessarily accept the underlying assumptions and processes as givens, and seeks consultation and validation from appropriate experts.

62. The main drawback of this approach is that it is fundamentally difficult to validate something as complex as an open-box model and even SMEs may not understand why certain people or groups behave as they do. Although the open-box understanding of why a farmer decides to cultivate opium poppy may be more representative of a real world situation, it is far more complicated than a black-box representation. Open-box modelling requires a deep understanding of the system being modelled and sufficient data to feed each of the variables. Black-box modelling may be less appropriate for non-routine situations but it is quicker and easier than open-box modelling. It should therefore be carefully considered whether an open-box model is required, whether it can actually be created, and whether the complex validation process is practicably feasible.

CONCEPTUO-LOGICAL VALIDATION

63. This approach encompasses concepts and validation techniques from various different theories and approaches. They have been grouped together under one heading as they incorporate components of a pragmatic approach to validation, considering how a concept is translated from the thoughts contained in people's minds to tools, techniques or models.[10] Conceptual validation is based on the acceptance that a conceptual model is an expression of the mental model that someone possesses and is concerned with how appropriate it is to obtain and use mental models. Logical validation is concerned with the way in which a conceptual model is translated into a formal model.[10]

64. The main strength of this approach is that it addresses several important areas that require validation. The conceptual validation addresses issues such as the identification of stakeholders, the expression of goals, the perspective of the problem, and the use of constructs to make sense of elements of the problem situation. The establishment of a critical reference group (comprised of a range of key stakeholders) for a study can provide a means of guarding against bias by the investigator, the privileging of one form of data over another, and the favouring of a particular perspective. This critical reference group can be involved in the design of a study, from the sample through to the interview questions and can help ensure that

the resultant model is a more accurate interpretation of any target socio-technical system, patter of behaviour or real world situation.[13]

65. The logical validation considers whether the modelling has been undertaken correctly and faithfully. This involves validating elements such as the mathematical equations used or the translation of concepts into algorithms and code. Logical validation also checks that the formal model captures the entirety and richness of the conceptual model. This involves evaluating the impact of any simplifying assumptions that may have to be made, and making the model user aware of this. It also involves evaluating any technical problems in the model, such as the internal coherence in terms of theorems, axioms and logic. Although introduced here (and in various literature sources) as logical validation, the process described is also known as verification.[13]

66. This approach validates through the process of reflexivity. This is when a research programme, the findings and the interpretation are subjected to continual critical evaluation. Reflexivity includes checking that an interpretation fits with all the original data, ensuring that topics are raised in all interviews and establishing what results can be generalised to a wider population.[13] It also involves checking modelling activities with peers and having regular review meetings. The purpose of these activities is to be a common sense check on modelling work, a second opinion on whether the modeller's conceptual model has been correctly represented and a check on whether transformations within the model have been correctly coded. The process of a modeller talking through their model with a third party can help to ensure that it is logical, transparent and sensible. The main drawback with this approach is that reflexivity does not necessarily confirm the validity of a model, particularly if both the modeller and the peers share similar underlying assumptions about human and social behaviour.

CONTINUOUS VALIDATION

67. Continuous validation is an overarching approach containing two elements that have been combined under one heading as each has the feature of being undertaken by the investigator in a continual, rolling manner. The first element is to take any opportunities to discuss with relevant people the work and validation currently underway and the results produced. This allows for maximal peer and SME review. It also increases the chance of serendipitous validation, such as discovering some other research using the same tool or case study, which would allow comparisons to be made. The second element can be loosely described as critical self review. The modeller should regularly take a step back from their modelling to critically consider what they are doing. This involves considering whether the work is answering the original question asked, whether it is the best or most suitable approach and if it is actually working. The two elements are linked as, in order to discuss modelling activities with third parties, some high level understanding of the purpose and progress of the activity is required.[13]

68. The main strength of continuous validation is that, because the research is continually reviewed, refined and updated, it paves the way for validity throughout. This approach involves a process of reflexivity similar to that in conceptuo-logical validation. However, with continuous validation this is undertaken throughout the research process, as opposed to post

[13] PYETT, P., M., 'Validation of Qualitative Research in the 'Real World'.' *Qualitative Health Research*, 2003, 13, (8), pp., 1170-1179.

hoc procedures that are applied or undertaken retrospectively.[14] Throughout the process of continuous validation, the responsibility for reliability and validity rests with the modeller, rather than being placed with external assessors. Continuous validation is essentially a term for a self-correcting mechanism to ensure the quality of the project.

69. The main weakness of this approach is that research and validation are only as good as the investigator. The investigator is required to be creative, sensitive, flexible and observant throughout the study to facilitate the development of appropriate validation of any model or technique developed. They must assess the data, processes, evidence and findings throughout to ensure that they remain pertinent to the study. For example, if research outputs lead to an initial theory or conjecture to be falsified, an investigator must be able to recognise this, accept it, and reject their initial thoughts, no matter how personally attached to them they may have become. They must then devise an alternative approach. If an investigator has this sole responsibility of ensuring a model is validated, the rigor of the validation process is entirely dependent on their ability to conduct this critical self-review effectively and efficiently.

VALIDATION BASED ON HISTORICAL APPLICATIONS

70. This approach uses the outputs, rather than the technique itself, for validation purposes. If a model or process led to recommendations that were implemented, it is possible to evaluate whether this implementation was deemed to be successful. Success or failure of a decision often cannot be attributed to the insights from a model, and the making of a specific decision may also not be wholly accredited to the model, but the fact that a recommendation was made and implemented, regardless of the reasons why, allows for some form of validation. If it was successful, then the model or process may be partially validated for that particular set of circumstances and type of implementation. Conversely, if unsuccessful then the model or process may have limitations set on where it should be used. This is similar to the instrumental approach except that it focuses primarily on previous historical applications of a model for different subjects.

71. The main strengths of this approach are that it can determine where a tool, technique or model has been used in the past, assess the degree of success it achieved, and use this to develop guidelines on when they should or should not be used again. Even if the decision made was not ideal, this process can make validation a quicker process allowing more time for consideration of other additional factors or other problems.

72. There are two key weaknesses with this approach. Firstly it does not validate the underlying assumptions. When considering human or social behaviour a particular hypothesis used for a specific subject may not be appropriate for another, even if the recommendations had been useful in this previous case. Secondly, it may be difficult to measure whether a decision made with the assistance of some tool or method is any better than the decision that would have been made without it. This approach doesn't fully account for the role that the tool or model played in the decision-making process as it only correlates the use of that tool with a previous case where a particular decision was taken.

[14] MORSE, J., M., BARRETT, M., MAYAN, M., OLSON, K., and SPIERS, J. 'Verification Strategies for Establishing Reliability and Validity in Qualitative Research.' International Journal of Qualitative Methods, 2002, 1, (2).

SUMMARY OF APPROACHES

73. All these approaches each offer an alternative to the realist approach to validation. Table 2 provides a short summary of each approach, with an indication of when it is appropriate to apply them, and the validation checks that are associated with their implementation. Many of these validation checks satisfy the validation assessments implied by the philosophical and theoretical perspectives, which were highlighted previously.

Table 2: Summary of Implementation of Validation Techniques

Validation Approach	When to Implement	Associated Validation Checks
Realist	Where historical data and/or outputs from experiments or trials exists.	Are the model outputs comparable to historical data, trials or experiments?
Interpretative	Where subjects have been modelled in detail	Have the subjects being modelled been consulted to determine the reasons for their actions? If subjects cannot be accessed, have suitable SMEs been consulted instead?
Constituent elements	Upon all models.	Do SMEs believe that relationships represented in models are reasonable? Do they believe data sources used are reliable? Do they believe that supporting information is reliable?
Constructivist	Upon all models.	Is there a community of people interested in the model? Are the model outputs accepted? Does the understanding / recommendations produce add value? Are the outputs useful?
Axiomatic	Where specific theories have been used within models.	Do SMEs agree that the model meets the axioms of the underlying theory?
Falsification	Upon underlying hypothesis of all models	Can the hypothesis be falsified?
Instrumentalist	Upon all models.	Does the user community believe the outputs to be useful? Do they also believe them to be usable?
Operational	On prescriptive models i.e. those which produce recommendations/ solutions.	Are the outputs produced usable and useful? Are they synergetic? Can they be applied in a timely and cost effective manner?
Experimental	Upon all models.	Do end users think the outputs are of a high quality e.g. increase insight, usable etc.
Aptness	Where there is a good relationship between modellers and end users.	From the perspective of the end user, does the model have insight generation capability, descriptive realism, mode reproduction, transparency, realism, ease of enrichment, fertility, formal correspondence with data and point prediction capability?
Black-box	Where outputs are on 'routine' situations and where non-critical decisions will be made on peripheral actors within a target system.	Do the input: outputs relationships within the model reflect reality? Will critical decisions involving key actors within the target system be made based on the outputs of this model? Will these decisions be on routine or non-routine situations?
Open-box	When there is a deep understanding of the central actor's decision-making processes.	Do the relationships programmed into the model make sense and ring true to SMEs?
Conceptuo-Logical	Upon all models.	Has the model had regular peer review? Have there been more formal review meetings?
Continuous	Upon all models.	Have relevant third parties been made aware of modelling and outputs? Has the modeller regularly undertaken critical self review?
Historical applications	When the technique has been used before.	Under what circumstances has the technique been used in the past and with what degree of success?
Identifying experts.	Where SME input is required.	Has the authenticity of the SME been confirmed?

IDENTIFYING EXPERTS FOR USE IN VALIDATION

74. A common theme within alternative approaches to validation is the use of expert advice and judgement. When searching for advice on 'best practice' on this topic, the response is frequently to use SMEs. Dependence upon SME input makes identifying a suitable SME an important component for validation. Before an SME is used, that individual must be identified and contact made with them. An effective means of achieving this is through utilisation of

existing contacts. The authenticity of the SME must also be established as some people profess to be experts when this status may be debatable. Simple checks of authenticity include the following:

a. Is the information being provided intuitively credible?

b. Is the information being provided consistent with other sources, such as the media or other experts?

c. Does the SME have a good reputation or do they come highly recommended?

d. Is their expertise based upon experience; does their historical record indicate that they should be providing expert advice?

e. Has their expertise been utilised in the past?

75. These checks may not all be appropriate and in some cases may not work at all (e.g. a SME may quite legitimately contradict the media), but the heavy reliance upon SME input dictates that sources should be carefully selected and authenticated. Research has already considered this issue and other organisations consider that authenticating experts is of paramount importance when validating models.

RELATIONSHIP WITH DG(S&A) GUIDELINES

76. DG(S&A) have produced guidelines for the validation and verification of models, some of which are appropriate to models of human behaviour. These guidelines from the perspective of alternative approaches to validating are summarised at Table 3. The feasibility of using each guideline is considered, comments made on whether and how it could be applied in the context of this paper and alternatives suggested where the extant guidance cannot be applied.

Table 3: Relationship with DG(S&A) Guidelines

DG(S&A) Guidelines	Feasible within the context of this paper?	Comments on application	Alternatives where extant guidance is not appropriate
Verification			
Management of specification	Yes	The purpose of the model should be kept in mind throughout development, meaning that what it is intended to achieve is known and documented. For example, if the model is to predict weather then the specification would include aspects such as wind speed and temperature. Although not as simple as this with human modelling, the modeller should know what they are trying to model and why, meaning that there should be some kind of specification for the model.	
Testing of components as they are added	No	Difficult, as the behaviour of the component on its own is not independent of the system as a whole when modelling human behaviour. For example, modelling the behaviour of a servant as an individual may be different to modelling their behaviour when in the presence of their master.	Each component described in the specification should be included in the model and simple checks of this undertaken.
Testing of model as components are added	No		As each component is added to the model, it should be checked that the model behaves as expected with the constituent elements. Although the behaviour of the system may not represent the real world when incomplete, checks for unexpected, inaccurate or unexplainable outputs could be carried out each time a component is added. It should be noted however that some systems are not built up on a component by component basis, in which case testing in this manner would not be possible.
Documented test plan	No	This covers the testing procedure and thus cannot be undertaken if formal testing is not possible.	Documentation of the model behaviour as additional components are added should be kept. By using fixed random numbers, it should be possible to track output changes resulting from each component addition.

DG(S&A) Guidelines	Feasible within the context of this paper?	Comments on application	Alternatives where extant guidance is not appropriate
Validation			
SME review	Yes	Concentrating on the workings of the model	
Military review	As appropriate	Concentrating on tactics and timescales - may not always be relevant	
Independent review	Yes	Should include testing whether changes in input variables produce sensible changes in output. Should also include comparison with trials, exercises or history. Comparison with a real world event or run from another model should also be conducted.	
Comparison with trials or exercises	Yes	Exercises may be useful in calibrating and understanding aspects of human models. Pre-planning of data collection is essential.	
Historical analysis	Partial	There is not always suitable historical data available to allow a comparison between historical events and model outputs.	HA may be able to support validation of specific components of the model, for example behaviour under stress. The limitations and caveats of data should be clearly stated and understood however.
Data dictionary	Yes	A record of data used in the model, its source and perceived limits on accuracy. In the case of human modelling, the source of data may be a theory rather than empirical observation or trial/exercise output.	
Audit trail of data	Yes	Builds on the data dictionary, but includes information on data history and quality.	
Configuration control	Yes	Ensures that only the most up to date data is used and that only authorised modifications to the data are made.	
Level 1 validation	Yes	Validation by review. The model data and model management have undergone full review. Areas where it cannot be used and areas where there are risks associated with use have been identified and documented.	
Level 2 validation	Partial	Validation based on real events. As with level 1 but with validation against trials, exercises or historical events having been undertaken.	Depends upon the scenario being modelled and the approach adopted.
Model V&V log books	Yes		

CONCLUSIONS

77. Validation is a method of increasing confidence in the accuracy and utility of information derived from a model. The realist approach to validation is to determine how accurately a model represents a real world situation by comparing model outputs with behaviour in that situation in reality. However, this approach may not be appropriate when trying to validate tools, techniques and models for understanding and predicting human and social behaviour. This is because defining "reality" is more complex in the social sciences than in the hard sciences and outputs of models concerned with human and social behaviour are often intended to facilitate understanding of a system rather than to definitively predict a future outcome.

78. This working paper introduced a number of problems with the realist approach that are raised by theoretical and philosophical perspectives on the possibility of explaining or understanding a social "reality". These perspectives challenge the objectivity of our knowledge of society and human behaviour. They argue that "truth" and social "reality" may not be indisputable fact, but may actually be an investigator's own interpretation that is affected by personal perceptions or bias. This "truth" may often be the result of research conducted within a certain cultural environment, adhering to inherent assumptions and directed towards achieving a certain goal. Furthermore, alternative theories of human and social behaviour may often be dismissed or disregarded because they do not meet the general consensus of what is considered to be "true" or "scientific" by fellow social scientists at that time.

79. These problems have implications for the development of a validation process that is more appropriate for the tools, techniques and methods for understanding and predicting human and social behaviour. Both the problems and implications raised by these perspectives were summarised in Table 1. These implications might lead the validation process to query:

    a. Whether the source data for models, tools and assessments is reliable and appropriate.
    b. Whether an investigator's interpretation of human behaviour or a society is appropriate.
    c. What scientific paradigm is underlying any theories used for models.
    d. Whether an alternative scientific paradigm may have been more appropriate.
    e. Whether the theories used for models, tools and techniques are credible.
    f. Whether theories become invalid if/when scientific paradigms change over time.
    g. Why the key elements in theories have been selected and prioritised over others.
    h. What particular view or interpretation these key elements encourage and reinforce.
    i. How a model might differ if the source knowledge offered a different interpretation.
    j. How the source knowledge has been interpreted and incorporated into a model.

80. This paper then introduced a range of alternative approaches that might be more appropriate for validating the tools, techniques and methods used for understanding and predicting human and social behaviour. Each approach has associated validation checks and many of these satisfy the validation checks implied by the philosophical and theoretical perspectives. The validation checks associated with these alternative approaches were summarised in Table 2 and include checking whether:

a. The subject(s) being modelled agree with the interpretation of their actions.

b. SMEs have been consulted if it is inappropriate or impossible to access the subject(s).

c. The primary data sources and supporting information are reliable.

d. SMEs agree that models or assessments meet the axioms of an underlying theory.

e. The underlying hypotheses of models and assessments can be falsified.

f. The models or tools have been subject to regular peer and SME review.

g. The input / output relationships of a model appropriately interpret the subject matter.

h. The ouputs of a model offer an appropriate interpretation of the subject matter.

i. The outputs of a model are comparable with historical data or previous experiments.

j. The models actually generate useful, reliable and valuable insight.

81. In summary, the perspectives introduced by this paper argue that in the social sciences there can be no objective "truth". This means that models cannot understand and predict human behaviour with the same certainty that is possible in the hard sciences. Therefore, the task when investigating and modelling human and social behaviour is to reach an accepted interpretation that is credible with peers, SMEs, and even the subjects of the model themselves where possible. The validation checks implied by the philosophical and theoretical perspectives, and many of the validation checks associated with alternative approaches, are concerned with ensuring that a model has achieved this accepted interpretation.

82. This paper does not attempt to offer any conclusive evaluation the best way to develop a practical approach validation that draws on broader philosophical and validation perspectives than those that are predominantly practiced in the MoD, though that is the ultimate aim of this project. Instead, it is hoped that this working paper will stimulate thought and debate around these theoretical and philosophical perspectives, and the issue of alternative approaches to validation, which we can incorporate in the guidance we subsequently develop. As such, the authors welcome and encourage comments on the paper.

REFERENCES

1. ROY, B., 'Decision science or decision-aid science.' *European Journal of Operational Research*, 1993, 66, pp., 184-203.

2. DERY, R., LANDRY, M., and BANVILLE, C., 'Revisiting the issue of model validation in OR: An epistemological view'. *European Journal of Operational Research*, 1993, 66, pp., 168-183.

3. HOLLIS, M., *The Philosophy of Social Science: An Introduction*, Cambridge University Press, 2002, pp. 84-86.

4. SMITH, S., 'Reflectivist and constructivist approaches to international theory' in BAYLISS, J., and SMITH, S., (eds.), *The Globalisation of World Politics: An Introduction to International Relations*, Oxford University Press, 2001, pp., 224-49.

5. HOLLIS, M., *The Philosophy of Social Science: An Introduction*, Cambridge University Press, 1994, p. 147.

6. FROST, M., *Ethics in International Relations*, Cambridge University Press, 1996, pp. 23-28.

7. MISER, H., J., 'A foundational concept of science appropriate for validation in operational research.' European Journal of Operational Research, 1993, 66, pp., 204-215.

8. ORAL, M., and KETTANI, O., 'The facets of the modelling and validation process in operations research.' European Journal of Operational Research, 1993, 66, pp., 216-234

9. PIDD, M. *Tools for Thinking*, John Wiley & Sons, Inc, 1996.

10. TURING, A., M., *Computing Machinery and Intelligence*. Mind, 1950, 49, 433-460.

11. PYETT, P., M., 'Validation of Qualitative Research in the 'Real World'.' Qualitative Health Research, 2003, 13, (8), pp., 1170-1179.

12. MORSE, J., M., BARRETT, M., MAYAN, M., OLSON, K., and SPIERS, J. 'Verification Strategies for Establishing Reliability and Validity in Qualitative Research.' International Journal of Qualitative Methods, 2002, 1, (2).